CLASSICAL FIELD THEORY
- A Quick Guide -

Huan Q. Bui

Colby College

PHYSICS & MATHEMATICS
Statistics

Class of 2021

June 25, 2019

Preface

Greetings,

Classical Field Theory, A Quick Guide to is compiled based on my independent study PH492: Topics in Classical Field Theory notes with professor Robert Bluhm. Sean Carroll's *Spacetime and Geometry: An Introduction to General Relativity*, along with other resources, serves as the main guiding text. This text also references a number of other texts such as *Quantum Field Theory* by Ryder, *Quantum Field Theory* by Mandl and Shaw, *Gauge Theories Of The Strong, Weak, And Electromagnetic Interactions* by Quigg, *A First Book of Quantum Field Theory* by A. Lahiri and P. B. Pal.

This text is a continuation of *General Relativity and Cosmology, A Quick Guide to*. Familiarity with classical mechanics, linear algebra, vector calculus, and especially general relativity is expected. There will be a quick review of general relativity where important concepts are revisited and important derivations highlighted, but familiarity with basic notions such as geodesics, Christoffel symbols, the Riemann curvature tensors, etc. is assumed.

Note: As a consequence of being developed from a variety of sources, there will be some overlapping among the sections. However, the objective of each section is unique.

Enjoy!

Contents

Chapter 1

Introduction to the Lagrangian and the Principle of Least Action

Proposition 1.0.1. All fundamental physics obeys least action principles.

The action S is defined as

$$S = \int_a^b \mathcal{L} \, dt. \tag{1.0.1}$$

where $\mathcal{L}$ is called the Lagrangian.

Refer for Farlow's *Partial Differential Equation*, page 353, for detailed explanation of Lagrange's calculus of variations.

I will derive the Euler-Lagrange equation(s) here, but we are not going to use it in the following subsection for the introduction to field theory for now.

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = 0. \tag{1.0.2}$$

1.1 A Classical-Mechanical Example

In this subsection we take a look at how the Lagrangian formulation of classical mechanics can give rise to Newton's second law of motion. In mechanics, the

Lagrangian often takes the form:

$$\mathcal{L} = K - V, \tag{1.1.1}$$

where K is the kinetic energy, and V is the potential energy. Let us consider a simple example where

$$K = \frac{1}{2}m\dot{x}^2 \tag{1.1.2}$$

$$V = V(x). \tag{1.1.3}$$

Variations on the Lagrangian gives

$$\delta\mathcal{L} = \delta\left(\frac{1}{2}m\dot{x}^2 - V(x)\right) \tag{1.1.4}$$

$$= m\dot{x}\delta\dot{x} - \frac{dV}{dx}\delta x \tag{1.1.5}$$

$$= m\dot{x}\dot{\delta x} - \frac{dV}{dx}\delta x \tag{1.1.6}$$

$$= m\left(-\ddot{x}\delta x + \frac{d}{dt}\dot{x}\delta x\right) - \frac{dV}{dx}\delta x \tag{1.1.7}$$

$$= -m\ddot{x}\delta x - m\frac{d}{dt}\dot{x}\delta x - \frac{dV}{dx}\delta x. \tag{1.1.8}$$

It follows that the variations on the action gives

$$\delta S = \int_a^b \delta L\, dt = -\int_a^b \left(m\ddot{x} + \frac{dV}{dx}\right)\delta x\, dt. \tag{1.1.9}$$

The principle of least action requires $\delta S = 0$ for all δx. Therefore it follows that

$$m\ddot{x} + \frac{dV}{dx} = 0, \tag{1.1.10}$$

which is simply Newton's second law of motion in disguise.

Before we move on, we should note that in order for the Lagrangian formulation to work in electromagnetism or in general relativity, we need to promote the Lagrangian to its relativistic version where the Lagrangian is given by

$$L = \int_a^b \mathcal{L}\, d^3x. \tag{1.1.11}$$

$\mathcal{L}$ is called the Lagrangian density, but we can colloquially refer to it as "the Lagrangian." The relativistic action hence takes the form

$$S = \int \mathcal{L}\, d^4x, \tag{1.1.12}$$

where d^4x implies integrating over all spacetime.

Chapter 2

Group Theory: a quick guide in a quick guide

1. Consider an N-dimensional vector with complex elements. A gauge transformation that takes $|z|^2 \to |z|^2$ (modulus preserving) is call a U(N) gauge transformation. The letter "U" denotes "unitary." These kinds of transformations can be represented by a unitary matrix, which is defined as a matrix whose conjugate transpose is the same as its inverse. If $z \to Uz$, then

$$|z|^2 \to (Uz)^\dagger(Uz) = z^\dagger U^\dagger U z = z^\dagger z = |z|^2, \qquad (2.0.1)$$

hence modulus preserving.

Note that the transformation $e^{i\alpha}$ is unitary. It is simply a 1×1 unitary matrix. This immediately implies that complex scalars have a U(1) gauge invariance.

The special group of U(N) transformations is denoted SU(N), which represents those with $\det U = 1$.

2. Consider another N-dimensional real-valued-element vector. We denote the group of orthogonal transformations O(N). Orthogonal transformations are orthogonal, i.e., length-preserving.

$$(Ox)^\top(Ox) = x^\top O^\top O x = x^\top x = \vec{x}^2 \qquad (2.0.2)$$

Once again, we denote the special group of the O(N) the SO(N) group. This group represents transformations with $\det O = 1$. The Lorentz group is a sub-group of SO(N), as the norm is defined differently, nevertheless it is still length-preserving. We call the Lorentz group SO(3,1), signifying that one sign is different from the other three.

Chapter 3

Introduction to Classical Field Theory

3.1 Relativistic Notation

Throughout this text, we will use the particle physics' Minkowski spacetime metric tensor:

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{3.1.1}$$

We shall assume knowledge of inner products and index-lowering/raising operations (please refer to *General Relativity & Cosmology: A Quick Guide to*) for more details about this "indexing" business.

Quick notes: the four-dimensional generalization of the gradient operator ∇ transforms like a four-vector. If $\phi(x)$ is a scalar function, so is

$$\delta\phi = \frac{\partial\phi}{\partial x^\mu}\delta x^\mu \tag{3.1.2}$$

and so

$$\frac{\partial\phi}{\partial x^\mu} \equiv \partial_\mu\phi \equiv \phi_{,\mu}\,. \tag{3.1.3}$$

Similarly, we also have something for the contravariant four-vector:

$$\frac{\partial\phi}{\partial x_\mu} \equiv \partial^\mu\phi \equiv \phi_{,}{}^\mu\,. \tag{3.1.4}$$

Lastly, we define the **d'Alembertian** as:

$$\Box \equiv \frac{\partial^2}{\partial t^2} - \partial^\mu \partial_\mu. \tag{3.1.5}$$

These will come in handy when we "vary" the action and the Lagrangian - but more on this later.

3.2   Classical Lagrangian Field Theory

In classical field theory, we are interested in fields, or systems of fields. We can consider a system which requires several fields, $\phi_r(x)$, where $r = 1, \ldots, N$. We are also interested in the Lagrangian density, which has a more general form than we have seen in the first section when we are still in the realm of Newtonian physics:

$$\mathcal{L} = \mathcal{L}(\phi_r, \phi_{r,a}), \tag{3.2.1}$$

where we have used the covariant notation to denote derivatives of ϕ with respect to the coordinates. The action, as integrated over some region Ω of four-spacetime, is then given by

$$S(\Omega) = \int_\Omega \mathcal{L}(\phi, \phi_{r,a}) \, d^4x. \tag{3.2.2}$$

Now, consider a variation in the field

$$\phi_r(x) \rightarrow \phi_r(x) + \delta\phi_r(x) \tag{3.2.3}$$

with the requirement that

$$\delta\phi_r(x)0 \tag{3.2.4}$$

on the boundary of Ω. As we will explore later on, the field ϕ can be real or complex. In the case of the complex field, we can simply treat ϕ and its complex conjugate ϕ^* as two independent fields. However, whether ϕ is complex or not should not matter in what we are doing now, which is requiring that the variation in the action to take a stationary value, i.e.,

$$\delta S = 0. \tag{3.2.5}$$

So,

$$0 = \delta S = \int_\Omega d^4x \left\{ \frac{\partial \mathcal{L}}{\partial \phi_r} \delta\phi_r + \frac{\partial \mathcal{L}}{\partial \phi_{r,a}} \delta\phi_{r,a} \right\} \tag{3.2.6}$$

$$= \int_\Omega d^4x \left\{ \frac{\partial \mathcal{L}}{\partial \phi_r} - \frac{\partial}{\partial x^a} \left(\frac{\partial \mathcal{L}}{\partial \phi_{r,a}} \right) \right\} \delta\phi_r + \int_\Omega d^4x \, \frac{\partial}{\partial x^a} \left(\frac{\partial \mathcal{L}}{\partial \phi_{r,a}} \delta\phi_r \right) \tag{3.2.7}$$

where the second line simply comes from doing integration by parts and the covariant notation:

$$\delta\phi_{r,a} = \frac{\partial}{\partial x^a}\delta\phi_r. \tag{3.2.8}$$

Now, the second term on the second line can be re-written under Gauss-Ostrogradsky's theorem as

$$\int_\Omega d^4x \, \frac{\partial}{\partial x^a}\left(\frac{\partial\mathcal{L}}{\partial\phi_{r,a}}\partial\phi_r\right) = \int_{\delta\Omega} d^3x \, \frac{\partial\mathcal{L}}{\partial\phi_{r,a}}\partial\phi_r \tag{3.2.9}$$

where the ∂_a denotes a divergence and $\delta\Omega$ denotes the boundary of the region in spacetime. But notice that since we require the variation to vanish at the boundary of Ω, this integral simply vanishes, so the variation in the action becomes

$$0 = \delta S = \int_\Omega d^4x \left\{\frac{\partial\mathcal{L}}{\partial\phi_r} - \frac{\partial}{\partial x^a}\left(\frac{\partial\mathcal{L}}{\partial\phi_{r,a}}\right)\right\}\delta\phi_r. \tag{3.2.10}$$

Now, because this equality has to hold for any variation $\delta\phi_r$, it must be true that the integrand has to be zero. We arrive at the Euler-Lagrange equation:

$$\frac{\partial\mathcal{L}}{\partial\phi_r} - \frac{\partial}{\partial x^a}\left(\frac{\partial\mathcal{L}}{\partial\phi_{r,a}}\right) = 0. \tag{3.2.11}$$

3.3 Quantized Lagrangian Field Theory (primer)

While the focus of this text is classical field theory, it can be worthwhile to have a little primer on what quantized Lagrangian field theory looks like. The main difference between classical and quantized Lagrangian field theory is the "degree of freedom" in the field. In classical field theory, we deal with systems with a continuously infinite number of degrees of freedom, i.e. at any point in spacetime, the field ϕ is assigned some value. In quantized field theory, we consider an "approximation" of this by imagining a system with a countable number of degrees of freedom, and then go to the continuum limit.

Consider flat spacetime, at an instance $t = Constant$. Let us now make space discrete by thinking of space as "cells" of equal volume called δx_i where i is just another index. Let the value of the field within each cell by the value of the field at the center of the cell. So now the system is discrete and can be described by a discrete set of generalized coordinates:

$$q_{ri}(t) \equiv \phi_r(\mathbf{x}_i, t) \tag{3.3.1}$$

where $q_{ri}(t)$ simply represents the value of the field ϕ_r at cell $\mathbf{x}_i$ at time t. Now, since the idea of the derivative no longer holds in the discrete world, the new

Lagrangian density takes into account the difference in the field values between neighboring cells. This gives the *Lagrangian, L*, NOT the *Lagrangian density* (the Lagrangian density is denoted $\mathcal{L}$):

$$L(t) = \sum_i \delta\mathbf{x}_i \mathcal{L}_i(\phi_r(i,t), \dot{\phi}_r(i,t), \phi_r(i',t)) \tag{3.3.2}$$

Now, recall from classical mechanics that the second term in the Euler-Lagrange equation represents the momentum, so we define the **momentum conjugate** to q_{ri} as

$$p_{ri} = \frac{\partial L}{\partial \dot{q}_{ri}} \tag{3.3.3}$$

$$= \frac{\partial L}{\partial \dot{\phi}_r(i,t)} \tag{3.3.4}$$

$$= \frac{\partial}{\partial \dot{\phi}_r(i,t)} \sum_i \delta\mathbf{x}_i \mathcal{L}_i \tag{3.3.5}$$

$$= \pi_r(i,t)\delta\mathbf{x}_i \tag{3.3.6}$$

where

$$\pi_r(i,t) = \frac{\partial \mathcal{L}_i}{\partial \dot{q}_r(i,t)} = \frac{\partial \mathcal{L}_i}{\partial \dot{\phi}_r(i,t)} \tag{3.3.7}$$

is called the **canonical momenta**. This definition makes sense by analogy to classical mechanics, as q_{ri} is just the coordinates. So, the Hamiltonian of the discrete system is given by

$$H = \sum_i p_i \dot{q}_{ri} - L \tag{3.3.8}$$

$$= \sum_i \delta\mathbf{x}_i \left\{ \pi_r(i,t)\dot{\phi}_r(i,t) - \mathcal{L}_i \right\}. \tag{3.3.9}$$

Now, letting the difference go to zero, i.e, $\delta\mathbf{x}_i \to 0$, which makes

$$\pi_r(x) = \frac{\partial \mathcal{L}_i}{\partial \dot{\phi}_r(i,t)} \to \frac{\partial \mathcal{L}_i}{\partial \dot{\phi}_r}, \tag{3.3.10}$$

i.e.,

$$\pi_r(i,t) \to \pi_r(\mathbf{x}_i,t), \tag{3.3.11}$$

in which case the Lagrangian turns from a sum into an integral

$$L(t) = \int d^3x \, \mathcal{L}(\phi_r, \phi_{r,a}) \tag{3.3.12}$$

and so does the Hamiltonian

$$H = \int d^3x \, \mathcal{H}(x), \tag{3.3.13}$$

where the Hamiltonian density is given by (again, turned from a sum into an integral)

$$\mathcal{H}(x) = \pi_r(x)\dot{\phi}_r(x) - \mathcal{L}(\phi_r, \phi_{r,a}). \tag{3.3.14}$$

3.4 Symmetries and Conservation Laws (primer)

Recall the Heisenberg's equation of motion of an operator $O(t)$:

$$i\hbar\frac{dO(t)}{dt} = [O(t), H], \tag{3.4.1}$$

where $[.,.]$ denotes the commutator. Note that

$$[O, H] = 0 \tag{3.4.2}$$

if O is a constant of motion, which often stem from invariance properties of systems under groups of transformation such as translation or rotations. These invariances relate to conservation laws, which we will explore in much more detail in the next sections. Consider an example with a wavefunction Ψ and an operator O undergoing a Lorentz transformation:

$$|\Psi\rangle \to |\Psi'\rangle = U|\Psi\rangle \tag{3.4.3}$$

$$O \to O' = UOU^\dagger. \tag{3.4.4}$$

U is a unitary transformation, which ensure that (i) the operator equations are covariant, and (ii) probability amplitudes and eigenvalues of operators (observables) are invariant (unitary matrices are length-preserving).

We can consider another example of a rotation in the complex plane:

$$U = e^{i\alpha T} \tag{3.4.5}$$

where U can be thought of as a unitary, continuous, 1×1 transformation, and $T = T^\dagger$. For an infinitesimal transformation

$$U \approx 1 + i\delta\alpha T \tag{3.4.6}$$

and hence

$$O' = O\delta O = UOU^\dagger = (1 + i\delta\alpha T)O(1 - i\delta\alpha T), \tag{3.4.7}$$

which we can show to be equivalent to

$$\delta O = i\delta\alpha[T, O]. \tag{3.4.8}$$

So, if the operator O is the Hamiltonian H, and by requiring the theory to be invariant under variations, i.e., $\delta H = 0$, which means

$$[T, H] = 0, \tag{3.4.9}$$

which implies that T is a constant of motion.

For a field theory derived from a Lagrangian density, we can construct conserved quantities from the invariance of $\mathcal{L}$ under symmetry transformations. We have a prescription to do this where very first step is to show that we can construct a function f^a of the field operators and derivatives such that

$$\frac{\partial f^a}{\partial x^a} = 0. \tag{3.4.10}$$

where a are just indices, and $x^0 \equiv t$. We will show how this is possible by first assuming that, well, it is possible - in which case if we define

$$F^a(t) = \int d^3x \, f^a(\mathbf{x}, t) = Constant. \tag{3.4.11}$$

(this equation really hints to us that something is conserved over time... hence a requirement to construct conserved quantities) then by $\partial f^a / \partial x^a = 0$ we have

$$\frac{dF^0}{dt} = \frac{d}{dt} \int d^3x \, f^0(\mathbf{x}, t) \tag{3.4.12}$$

$$= \frac{d}{dt} Constant - \frac{d}{dt} \partial_j f^j(\mathbf{x}, t), \quad j = 1, 2, 3 \tag{3.4.13}$$

$$= -\int d^3x \, \partial_j f^j(\mathbf{x}, t) \tag{3.4.14}$$

$$= 0. \tag{3.4.15}$$

This means

$$F^0 = \int d^3x \, f^0(\mathbf{x}, t) = Constant, \tag{3.4.16}$$

i.e, F^0 is a conserved quantity we are talking about. But what exactly are the quantities f and F and how can we interpret them? As we go along, it will make more sense to call f^a the *conserved current*. We will also explore Noether's theorem that relates transformation symmetries and conservations later in the following sections.

Let us apply this result to the transformation (or rather, the variation):

$$\phi_r(x) \rightarrow \phi_r'(x) = \phi_r(x) + \delta\phi_r(x). \tag{3.4.17}$$

This leads the variation of the Lagrangian to take the form

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi_r}\delta\phi_r + \frac{\partial\mathcal{L}}{\partial\phi_{r,a}}\delta\phi_{r,a} \tag{3.4.18}$$

$$= \frac{\partial\mathcal{L}}{\partial\phi_r}\delta\phi_r + \frac{\partial\mathcal{L}}{\partial_a(\partial\phi_r)}\delta\partial_a(\phi_r) \tag{3.4.19}$$

$$= \left(0 + \frac{\partial\mathcal{L}}{\partial(\partial_a\phi_r)}\right)\delta\phi_r + \frac{\partial\mathcal{L}}{\partial_a(\partial\phi_r)}\delta\partial_a(\phi_r) \tag{3.4.20}$$

$$= \delta\phi_r\partial_a\frac{\partial\mathcal{L}}{\partial(\partial_a\phi_r)} + \frac{\partial\mathcal{L}}{\partial(\partial_a\phi_r)}\partial_a\delta\phi_r \tag{3.4.21}$$

$$= \delta\phi_r\partial_a\frac{\partial\mathcal{L}}{\partial_a(\partial\phi_r)} + \frac{\partial\mathcal{L}}{\partial_a(\partial\phi_r)}\partial_a\delta\phi_r \tag{3.4.22}$$

$$= \frac{\partial}{\partial x^a}\left(\frac{\partial\mathcal{L}}{\partial_a(\partial\phi_r)}\delta\phi_r\right) \tag{3.4.23}$$

$$= \frac{\partial}{\partial x^a}\left(\frac{\partial\mathcal{L}}{\partial\phi_{r,a}}\delta\phi_r\right) \tag{3.4.24}$$

where the third line comes from the Euler-Lagrange equation. Now, since we require $\delta\mathcal{L} = 0$, in hopes of deriving a conserved quantity, we can match f to the last term in the equality:

$$f^a = \frac{\partial\mathcal{L}}{\partial\phi_{r,a}}\delta\phi_r = \frac{\partial\mathcal{L}}{\partial_a(\partial\phi_r)}\delta\phi_r, \tag{3.4.25}$$

in which case

$$F^a = \int d^3x\, f^a(x,t) = \int d^3x\, \frac{\partial\mathcal{L}}{\partial_a(\partial\phi_r)}\delta\phi_r, \tag{3.4.26}$$

so

$$F^0 = \int d^3x\, \frac{\partial\mathcal{L}}{\partial_0(\partial\phi_r)}\delta\phi_r \tag{3.4.27}$$

$$= \int d^3x\, \frac{\partial\mathcal{L}}{\partial\dot{\phi}_r}\delta\phi_r \tag{3.4.28}$$

$$= \int d^3x\, \pi_r(x)\delta\phi_r(x) \tag{3.4.29}$$

where π_r is the momentum as defined earlier.

Now, some interesting things can happen is our field is complex. Recall that we can treat a complex fields as two independent fields. Consider the following transformations:

$$\phi_r\, to\phi'_r = e^{i\epsilon}\phi_r \approx (1 + i\epsilon)\phi_r \tag{3.4.30}$$

$$\phi_r^\dagger \to \phi\,\dagger'_r\, e^{0ie\epsilon}\phi_r^\dagger \approx (1 - i\epsilon)\phi_r^\dagger \tag{3.4.31}$$

where ϵ is real, and suppose that $\mathcal{L}$ is invariant under this transformation. The transformations, assuming that ϵ is very small, give

$$\delta\phi_r = i\epsilon\phi_r \tag{3.4.32}$$

$$\delta\phi_r^\dagger = -i\epsilon\phi_r^\dagger. \tag{3.4.33}$$

If this is the case, then F^0 can be written as

$$F^0 = \int d^3x\,\pi_r(x)\delta\phi_r(x) = i\epsilon\int d^3x\,[\pi_r(x)\phi_r(x) - \pi_r^\dagger(x)\phi_r^\dagger(x)]. \tag{3.4.34}$$

Now, recall that F^0 is simply some constant. So if we define Q as an "educated" scalar multiplication of F^0, then Q should also be a conserved quantity just like F^0, which we have shown. So, let

$$Q = -\frac{iq}{\hbar}\int d^3x\,[\pi_r(x)\phi_r(x) - \pi_r^\dagger(x)\phi_r^\dagger(x)], \tag{3.4.35}$$

where q is just another constant. There are two reasons why we might want to do this. First, we would like to introduce q because we will ultimately show that it is the electric charge. And second, we want to introduce the term $i/\hbar$ because we can observe that the integrand looks suspiciously like it has something to do with a commutator. In fact, we will now evaluate the commutator $[Q, \phi_r(x)]$. However, there are important steps and concepts we need to understand before doing this.

First, recall the Heisenberg uncertainty principle can be written in commutator form

$$[\hat{x}, \hat{p}] = i\hbar \tag{3.4.36}$$

where $\hat{x}$ is just the position operator, and $\hat{p}$ is the momentum operator. Now, while we are not actually dealing with position and momentum operators, we are dealing the **canonical coordinate** ϕ_r and **canonical momenta** π_r, to which the canonical commutation relations still apply:

$$[\hat{x}_i, \hat{\pi}_j] = i\hbar\delta_j^i \tag{3.4.37}$$

$$[\hat{\pi}_i, \hat{\pi}_j] = [\hat{x}_i, \hat{x}_j] = 0 \tag{3.4.38}$$

Specifically in our discrete case,

$$[\phi_r(i,t), \pi_s(j',t)] = i\hbar\frac{\delta_s^r\delta_j^{j'}}{\delta\mathbf{x}_j} \tag{3.4.39}$$

$$[\phi_r(j,t), \phi_s(j',t)] = [\pi_r(j,t), \pi_s(j',t)] = 0. \tag{3.4.40}$$

In the continuous case:

$$[\phi_r(\mathbf{x},t), \pi_s(\mathbf{x}',t)] = i\hbar\delta_s^r\delta(\mathbf{x}-\mathbf{x}') \tag{3.4.41}$$

$$[\phi_r(\mathbf{x},t), \phi_s(\mathbf{x}',t)] = [\pi_r(\mathbf{x},t), \pi_s(\mathbf{x}',t)] = 0. \tag{3.4.42}$$

Now, back to our case. We want to find

$$[Q, \phi_r(x)]. \tag{3.4.43}$$

(more to come... this section in Mandl and Shaw is actually pretty long and I don't know what the "narrative" is. I'll come back to this later.)

Chapter 4

Gauge Invariance

4.1 Introduction

From studying Maxwell's equations, we know that there is more than one way to choose a vector potential A_μ that describes the same electromagnetic field. This freedom is called *gauge invariance*. We will explore the idea of gauge invariance in the context of a continuous symmetry of the Lagrangian, which leads to the conservation of electric charge and other important consequences under Noether's theorem.

   In the subsequent subsections, we will explore and try to understand the idea of gauge invariance. We will start out with gauge invariance in classical electrodynamics. Then, we will look briefly at phase invariance in quantum mechanics and ultimately field theory.

4.2 Gauge Invariance in Classical Electrodynamics

Recall the second of Maxwell's equations, written in differential form:

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0. \tag{4.2.1}$$

Following from a vector calculus identity, this means the magnetic field can be expressed as a curl of some vector potential, called $\mathbf{A}$:

$$\boldsymbol{\nabla} \cdot \mathbf{B} = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} \times \mathbf{A} = 0. \tag{4.2.2}$$

Now, we also know that the curl of a conservative vector field is zero, it is also true that

$$\boldsymbol{\nabla} \times (\mathbf{A} + \boldsymbol{\nabla}\Lambda) = \boldsymbol{\nabla} \times \mathbf{A} + \boldsymbol{\nabla} \times \boldsymbol{\nabla}\Lambda = \boldsymbol{\nabla} \times \mathbf{A}. \qquad (4.2.3)$$

It follows that this new vector potential $\mathbf{A} + \boldsymbol{\nabla}\Lambda$ also describes the same magnetic field $\mathbf{B}$:

$$\mathbf{B} = \boldsymbol{\nabla} \times (\mathbf{A} + \boldsymbol{\nabla}\Lambda) = \boldsymbol{\nabla} \times \mathbf{A}. \qquad (4.2.4)$$

Next, consider Faraday-Lenz law:

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \qquad (4.2.5)$$

Since $\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}$, this can be re-written as:

$$\boldsymbol{\nabla} \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right) = 0. \qquad (4.2.6)$$

This suggests that

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \qquad (4.2.7)$$

is some conservative vector field, which we shall identity as $-\boldsymbol{\nabla}V$, where V is the scalar potential:

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\boldsymbol{\nabla}V. \qquad (4.2.8)$$

Now, in order for the electric field $\mathbf{E}$ to remain invariant under the transformation

$$\mathbf{A} \to \mathbf{A} + \boldsymbol{\nabla}\Lambda, \qquad (4.2.9)$$

we must require that

$$\mathbf{E}' = \mathbf{E} \qquad (4.2.10)$$

$$-\boldsymbol{\nabla}V' - \frac{\partial \mathbf{A}'}{\partial t} = -\boldsymbol{\nabla}V - \frac{\partial \mathbf{A}}{\partial t} \qquad (4.2.11)$$

$$-\boldsymbol{\nabla}V' - \frac{\partial}{\partial t}(\mathbf{A} + \boldsymbol{\nabla}\Lambda) = -\boldsymbol{\nabla}V - \frac{\partial \mathbf{A}}{\partial t} \qquad (4.2.12)$$

$$\boldsymbol{\nabla}V' = \boldsymbol{\nabla}V + \frac{\partial}{\partial t}\boldsymbol{\nabla}\Lambda \qquad (4.2.13)$$

$$\boldsymbol{\nabla}V' = \boldsymbol{\nabla}V + \boldsymbol{\nabla}\left(\frac{\partial \Lambda}{\partial t}\right), \qquad (4.2.14)$$

i.e,

$$V' = V + \frac{\partial \Lambda}{\partial t}. \qquad (4.2.15)$$

So, the scalar potential has to undergo the transformation

$$V \rightarrow V + \frac{\partial \Lambda}{\partial t}. \tag{4.2.16}$$

Now, if we define the 4-vector potential as

$$A^\mu = (V, \mathbf{A}), \tag{4.2.17}$$

then everything we have done so far can be encoded in the 4-curl of A^μ, which we define as the anti-symmetric electromagnetic field strength tensor

$$F^{\mu\nu} = -F^{\nu\mu} = \partial^\mu A^\nu - \partial^\nu A^\mu = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & B_1 & 0 \end{pmatrix}. \tag{4.2.18}$$

We can verify that $F^{\mu\nu}$ is invariant under the transformation

$$A^\mu \rightarrow A^\mu - \partial^\mu \Lambda \tag{4.2.19}$$

or

$$A_\mu \rightarrow A_\mu + \partial_\mu \Lambda, \tag{4.2.20}$$

where $\Lambda(x)$ is an arbitrary scalar field of the coordinate. The fact that many different four-vector potentials yield the same electromagnetic fields, and thus describe the same physics, is a manifestation of the gauge invariance of classical electrodynamics.

From *General Relativity & Cosmology, A Quick Guide to...*, we have verified how the Maxwell's equations in covariant notation:

$$\partial_\mu F^{\mu\nu} = -J^\nu \tag{4.2.21}$$
$$\partial_\sigma F_{\mu\nu} + \partial_\mu F_{\nu\sigma} + \partial_\nu F_{\sigma\mu} = 0 \tag{4.2.22}$$

where the *4-current J^ν* is defined as

$$J^\nu = (\partial, \mathbf{J}) \tag{4.2.23}$$

give rise to the same Maxwell's equations in differential form. Now, we look at two important consequences of these two covariant equations.

First, consider the 4-divergence of the 4-current. We can readily from the definition of $F^{\mu\nu}$ that

$$\partial_\nu J^\nu = -\partial_\nu \partial_\mu F^{\mu\nu} \tag{4.2.24}$$
$$= -\partial_\nu \partial_\mu (\partial^\mu A^\nu - \partial^\nu A^\mu) \tag{4.2.25}$$
$$= -\partial_\nu \partial_\mu \partial^\mu A^\nu + \partial_\nu \partial_\mu \partial^\nu A^\mu) \tag{4.2.26}$$
$$= [-\partial_\nu \partial_\mu + \partial_\mu \partial_\nu] \partial^\mu A^\nu \tag{4.2.27}$$
$$= 0. \tag{4.2.28}$$

Note that operators $\partial_\nu \partial_\mu$ and $\partial_\mu \partial_\nu$ commute because they are just partial derivatives. This tells us that **the electromagnetic current is conserved**.

Second,

$$\partial_\mu F^{\mu\nu} = -J^\nu \qquad (4.2.29)$$

can be expanded as

$$\partial_\mu \partial^\mu A^\nu - \partial_\mu \partial^\nu A^\mu = J^\nu \qquad (4.2.30)$$

and re-written as

$$\Box A^\nu - \partial^\nu (\partial_\mu A^\mu) = J^\nu. \qquad (4.2.31)$$

In the absence of any sources, and in the Lorentz gauge where $\partial_\mu A^\mu = 0$, this reduces to

$$\Box A^\nu = 0. \qquad (4.2.32)$$

So A^ν satisfies the Klein-Gordon equation for a massless particle (photon).

4.3   Phase Invariance in Quantum Mechanics

It is in fact possible for us to guess Maxwell's equations from a gauge principle based on the Schrödinger equation, even in the absence of electrodynamics knowledge. Let a wavefunction ψ be given, recall that a quantum mechanical observable is of the form

$$\langle \mathcal{O} \rangle = \int \psi^* \mathcal{O} \psi \, d^n x, \qquad (4.3.1)$$

which can be readily verified to be invariant under the global phase rotation:

$$\psi(x) \to e^{i\theta} \psi(x). \qquad (4.3.2)$$

This result implies that there is no such thing as "the absolute phase." Rather, the only thing we can measure is the *relative phase* between wavefunctions, which is unaffected by global rotation.

Now, what about local rotation $e^{i\alpha(x)}$? Are we free to choose different phase conventions at different locations? Can quantum mechanics be formulated such that it is invariant under *local*, i.e., spatio-dependent phase rotations

$$\psi \to e^{i\alpha(x)} \psi? \qquad (4.3.3)$$

The answer turns out to be yes, and we shall explore how this is done. Consider the same local transformation. Since the Schrödinger equation involves the derivatives of ψ, we should consider how it transforms:

$$\partial_\mu \psi(x) \to \partial_\mu \psi' = \partial_\mu \left(e^{i\alpha(x)} \psi(x) \right) \tag{4.3.4}$$

$$= i \left(\partial_\mu \alpha(x) \right) \left(e^{i\alpha(x)} \psi(x) \right) + e^{i\alpha(x)} \partial_\mu \psi(x) \tag{4.3.5}$$

$$= e^{i\alpha(x)} \left[\partial_\mu \psi(x) + i(\partial_\mu \alpha(x)) \psi(x) \right]. \tag{4.3.6}$$

So there is an additional gradient term apart from just the phase change as we have seen before. The problem here is that our current notion of the derivation is not *gauge-covariant*, just like how the "normal" derivative does not work in curved spacetime and has to be replaced by covariant derivative in general relativity. The same thing is happening here, and the "fix" is also the same. We shall change our notion of the derivative and define a new, *gauge covariant* one as

$$D_\mu \equiv \partial_\mu + ieA_\mu \tag{4.3.7}$$

such that

$$D_\mu \psi(x) \to D_\mu \psi'(x) = e^{i\alpha(x)} D_\mu \psi(x). \tag{4.3.8}$$

To find how A_μ transforms (it should because it acts like the Christoffel symbol in general relativity, connecting the coordinate systems), we let D_μ act on $\psi(x)$ in their respective frames, and require that

$$D'_\mu \psi'(x') = e^{i\alpha(x)} D_\mu \psi(x). \tag{4.3.9}$$

First, we look at the rotated frame:

$$D'_\mu \psi'(x') = \left(\partial_\mu + ieA'_\mu \right) \psi(x) e^{i\alpha(x)} \tag{4.3.10}$$

$$= \partial_\mu \left(\psi(x) e^{i\alpha(x)} \right) + ieA'_\mu \psi(x) e^{i\alpha(x)} \tag{4.3.11}$$

$$= e^{i\alpha(x)} \partial_\mu \psi(x) + \psi(x) \left(\partial_\mu e^{i\alpha(x)} \right) + ieA'_\mu \psi(x) e^{i\alpha(x)} \tag{4.3.12}$$

$$= e^{i\alpha(x)} \left(\partial_\mu \psi(x) + i\psi(x) \partial_\mu \alpha(x) + ieA'_\mu \psi(x) \right). \tag{4.3.13}$$

Next, we look at the original frame:

$$D_\mu \psi(x) = \left(\partial_\mu + ieA_\mu \right) \psi(x) \tag{4.3.14}$$

$$= \partial_\mu \psi(x) + ieA_\mu \psi(x). \tag{4.3.15}$$

Since $D'\psi'(x') = e^{i\alpha(x)} D_\mu \psi(x)$,

$$\partial_\mu \psi(x) + ieA_\mu \psi(x) = \partial_\mu \psi(x) + i\psi(x)\partial_\mu \alpha(x) + ieA'_\mu \psi(x) \tag{4.3.16}$$

$$ieA_\mu = i\partial_\mu \alpha(x) + eA'_\mu. \tag{4.3.17}$$

Therefore, the transformation rule for A_μ must be

$$A'_\mu = A_\mu - \frac{1}{e}\partial_\mu\alpha(x). \tag{4.3.18}$$

This transformation rule has exactly the same form of a gauge transformation in electrodynamics, which we have discussed in the previous subsection

$$A_\mu \to A_\mu + \partial_\mu\Lambda. \tag{4.3.19}$$

In fact, A_μ is the electromagnetic field 4-vector. Now, while the form is the same, there is an additional coupling constant e, which is the charge in natural units of the particle described by $\psi(x)$. Therefore, the form of the coupling $D_\mu\psi(x)$ between the electromagnetic field and matter is suggested, if not uniquely dictated, by local phase invariance.

4.4   Significance of Potentials in Quantum Theory

In this subsection, we will look at the role of the vector potential A_μ in quantum mechanics, specifically quantum-mechanical interference phenomena. First, we will briefly discuss the Aharonov-Bohm effect. Second, we will show that the vector potential, though regarded as a purely mathematical device which lacks physical significance, does play a role in quantum mechanics. Third, we will learn a little bit about path-dependent phase factors in quantum mechanics.

4.4.1   The Aharonov-Bohm Effect & The Physical Vector Potential

The Ahanorov-Bohm effect shows that the vector potential is not just a mathematical construct to simplify calculations and that it does not have any physical significance. In 1959, Aharonov and Bohm proposed an experiment to resolve the question of the physical significance of the vector potential. The gist of the effect is that fact that wavefunctions of quantum mechanical objects acquire additional phase when traveling through space with no electromagnetic fields, only potentials. To show this, we first recall the Schröndinger equation for a free particle:

$$-\frac{\hbar^2}{2m}\nabla^2\psi^0 = i\hbar\frac{\partial\psi^0}{\partial t}. \tag{4.4.1}$$

In the presence of a vector potential, the Schrödinger equation becomes

$$\frac{(-i\hbar\nabla - e\mathbf{A})^2}{2m}\psi(x) = i\hbar\frac{\partial\psi(x)}{\partial t}. \tag{4.4.2}$$

The solution to this new Schrödinger is

$$\psi(\mathbf{x}, t) = \psi^0(\mathbf{x}, t) e^{\frac{iS}{\hbar}} \tag{4.4.3}$$

where

$$S = e \int \mathbf{A} \, d\mathbf{x} \tag{4.4.4}$$

is a line integral of $\mathbf{A}$ along the trajectory of the particle. Now, recall that

$$\boldsymbol{\nabla} \times \mathbf{A} = \mathbf{B}, \tag{4.4.5}$$

which means it is possible to have $\mathbf{A} \neq 0$ but $\mathbf{B} = 0$. So, if we allow particles to move in a potential, say around a vertical rod with a solenoidal $\mathbf{A}$, even without the magnetic field, there will be a phase difference between particles that go *along* $\mathbf{A}$ and those that go *against* $\mathbf{A}$. This creates a phase shift between the wavefunctions, giving rise to interference that can be shifted by turning on and off the magnetic field. We can argue this rigorously as follows. Let

$$\psi^0(\mathbf{x}, t) = \psi^0_1(\mathbf{x}, t) + \psi^0_2(\mathbf{x}, t) \tag{4.4.6}$$

represent the wavefucntion in the absence of a vector potential, where ψ^0_1 and ψ^0_2 denote the components of the beam that pass on the right or left of the solenoid: When we turn the current on, there is a magnetic field $\mathbf{B}$ confined
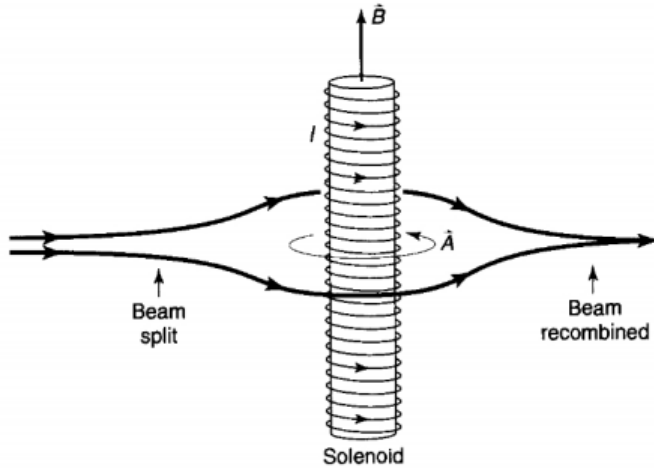
within the solenoid, i.e., there is no lateral magnetic field $\mathbf{B}$ on the side of the solenoid. So, according to the figure, electrons pass through a magnetic-field-free region before landing on the detector sheet. However, there exists a vector potential $\mathbf{A}$ outside of the solenoid, as the curl of $\mathbf{A}$ has to be such that the

produced magnetic field **B** is contained within the coil. This can be illustrated mathematically by Stokes' theorem:

$$\int_D \mathbf{B} \cdots d\sigma = \int_{\sigma D} \mathbf{A} \cdot d\mathbf{x}. \tag{4.4.7}$$

The perturbed wavefunction as a result of passing through this region is

$$\psi = \psi_1^0 e^{iS_1\hbar} + \psi_2^0 e^{iS_2/\hbar} \tag{4.4.8}$$

where

$$S_i = e \int_{\text{path } i} \mathbf{A} \cdot d\mathbf{x}. \tag{4.4.9}$$

It is clear now that the two components of the perturbed wavefunction has a phase difference of

$$\phi = \frac{S_1 - S_2}{\hbar} = \frac{e}{\hbar} \oint \mathbf{A}(\mathbf{x}) \cdot d\mathbf{x} = \frac{e}{\hbar} \oint A^\mu \, dx_\mu, \tag{4.4.10}$$

which is dependent on the vector potential **A**. Therefore, there is a physical effect from **A** even though the electrons do not experience any electromagnetic force.

4.4.2   Path-dependent Phase Factors

Results from the Aharonov-Bohm experiment showed that knowing the electromagnetic field strength tensor $F^{\mu\nu}$ is not enough to determine all electrodynamic phenomena in quantum mechanics. In fact, a phase factor of the form

$$e^{\frac{-ie}{\hbar} \oint A^\mu \, dx_\mu} \tag{4.4.11}$$

must be known in order to give correct predictions. Notice that the given integral is path-dependent line integral.

4.5   Phase Invariance in Field Theory

In the previous sections we have seen somewhat of the connection between electromagnetic gauge invariance, and have generalized global phase invariance to local phase symmetry. Now, we will bring these results into Lagrangian field theory. Detailed derivations will not be focused here, as the main idea of this section is to understand what gauge transformation and symmetry are. So, we will mainly focus on results, and only derivative only when necessary. Consider the Lagrangian for the free complex scalar field:

$$\mathcal{L} = |\partial^\mu \phi|^2 = m^2 |\psi|^2. \tag{4.5.1}$$

We will show that by the principle of least action ϕ and its complex conjugate satisfies the Klein-Gordon equation:

$$(\Box + m^2)\phi(x) = 0 \tag{4.5.2}$$

$$(\Box + m^2)\phi^*(x) = 0 \tag{4.5.3}$$

Next, consider a global phase shift:

$$\phi(x) \to e^{iq\alpha}\phi(x) \tag{4.5.4}$$

$$\phi^*(x) \to e^{-iq\alpha}\phi^*(x). \tag{4.5.5}$$

We can readily verify that the infinitesimal variations are

$$\delta\phi = iq(\delta\alpha)\phi \tag{4.5.6}$$

$$\delta(\partial_\mu\phi) = iq(\delta\alpha)\partial_\mu\phi \tag{4.5.7}$$

$$\delta\phi^* = -iq(\delta\alpha)\phi^* \tag{4.5.8}$$

$$\delta(\partial_\mu\phi^*) = -iq(\delta\alpha)\partial_\mu\phi^*. \tag{4.5.9}$$

Global phase invariance requires the Lagrangian to remain unchanged, i.e.,

$$\delta\mathcal{L} = 0, \tag{4.5.10}$$

i.e.,

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta(\partial_\mu\phi) + \frac{\partial\mathcal{L}}{\partial\phi^*}\delta\phi^* + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^*)}\delta(\partial_\mu\phi^*) \tag{4.5.11}$$

$$= \left[\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right]iq(\delta\alpha)\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}iq(\delta\alpha)\partial_\mu\phi - (\phi \to \phi^*) \tag{4.5.12}$$

$$= iq(\delta\alpha)\partial_\mu\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\phi - \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^*)}\phi^*\right] \tag{4.5.13}$$

$$\equiv 0. \tag{4.5.14}$$

We can identity a conserved Noether's current (we will discuss Noether's theorem later, but the point here is to show the existence of a conserved quantity) as:

$$J^\mu = -iq\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\phi - \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^*)}\phi^*\right] \tag{4.5.15}$$

$$= iq\left[\phi^*\partial_\mu\phi - (\partial_\mu\phi^*)\phi\right] \tag{4.5.16}$$

$$\equiv iq\phi^*\bar{\partial}^\mu\phi, \tag{4.5.17}$$

which satisfies

$$\partial_\mu J^\mu = 0. \tag{4.5.18}$$

So, we have shown the connection between global phase invariance and current conservation. But what if the transformation is local? Consider the following local phase rotation:

$$\phi(x) \to e^{iq\alpha(x)}\phi(x). \tag{4.5.19}$$

The same "problem" arises as earlier in the section where we have some extra terms in the gradient of the scalar field:

$$\partial^\mu\phi \to e^{iq\alpha(x)}\left[\partial_\mu\phi + iq(\partial_\mu\alpha(x))\phi\right], \tag{4.5.20}$$

which necessitates the introduction of the gauge-covariant derivative

$$D_\mu \equiv \partial_\mu + iqA_\mu \tag{4.5.21}$$

such that

$$D_\mu\phi \to e^{iq\alpha(x)}D_\mu\phi, \tag{4.5.22}$$

which demands the A_μ has to transform as

$$A_\mu(x) \to A_\mu(x) - \partial_\mu\alpha(x), \tag{4.5.23}$$

which we have verified earlier in the section. So, by requiring local phase symmetry (replacing ∂_μ to D_μ), we require some form of interaction between the gauge field (vector potential) A_μ and matter.

Chapter 5

Lagrangian Field Theory in Flat Spacetime

5.1 Real Scalar Fields

A scalar field can be used to describe particles of spin 0. A scalar field has only one component, or one degree of freedom, making it the "simplest case" of the fields we will discuss. Let us now consider a moving field in one dimension, which has the form

$$\phi(s) \sim e^{-i\mathbf{k}\cdot\mathbf{x}}, \tag{5.1.1}$$

where

$$\mathbf{k} = K^\mu = (K^0, \vec{K}) \tag{5.1.2}$$

$$\mathbf{x} = X^\mu = (X^0, \vec{X}). \tag{5.1.3}$$

Remember that K^μ is the wavenumber vector, and X^μ is the position vector. Also recall that the metric is Minkowskian at this point of consideration (we are still in flat spacetime. General curved spacetime will come later):

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{5.1.4}$$

Doing the inner product of X^μ and K^μ gives

$$\phi(x) = e^{-iK^0 t + i\vec{k}\cdot\vec{x}}. \tag{5.1.5}$$

We shall choose "natural units" such that $\hbar = c = 1$. This gives

$$\phi(x) = e^{-i\omega t} e^{i\vec{k}\cdot\vec{x}}. \tag{5.1.6}$$

Now, particles obey the following Einstein mass-energy equivalence:

$$E^2 = m^2 + \vec{p}^2. \tag{5.1.7}$$

But because of our choice of units, $E = c\hbar K^0 = K^0$, and $\vec{p} = \hbar\vec{k} = \vec{k}$. This gives

$$\left(K^0\right)^2 - \vec{k}^2 = m^2 \tag{5.1.8}$$

$$K^\mu K_\mu = m^2. \tag{5.1.9}$$

So, massive particles obey $K^\mu K_\mu = m^2$, while massless particles obey $K^\mu K_\mu = 0$.

Now, we might wonder how we know that the scalar field has the above form. The answer is derived from, you guessed it, the Lagrangian for a scalar field. Let us consider a single scalar field in classical mechanics where

$$\text{Kinetic energy: } K = \frac{1}{2}\dot{\phi}^2 \tag{5.1.10}$$

$$\text{Gradient energy: } G = \frac{1}{2}\left(\nabla\phi\right)^2 \tag{5.1.11}$$

$$\text{Potential energy: } P = V(\phi). \tag{5.1.12}$$

Note: I haven't found a satisfactory explanation to what a "gradient energy" is. I'll come back to this term later.

We currently have three terms, but we would like our Lagrangian density to have the form $\mathcal{L} = K - V$. So, let us combine the kinetic energy and gradient energy terms into one:

$$K' = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}\left(\nabla\phi\right)^2. \tag{5.1.13}$$

We shall verify that

$$K' = -\frac{1}{2}\left(\partial_\mu\phi\right)\left(\partial^\mu\phi\right) = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}\left(\nabla\phi\right)^2. \tag{5.1.14}$$

This turns out to be quite straightforward:

$$\left(\partial_\mu\phi\right)\left(\partial^\mu\phi\right) = \eta^{\mu\nu}\left(\partial_\mu\phi\right)\left(\partial_\nu\phi\right) \tag{5.1.15}$$

$$= \left(\partial_0\phi\right)^2 - \left(\partial_j\phi\right)^2 \tag{5.1.16}$$

$$= \dot{\phi}^2 - \left(\nabla\phi\right)^2. \tag{5.1.17}$$

So, a good choice of Lagrangian for our scalar field would be

$$\mathcal{L} \sim K' - V = -\frac{1}{2} \left(\partial_\mu \phi \right) \left(\partial^\mu \phi \right) - V(\phi). \tag{5.1.18}$$

In order for the action to be extremized, i.e. $\delta S = 0$, we require that $\delta \mathcal{L} = 0$ for any $\delta \phi$. Varying $\mathcal{L}$ with respect to ϕ gives

$$\delta \mathcal{L} = \delta \left(-\frac{1}{2} \left(\partial_\mu \phi \right) \left(\partial^\mu \phi \right) - V(\phi) \right) \tag{5.1.19}$$

$$= -\frac{1}{2} \left(\partial_\mu \delta \phi \, \partial^\mu \phi + \partial_\mu \phi \, \partial^\mu \delta \phi \right) - \frac{dV(\phi)}{d\phi} \delta \phi \tag{5.1.20}$$

$$= -\partial_\mu \delta \phi \, \partial^\mu \phi - \frac{dV(\phi)}{d\phi} \delta \phi. \tag{5.1.21}$$

Now, integration by parts tells us that

$$\partial_\mu \left(\partial^\mu \phi \, \delta \phi \right) = \partial^\mu \partial_\mu \phi \, \delta \phi + \partial_\mu \delta \phi \, \partial^\mu \phi. \tag{5.1.22}$$

So,

$$\partial_\mu \delta \phi \, \partial^\mu \phi = \partial_\mu \left(\partial^\mu \phi \, \delta \phi \right) - \partial^\mu \partial_\mu \phi \, \delta \phi. \tag{5.1.23}$$

Therefore, variations on $\mathcal{L}$ is:

$$\delta \mathcal{L} = - \left[\partial_\mu \left(\partial^\mu \phi \, \delta \phi \right) - \partial^\mu \partial_\mu \phi \, \delta \phi \right] - \frac{dV(\phi)}{d\phi} \delta \phi. \tag{5.1.24}$$

It follows that the action is

$$S = \int_a^b \delta \mathcal{L} \, d^4 x = \int_a^b \left\{ - \left[\partial_\mu \left(\partial^\mu \phi \, \delta \phi \right) - \partial^\mu \partial_\mu \phi \, \delta \phi \right] - \frac{dV(\phi)}{d\phi} \delta \phi \right\} \, d^4 x. \tag{5.1.25}$$

The total derivative term $\partial_\mu \left(\partial^\mu \phi \, \delta \phi \right)$ vanishes as we require the variations $\delta \phi = 0$ at a and b. This leaves us with

$$S = \int_a^b \left\{ \partial^\mu \partial_\mu \phi - \frac{dV(\phi)}{d\phi} \right\} \delta \phi \, d^4 x. \tag{5.1.26}$$

We require that this equality hold for any variation $\delta \phi$. So it must be true that

$$\partial^\mu \partial_\mu \phi - \frac{dV(\phi)}{d\phi} = 0. \tag{5.1.27}$$

We introduce a new operator, the **d'Alembertian**:

$$\Box \equiv \partial^\mu \partial_\mu \equiv \partial_\nu \partial^\mu \equiv \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2. \tag{5.1.28}$$

The requirement we just derived now becomes the **Klein-Gordon equation**:

$$\Box \phi - \frac{dV}{d\phi} = 0. \tag{5.1.29}$$

Remember that we are working with Lagrangian for a scalar field. It can easily be shown the connection between the Klein-Gordon equations and Newton's second law of motion, by separating the temporal and spatial derivatives from the d'Alembertian and rewriting a few things:

$$\Box\phi - \frac{dV}{d\phi} = \ddot{\phi} - \vec{\nabla}^2\phi - \frac{dV(\phi)}{d\phi} = 0. \tag{5.1.30}$$

We can see the time second derivative on the field ϕ and the ϕ-derivative on the potential field resemble "acceleration" and "force" in Newton's second law.

Let us return to our original question of why a scalar field has the form $\phi \sim e^{-i\mathbf{k}\cdot\mathbf{x}}$. From our derivation of the Klein-Gordon equation, we observe that a scalar field ϕ must be a solution to the Klein-Gordon equation. Now, we verify that

$$\phi = e^{-i\mathbf{k}\cdot\mathbf{x}} \tag{5.1.31}$$

is a solution to the KG equation. Note that even though we are concerned with real scalar field for the time being, it is reasonable to have ϕ of that particular complex form, simply because it makes taking derivatives easier. One can certainly work with only the real part of ϕ and get the same result. Now, to check when ϕ satisfies the Klein-Gordon equation, we simply unpack the d'Alembertian and attack the derivatives step-by-step. The first derivative is

$$\partial_\mu\phi = \partial_\mu\left(e^{-i\mathbf{k}\cdot\mathbf{x}}\right) \tag{5.1.32}$$

$$= -i\partial_\mu\left(\mathbf{k}\cdot\mathbf{x}\right)e^{-i\mathbf{k}\cdot\mathbf{x}} \tag{5.1.33}$$

$$= -i\partial_\mu\left(K_\nu X^\nu\right)e^{-iK_\alpha X^\alpha} \tag{5.1.34}$$

$$= -iK_\nu\,\partial_\mu X^\nu\,\phi \tag{5.1.35}$$

$$= -iK_\nu\delta_\mu^\nu\phi \tag{5.1.36}$$

$$= -iK_\mu\phi \tag{5.1.37}$$

Next, we attack the second derivative:

$$\partial^\mu\partial_\mu\phi = \eta^{\mu\nu}\partial_\nu\partial_\mu\phi \tag{5.1.38}$$

$$= \eta^{\mu\nu}\partial_\nu\left(-iK_\mu\phi\right) \tag{5.1.39}$$

$$= -iK_\mu\eta^{\mu\nu}\left(-iK_\nu\phi\right) \tag{5.1.40}$$

$$= (-i)^2 K^\mu K_\mu\phi. \tag{5.1.41}$$

If $K^\mu K_\mu = m^2$ (as we have shown before), then

$$\Box\phi + m^2\phi = (-m^2 + m^2)\phi = 0, \tag{5.1.42}$$

which satisfies the Klein-Gordon equation. So, as long as $K^\mu K_\mu = m^2$ is satisfied, ϕ of the given form is a solution to the KG equation and is a legitimate

scalar field.

Without knowing the solution to the Klein-Gordon equation, we can also verify that the Klein-Gordon equation is the equation of motion via the Euler-Lagrange equation, which says that

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) = 0. \tag{5.1.43}$$

Recall the Lagrangian:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - V(\phi), \tag{5.1.44}$$

we get

$$\frac{\partial \mathcal{L}}{\partial \phi} = -\frac{dV}{d\phi} \tag{5.1.45}$$

$$\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} = \frac{1}{2}\partial^\mu \phi. \tag{5.1.46}$$

So the Euler-Lagrange equation gives:

$$-\frac{dV}{d\phi} - \frac{1}{2}\partial_\mu(\partial^\mu \phi) = 0. \tag{5.1.47}$$

If we reasonably take

$$V(\phi) = \frac{1}{2}m^2 \phi^2, \tag{5.1.48}$$

then we get $-m^2\phi - \Box\phi = 0$, i.e.

$$(\Box + m^2)\phi = 0. \tag{5.1.49}$$

5.2   Complex Scalar Fields and Electromagnetism

5.2.1   Gauge transformation of the first kind - Global Symmetry

In this section, we will see how things "come together." As real scalar fields describe massive neutral particles and how vector fields describe massless photons, complex scalar fields somehow fills in and completes the picture by bringing in the charge and somehow joins the real scalar and vector field pictures. At the end of this section, we will see how the "full theory" where the electromagnetic field with charge is described by a single Lagrangian.

If our scalar field is complex, we can think of it and its complex conjugate as:

$$\phi = \frac{\phi_r + i\phi_i}{\sqrt{2}} \tag{5.2.1}$$

$$\phi^* = \frac{\phi_r - i\phi_i}{\sqrt{2}}. \tag{5.2.2}$$

Or, we can think of ϕ and ϕ^* as independent fields. Since the Lagrangian has to be real, plugging ϕ and ϕ^* into the Lagrangian introduced in the previous section gives:

$$\mathcal{L} = (\partial_\mu \phi)(\partial^\mu \phi)^* - m^2 \phi^* \phi. \tag{5.2.3}$$

Without much inspection, we can see that the fields satisfy the Klein-Gordon equation:

$$\begin{cases} (\Box + m^2)\phi & = 0 \\ (\Box + m^2)\phi^* & = 0. \end{cases} \tag{5.2.4}$$

Now, we can ask ourselves: "Is this Lagrangian invariant under the transformation

$$\phi \to e^{-i\Lambda}\phi \tag{5.2.5}$$

$$\phi^* \to e^{i\Lambda}\phi^*, \tag{5.2.6}$$

where Λ is a constant?" Again, by inspection, the answer is yes, simply because the transformation is nothing but an orthogonal rotation (length preserving). And since because we changed the field the same way everywhere by a rotation induced by $e^{\pm i\Lambda}$ while the Lagrangian remains invariant, we say that the theory has a **global U(1) symmetry**. Global transformations like this one are called *gauge transformation of the first kind*.

5.2.2 Gauge transformation of the second kind - Local Symmetry

Now, consider the case where the rotation operation $e^{\pm i\Lambda}$ has a spatio-dependence, i.e., $\Lambda \to \Lambda(x)$. We call this a *gauge transformation of the second kind*. This leads to the fact that ϕ transform differently at different places. We want our theory to also be invariant under this "local" transformation. But first, we have to show that our previous Lagrangian no longer retains its gauge symmetry and try to come up with ways to "fix" the theory. Let us say that ϕ transforms as $\phi \to \phi e^{i\Lambda(x)}$, then

$$\partial_\mu \phi \to \partial_\mu(\phi e^{i\Lambda(x)}) = i(\partial_\mu \Lambda(x))\phi + e^{i\alpha}\partial_\mu \phi. \tag{5.2.7}$$

It is already clear that

$$(\partial_\mu \phi)(\partial_\mu \phi)^* \nrightarrow (\partial_\mu \phi)(\partial_\mu \phi)^* e^{i\Lambda(x)} \tag{5.2.8}$$

since there will always be some $\Lambda(x)$-dependent residual terms. We can stop here and surrender, but we can also "fix" our notion of the derivative (just like what we did in general relativity), introducing the **gauge-covariant derivative**:

$$D_\mu = \partial_\mu + iqA_\mu, \tag{5.2.9}$$

where q is the charge, which acts as a *coupling term* like the Christoffel symbols in general relativity, and A_μ is the vector potential, which is also the gauge field that has symmetry:

$$A_\mu \sim A_\mu + \partial_\mu \Lambda. \tag{5.2.10}$$

Once again, recall we have discussed in the gauge invariance section. By adding the 4-dimensional gradient of a scalar field Λ (so adding the $\partial_\mu \Lambda$ term), we are not changing the 4-dimensional curl of A_μ, i.e. the electromagnetic field tensor $F^{\mu\nu}$ is divergence-free. We can refer to section on gauge invariance for more details about this is true. But focusing on the results here, if we cleverly (once again) pick Λ such that

$$\Lambda(x) = -\frac{\alpha(x)}{q} \tag{5.2.11}$$

then with the gauge transformations

$$\begin{cases} \phi \to \phi e^{i\Lambda(x)} \\ A_\mu \to A_\mu - \frac{1}{q}\partial_\mu \Lambda(x) \end{cases} \tag{5.2.12}$$

the new derivative $D_\mu \phi = (\partial_\mu + iqA_\mu)\phi$ is transformed into

$$\to \left(\partial_\mu iq \left[A_\mu - \frac{1}{q}\partial_\mu \Lambda(x)\right]\right)\phi e^{i\Lambda(x)} \tag{5.2.13}$$

$$= e^{i\Lambda(x)}\left(\partial_\mu \phi + i(\partial_\mu \Lambda(x)\phi) + iqA_\mu \phi - i(\partial_\mu \Lambda(x))\phi\right) \tag{5.2.14}$$

$$= e^{i\Lambda(x)}(\partial_\mu \phi + iqA_\mu \phi) \tag{5.2.15}$$

$$= e^{i\Lambda(x)} D_\mu \phi. \tag{5.2.16}$$

So, $D_\mu \phi$ transforms correctly. Since we have added the electromagnetic vector potential to the theory, adding the electromagnetic field tensor term completes the classical field theory for the charged scalar field in electromagnetism.

$$\boxed{\mathcal{L} = |D_\mu \phi|^2 - m^2 |\phi|^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}} \tag{5.2.17}$$

In the next section, we will look at the underlying motivations for the choices we have made regarding "fixing" our notion of the derivatives and Λ.

5.2.3   Motivations in the derivation of the E&M Lagrangian

In the previous section, we have look at the basics of gauge transformations and global and local symmetry. We have also come across a number of "clever" choices, but those choices have so far been unjustified. in this section, we will try to develop an understanding for "why" the complete Lagrangian of electromagnetism and matter has the following form

$$\mathcal{L} = |D_\mu \phi|^2 - m^2 |\phi|^2 - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}. \tag{5.2.18}$$

First, we revisit the gauge transformation of the second kind, i.e., local gauge transformation where

$$\phi \to \phi' = e^{i\Lambda(x)} \phi \approx (1 - i\Lambda(x))\phi \tag{5.2.19}$$

for $\Lambda(x) \ll 1$. Variations on the field give

$$\delta\phi = -i\Lambda\phi \tag{5.2.20}$$

and

$$\partial_\mu \phi \to \partial_\mu \phi - i(\partial_\mu \Lambda(x))\phi - i\Lambda(x)(\partial_\mu \phi). \tag{5.2.21}$$

Therefore

$$\delta(\partial_\mu \phi) = -i\Lambda(\partial_\mu \phi) - i(\partial_\mu \Lambda)\phi. \tag{5.2.22}$$

We can do the same with the complex conjugate of ϕ, giving

$$\delta\phi^* = i\Lambda\phi^* \tag{5.2.23}$$

$$\delta(\partial_\mu \phi^*) = i\Lambda(\partial_\mu \phi^*) + i(\partial_\mu \Lambda)\phi^*. \tag{5.2.24}$$

Assuming that the Lagrangian for electromagnetism takes the form like that of the massive scalar field case:

$$\mathcal{L} = (\partial_\mu \phi)(\partial^\mu \phi^*) - m^2 \phi^* \phi, \tag{5.2.25}$$

variations on the Lagrangian gives

$$\delta\mathcal{L} = \delta[(\partial_\mu \phi)(\partial^\mu \phi^*)] - m^2 \delta(\phi^* \phi) \tag{5.2.26}$$

$$= [\delta(\partial_\mu \phi)](\partial^\mu \phi^*) + (\partial_\mu \phi)[\delta(\partial^\mu \phi^*)] - m^2[(\delta\phi)\phi^* + \phi(\delta\phi^*)]. \tag{5.2.27}$$

We can readily verify that $(\delta\phi)\phi^* + \phi(\delta\phi^*)$ is zero by the above identities:

$$(\delta\phi)\phi^* + \phi(\delta\phi^*) = -i\Lambda\phi\phi^* + \phi i\Lambda\phi^* = 0. \tag{5.2.28}$$

So we're left with

$$\delta\mathcal{L} = [-i\Lambda(\partial_\mu\phi) - i(\partial_\mu\Lambda)\phi](\partial^\mu\phi^*) + (\partial_\mu\phi)[i\Lambda(\partial_\mu\phi^*) + i(\partial_\mu\Lambda)\phi^*] \quad (5.2.29)$$
$$= -i\Lambda(\partial_\mu\phi)(\partial^\mu\phi^*) - i(\partial_\mu\Lambda)\phi(\partial^\mu\phi^*) \quad (5.2.30)$$
$$+ i\Lambda(\partial^\mu\phi^*)(\partial_\mu\phi) + i\phi^*(\partial^\mu\Lambda)(\partial_\mu\phi) \quad (5.2.31)$$
$$= i\phi^*(\partial^\mu\Lambda)(\partial_\mu\phi) - i(\partial_\mu\Lambda)\phi(\partial^\mu\phi^*) \quad (5.2.32)$$
$$= i\phi^*(\partial_\mu\Lambda)(\partial^\mu\phi) - i(\partial_\mu\Lambda)\phi(\partial^\mu\phi^*) \quad (5.2.33)$$
$$= (\partial_\mu\Lambda)[i\phi^*\partial^\mu\phi - i\phi\partial^\mu\phi^*]. \quad (5.2.34)$$

By Noether's theorem, which we will into much more detail in the following sections:

$$J^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}(-i\phi) + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^*)}(i\phi^*) \quad (5.2.35)$$
$$= i(\phi^*\partial^\mu\phi - \phi\partial^\mu\phi^*). \quad (5.2.36)$$

So the variation in the Lagrangian becomes

$$\delta\mathcal{L} = (\partial_\mu\Lambda)J^\mu. \quad (5.2.37)$$

Now, our ultimate goal is achieve some sort of invariance. But by the look of $\delta\mathcal{L}$ we are already failing to have invariance. However, we can fix this, by adding terms to the Lagrangian such that this J^μ term goes away (this is exactly the idea of *gauge invariance* which we have discussed earlier). So, consider

$$\mathcal{L}_1 = -eJ^\mu A_\mu \quad (5.2.38)$$
$$= -e[i\phi^*\partial^\mu\phi - i\phi\partial^\mu\phi^*]A_\mu. \quad (5.2.39)$$

where we have introduced the vector potential A_μ. Recall in the section on gauge invariance we have require A_μ to transform as

$$A'_\mu = A_\mu + \frac{1}{e}\partial_\mu\Lambda \quad (5.2.40)$$

in order to achieve gauge invariance. (Note that there is a subtle sign change due to different sign conventions, but the idea is the same). Thus variations on A_μ is given by

$$\delta A_\mu = A'_\mu - A_\mu = \frac{1}{e}\partial_\mu\Lambda \quad (5.2.41)$$

Plug this into the variation of this extra term $\mathcal{L}_1$ we get

$$\delta\mathcal{L}_1 = -\delta[eJ^\mu A_\mu] \quad (5.2.42)$$
$$= -e(\delta J^\mu)A_\mu - eJ^\mu(\delta A_\mu) \quad (5.2.43)$$
$$= -e(\delta J^\mu)A_\mu - J^\mu\partial_\mu\Lambda \quad (5.2.44)$$

We shall compute δJ^μ here:

$$\delta J^\mu = i\delta(\phi^*\partial^\mu\phi - \phi\partial^\mu\phi^*) \tag{5.2.45}$$

$$= i[(\delta\phi^*)\partial^\mu\phi + \phi^*\partial^\mu(\delta\phi) - \phi(\partial^\mu\delta\phi^*) - (\delta\phi)\partial^\mu\phi^*] \tag{5.2.46}$$

$$= i[i\Lambda\phi^*\partial^\mu\phi - \phi^*\partial^\mu(i\Lambda\phi) - \phi\partial^\mu(i\Lambda\phi^*) + i\Lambda\phi\partial^\mu\phi^*] \tag{5.2.47}$$

$$= -\Lambda\phi^*\partial^\mu\phi + \phi^*\partial^\mu(\Lambda\phi) + \phi\partial^\mu(\Lambda\phi^*) - \Lambda\phi\partial^\mu\phi^* \tag{5.2.48}$$

$$= -\Lambda\phi^*\partial^\mu\phi + \Lambda\phi^*\partial^\mu\phi + \phi^*\phi\partial^\mu\Lambda + \phi\Lambda\partial^\mu\phi^* + \phi\phi^*\partial^\mu\Lambda - \Lambda\phi\partial^\mu\phi^* \tag{5.2.49}$$

$$= 2\phi^*\phi\partial^\mu\Lambda. \tag{5.2.50}$$

So, variations on the modified Lagrangian is now

$$\delta\mathcal{L} + \delta\mathcal{L}_1 = (\partial_\mu\Lambda)J^\mu - e(2\phi^*\phi\partial^\mu\Lambda) - J^\mu(\partial_\mu\Lambda) \tag{5.2.51}$$

$$= -2eA_\mu\phi^*\phi(\partial^\mu\Lambda). \tag{5.2.52}$$

But it looks like we are *still* achieving invariance. So, we have to repeat the procedure, adding more terms to the Lagrangian. Let's call this additional term $\mathcal{L}_2$, defined as

$$\mathcal{L}_2 = e^2 A_\mu A^\mu \phi^*\phi. \tag{5.2.53}$$

Varying $\mathcal{L}_2$ gives

$$\delta\mathcal{L}_2 = \delta(e^2 A_\mu A^\mu \phi^*\phi) \tag{5.2.54}$$

$$= e^2\phi^*\phi A_\mu \delta A^\mu + e^2\phi^*\phi A^\mu \delta A_\mu \tag{5.2.55}$$

$$= 2e^2 A_\mu(\partial^\mu\Lambda)\phi^*\phi. \tag{5.2.56}$$

Now, we should achieve invariance:

$$\delta\mathcal{L} + \delta\mathcal{L}_1 + \delta\mathcal{L}_2 = -2eA_\mu\phi^*\phi(\partial^\mu\Lambda) + 2e^2 A_\mu(\partial^\mu\Lambda)\phi^*\phi \tag{5.2.57}$$

$$= 0. \tag{5.2.58}$$

The total Lagrangian $\mathcal{L} + \mathcal{L}_1 + \mathcal{L}_2$ is now invariant under local gauge transformation, but at the price of introducing the vector potential A_μ, which couples to the current J_μ of the complex ϕ field. Now, A_μ must also be contributing to the Lagrangian, so we need to take into account for that contribution by requiring an additional (invariant) term that involves nothing but the electromagnetic field strength tensor:

$$\mathcal{L}_3 = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{5.2.59}$$

So, the total Lagrangian becomes:

$$\mathcal{L}_{\text{tot}} = \mathcal{L} + \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 \tag{5.2.60}$$

$$= (\partial_\mu\phi)(\partial^\mu\phi^*) - ie(\phi^*\partial^\mu\phi - \phi\partial^\mu\phi^*)A_\mu \tag{5.2.61}$$

$$e^2 A_\mu A^\mu \phi^*\phi - m^2\phi * \phi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{5.2.62}$$

Some simplification gives

$$\mathcal{L}_{\text{tot}} = (\partial_\mu \phi + ieA_\mu \phi)(\partial^\mu \phi^* - ieA^\mu \phi^*) - m^2 \phi^* \phi - \frac{1}{4} F^{\mu\nu} F_{\mu\nu}. \qquad (5.2.63)$$

Now, the form of the Lagrangian is starting to look like the Lagrangian we wish to achieve in the beginning of this section. Indeed, we can see that our derivative has been slightly modified to include the vector potential:

$$\partial_\mu \to \partial_\mu + ieA_\mu. \qquad (5.2.64)$$

So we if define a new operator

$$D_\mu = \partial_\mu + ieA_\mu, \qquad (5.2.65)$$

then we see that it is indeed invariant under local gauge transformations (as already been shown numerous times in this text). It is worthwhile, however, to show again, but in a slightly different light:

$$\delta(D_\mu \phi) = \delta(\partial_\mu \phi) + ie(\delta A_\mu)\phi + ieA_\mu \delta\phi \qquad (5.2.66)$$
$$= -i\Lambda(\partial_\mu \phi + ieA_\mu \phi) \qquad (5.2.67)$$
$$= -i\Lambda(D_\mu \phi). \qquad (5.2.68)$$

In fact, this is true even without the approximation $\Lambda \ll 1$. To show this, we also have to take into account the transformation of $A_\mu \to A'_\mu$, i.e., the operators are not the same at different locations in space. But it can easily be done. In fact, we have shown this in the section on gauge invariance.

Now, we are getting really close to coming a full circle to justify the numerous of "clever choices" in the previous sections. However, we shall still be careful and look at a subtle difference:

$$D_\mu \equiv \partial_\mu + ieA_\mu \qquad (5.2.69)$$

if applied to the scalar field ϕ, while

$$D_\mu \equiv \partial_\mu - ieA_\mu \qquad (5.2.70)$$

if applied to the complex conjugate scalar field ϕ^*. But the operator itself shouldn't know in advance what kind of function is being fed to it so that it can change accordingly. Also, the role of "complex conjugate" is exchangeable, thus a "sign-varying" definition of D_μ certainly does not work. To resolve this, we must recognize this: because e is the electric charge, changing the sign of e changes the sign of the charge. Now, because this goes hand-in-hand with the field ϕ and its complex conjugate ϕ^*, i.e., whenever we have the complex conjugate, the sign of e in the operator changes, we might as well associate the field ϕ with having the charge e, and the field ϕ^* with having the charge $-e$. To summarize, we simply note: ϕ describes a field with charge e, and ϕ^* describes

a field with charge $-e$. So, we can now be satisfied with the Lagrangian for electromagnetism and classical field theory (in flat spacetime):

$$\boxed{\mathcal{L} = (D_\mu \phi)(D^\mu \phi^*) - m^2 A_\mu A^\mu - \frac{1}{4} F^{\mu\nu} F_{\mu\nu}} \tag{5.2.71}$$

5.2.4   A few remarks

Here are a few key concepts we should clarify and make sure we understand correctly. These can be treated as important takeaways from this section. It also helps to have some "intuitive" grounding and a good "narrative" in mind when we talk about the theory. Sometimes, while the mathematics is straightforward, it can be tremendously difficult to interpret a theory and "understand" it. So, while not as technical as the other sections, this subsubsection serves to summarize and hopefully explain why we did what we have done and defined what we have defined. I hope this section sort of "brings everything together" before we move on to other topics.

1. $D_\mu \phi^*$ is a covariant derivative of ϕ^* not because we conjugated $D_\mu \phi$ but because it transforms in the same way that ϕ^* does under gauge transformations. We can readily verify this (in fact, we have verified this many times).

2. How can we be sure that A_μ is actually the vector potential? This is possibly a question that a careful reader would ask right from the beginning. But no worries, as we know for sure that A_μ must be the (compensating) vector potential. Why? Because of the definition of the electromagnetic field strength tensor as the 4-dimensional curl of A_μ.

3. As a result of the previous item, we have a new interpretation for the electromagnetic field: The electromagnetic field is the gauge field which has to be introduced to guarantee invariance under local U(1) gauge transformations.

4. The Maxwell's inhomogeneous equations (i.e., not in a vacuum, i.e., there exists some current $\mathcal{J}^\mu$) arise as a result of varying the vector potential A_μ, i.e., by requiring that A_μ satisfy the Euler-Lagrange equation for A_μ:

$$\frac{\partial \mathcal{L}}{\partial A_\mu} - \partial_\nu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\nu A_\mu)} \right) = 0 \tag{5.2.72}$$

   where the Lagrangian is what we just derived

$$\mathcal{L} = (D_\mu \phi)(D^\mu \phi) - m^2 A_\mu A^\mu - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} \tag{5.2.73}$$

$$= (\partial_\mu \phi + ieA_\mu \phi)(\partial^\mu \phi^* - ieA^\mu \phi^*) - m^2 \phi^* \phi - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} \tag{5.2.74}$$

we obtain (which the unconvinced and diligent reader can and should readily verify by attacking the Lagrangian with derivatives)

$$\partial_\nu F^{\mu\nu} = -ie(\phi^* \partial^\mu \phi - \phi \partial^\mu \phi^*) + 2e^2 A^\mu |\phi|^2 \qquad (5.2.75)$$

$$= -ie(\phi^* D^\mu \phi - \phi D^\mu \phi^*) \qquad (5.2.76)$$

$$= -e\mathcal{J}^\mu, \qquad (5.2.77)$$

where the definition of the **covariant current** $\mathcal{J}^\mu$

$$\mathcal{J}^\mu \equiv \phi^* D^\mu \phi - \phi D^\mu \phi^* \qquad (5.2.78)$$

is "covariantly motivated" by the definition of J^μ. Now, because

$$\partial_\mu(\partial_\nu F^{\mu\nu}) = \partial_\mu \partial_\nu (\partial^\mu A^\nu - \partial^\nu A^\mu) \qquad (5.2.79)$$

$$= \partial_\mu \partial_\nu \partial^\mu A^\nu - \partial_\mu \partial_\nu \partial^\nu A^\mu \qquad (5.2.80)$$

$$= \partial_\mu \partial_\nu \partial^\mu A^\nu - \partial_\nu \partial_\mu \partial^\mu A^\nu \qquad (5.2.81)$$

$$= 0, \qquad (5.2.82)$$

it must also be true that

$$\partial_\mu \mathcal{J}^\mu = 0, \qquad (5.2.83)$$

i.e., the covariant current is **conserved** when the electromagnetic field is present, NOT the current J^μ.

5. We have shown very early in the section that the electromagnetic field is **massless**, but it is worth emphasizing this point. If we add the mass term:

$$\mathcal{L}_M = M^2 A_\mu A^\mu \qquad (5.2.84)$$

to the Lagrangian, we no longer get invariance under gauge transformation. This means that *gauge invariance requires the electromagnetic field - the gauge field - to be massless.* This turns out to be very important particle physics.

6. What does the charge e do in the theory? We can indeed think of e as a coupling constant, as we have discussed. If we look at the Lagrangian, we can see clearly that the field ϕ couples to the electromagnetic field with strength e. So, it seems that the electric charge plays **two roles** in theory: (i) measuring the strength with which a particle interacts with electric and magnetic fields, and (ii) serves as a conserved quantity. These two roles in fact arise as a consequence of the "gauge principle," which is also very important in particle physics.

5.3   Vector Fields and Photons

In this section, we will see how electromagnetism naturally arises from the previous section where we demand invariance of the action under gauge transformation of the second kind (local rotations). In particular, we will work with

a Lagrangian density that describes the massless and neutral photon, and see how variations on the action gives rise to the Maxwell's equations. We will also confirm that fact that photons must be massless in the theory, by enforcing conditions on the vector potential A_μ. The section is more of synopsis of the methods we have learned so far about variations on the action, gauge transformations, symmetries, and invariances. The goal of this section is to show the theory actually works to describe a physical, real, particle. We will be working with vector field in the formulation the action, but the main principles regarding variations and gauge-stuff are the same.

Vector fields describe particles of spin 1 such as photons. Unlike real scalar fields ϕ where there is only one degree of freedom, a vector field is represented by A_μ with $\mu = 0, 1, 2, 3$, hence having 4 degrees of freedom. Electromagnetism is a field theory where the relevant field is a vector field, A_μ, called the vector potential.

$$A_\mu = (A_0, \vec{A}). \tag{5.3.1}$$

The first component of the vector potential, A_0 is the electrostatic potential V where $\vec{E} = -\vec{\nabla}V$. The other spatial components of A_μ, forming $\vec{A}$, form the vector potential from which the magnetic field and full electric field is derived:

$$\vec{B} = \vec{\nabla} \times \vec{A} \tag{5.3.2}$$

$$\vec{E} = -\vec{\nabla}V - \frac{\partial \vec{A}}{\partial t} \tag{5.3.3}$$

Let us consider the following Lagrangian density:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - j^\mu A_\mu, \tag{5.3.4}$$

where $j^\mu = (\rho, \vec{J})$ is a combination of the charge density and current density. The electromagnetic field strength tensor is given by:

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{5.3.5}$$

$$= \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix}. \tag{5.3.6}$$

With this definition, we can also have an equivalent definition:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{5.3.7}$$

Recall the cyclic identity (this can be readily verified - we in fact have covered this in the GR notes):

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0. \tag{5.3.8}$$

We can easily show that this identity yields two of four Maxwell's equations:

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{5.3.9}$$

$$\vec{\nabla} \cdot \vec{B} = 0. \tag{5.3.10}$$

The remaining Maxwell equations come from varying the action and minimizing the action: $\delta S = 0$ with respect to the vector potential A_μ. Similar to what we have done before, we want to vary the Lagrangian. Now, the E&M Lagrangian has two terms. The term involving the vector potential is simple:

$$\delta \left(j^\mu A_\mu \right) = j^\mu \, \delta A_\mu \tag{5.3.11}$$

true for all δA_μ, so if the field strength tensor is zero, then $j^\mu = 0$. The term involving the field strength tensor is a little more complicated, but certainly doable:

$$\delta \left(\frac{-1}{4} F^{\mu\nu} F_{\mu\nu} \right) = \frac{-1}{4} \delta \left[\left(\partial^\mu A^\nu - \partial^\nu A^\nu \right) \left(\partial_\mu A_\nu - \partial_\nu A_\mu \right) \right] \tag{5.3.12}$$

$$= \frac{-1}{2} \delta \left(\partial^\mu A^\nu \, \partial_\mu A_\nu - \partial_\mu A_\nu \, \partial^\mu A^\nu \right) \tag{5.3.13}$$

$$= \frac{-1}{2} \left(\partial^\mu \, \delta A^\nu \, \partial_\mu A_\nu + \partial^\mu A^\nu \, \partial_\mu \, \delta A_\nu - \partial_\mu \, \delta A_\nu \, \partial^\mu A^\nu - \partial_\mu A_\nu \, \partial^\mu \, \delta A^\nu \right). \tag{5.3.14}$$

Raising and lowering indices gives

$$\delta \left(\frac{-1}{4} F^{\mu\nu} F_{\mu\nu} \right) = \frac{-1}{2} \left(\partial^\mu \, \delta A^\nu \, \partial_\mu A_\nu + \partial_\nu A_\mu \, \partial^\nu \, \delta A^\mu - \partial_\mu \, \delta A_\nu \, \partial^\mu A^\nu - \partial^\nu A^\mu \, \partial_\nu \, \delta A_\mu \right) \tag{5.3.15}$$

$$= \partial^\nu \, \delta A^\mu \, \partial_\nu A_\mu - \partial_\mu \, \delta A_\nu \, \partial^\mu A^\nu. \tag{5.3.16}$$

We can again integrate by parts on the two terms similar to the following steps

$$\partial^\nu \, \delta A^\mu \, \partial_\nu A_\mu = \partial^\nu (\partial_\mu A_\nu \, \delta A^\mu) - (\partial^\nu \partial_\mu A_\nu) \delta A^\mu = \partial_\mu (\partial^\nu A^\mu \, \delta A_\nu) - (\partial^\nu \partial_\mu A_\nu) \delta A^\mu \tag{5.3.17}$$

$$\partial_\mu \, \delta A_\nu \partial^\mu A^\nu = \partial_\mu (\partial^\mu A^\nu \, \delta A_\nu) - \partial_\mu (\partial^\mu A^\nu) \, \delta A_\nu = \partial_\mu (\partial^\mu A^\nu \, \delta A_\nu) - \partial_\mu (\partial^\mu A^\nu) \, \delta A_\nu. \tag{5.3.18}$$

and eliminate the total derivative from the action integral. Assuming that the term with the current density and vector potential is zero, we are eventually (after lowering/raising the indices correctly, of course) left with the requirement

$$\partial_\mu (\partial^\mu A^\nu) \, \delta A_\nu - (\partial^\nu \partial_\mu A_\nu) \delta A^\mu \equiv \left(\Box A^\mu - \partial^\mu \partial^\nu A_\nu \right) \delta A_\mu = 0 \tag{5.3.19}$$

for all δA_μ, which forces the following identity:

$$\Box A^\mu - \partial^\nu \partial^\mu A_\nu = \Box A^\mu - \partial_\nu \partial^\mu A^\nu = \partial_\nu (\partial^\nu A^\mu - \partial^\mu A^\nu) = \partial_\nu F^{\mu\nu} = 0. \tag{5.3.20}$$

Now, with the current density and vector potential terms, we get the requirement

$$\partial_\nu F^{\mu\nu} = j^\mu. \tag{5.3.21}$$

This identity gives the remaining two Maxwell's equations.

We can look at photons as an example. Photons do not carry a current/charge, so $j^\mu = 0$. Therefore the equation of motion can be derived from just

$$\partial_\nu F^{\mu\nu} = 0. \tag{5.3.22}$$

Now, we have an interesting problem to think about: We know that photons can have 2 independent transverse polarizations, i.e. there are 2 massless modes for photons. However, A^μ has 4 degrees of freedom, not 2. So why does our theory require more than 2 degrees of freedom to describe a physical quantity that only has 2 degrees of freedom? The answer to this is that there are 2 degrees of freedom in A_μ that don't matter. The first is the A_0 factor - the electrostatic potential. Why A_0 does not matter in describing photons can be illustrated if we look at the case where $\mu = 0$:

$$\Box A_0 - \partial_0 \, \partial^\nu A_\nu = \partial^0 \partial_0 A_0 + \partial^j \partial_j A_0 - \partial_0 \partial^0 A_0 - \partial_0 \partial^j A_j \tag{5.3.23}$$

$$= \partial^j \partial_j A_0 - \partial_0 \partial^j A_j \tag{5.3.24}$$

$$= 0. \tag{5.3.25}$$

We see that A_0 is not a propagating mode, or the **ghost** mode, or the **auxiliary** mode. This is actually a good thing in our theory. In fact, the Lagrangian is actually chosen such that the time second derivative vanishes.

Now that we have sort of explained why one degree of freedom of A_μ does not matter. What about the other one that shouldn't matter? The short answer to this is the keyword **gauge symmetry** in the theory. If we look back at how the field strength tensor is defined as the 4-dimensional curl of the vector potential A_μ:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{5.3.26}$$

As an aside, it makes sense to define the 4-dimensional curl this way, because the curl of a 2-vector field in two dimensions gives a scalar (a zero-index object), the curl of a 3-vector field in three dimensions gives a 3-vector field (a one-index object). So it is reasonable to define the 4-dimensional curl of a 4-vector field as a tensor (a two-index object). Now, back to our problem. We should also attempt to gauge transform

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu \Lambda(x), \tag{5.3.27}$$

then we observe that

$$F'_{\mu\nu} = \partial_\mu(A_\nu + \partial_\nu\Lambda(x)) - \partial_\nu(A_\mu + \partial_\mu\Lambda(x)) \tag{5.3.28}$$

$$= \partial_\mu A_\nu - \partial_\nu A_\mu \tag{5.3.29}$$

$$= F^{\mu\nu}, \tag{5.3.30}$$

i.e. there is a way to choose $\Lambda(x)$ such that we eliminate one A_μ mode, leaving just 4-1-1=2 modes.

The goal of the above gauge transformation is to pick a $\Lambda(x)$ to remove an extra degree of freedom in the theory. Let us suppose that $\partial^\mu A_\mu \neq 0$, i.e. A_μ is NOT divergence-free. We can pick $\Lambda(x)$ such that after the transformation, $\partial^\mu A'_\mu = 0$. Let's start from the very beginning, repeat what we just done earlier:

$$A'_\nu = A_\nu + \partial_\nu\Lambda. \tag{5.3.31}$$

It follows that

$$\partial^\nu A'_\nu = \partial^\nu(A_\nu + \partial_\nu\Lambda) \tag{5.3.32}$$

$$= \partial^\nu A_\nu + \partial^\nu \partial_\nu\Lambda \tag{5.3.33}$$

$$= \partial^\nu A_\nu + \Box\Lambda. \tag{5.3.34}$$

This suggests that if we pick $\Lambda(x)$ such that $\Box\Lambda(x) = -\partial^\nu A_\nu$ then in this gauge $\partial^n u A'_\nu = 0$. After that, we can just drop the prime and get $\partial^\nu A_\nu = 0$. As a consequence, in this fixed gauge, the equations of motion can be deduced from

$$\begin{cases} \Box A_\mu = 0 \\ \partial^\nu A_\nu = 0 \end{cases} \tag{5.3.35}$$

each of which adds one constraint, reducing the degree of freedom by 1. To solve for A_μ, we can assume the form of A_μ:

$$A_\mu = \epsilon_\mu e^{-i\mathbf{k}\cdot\mathbf{x}}, \tag{5.3.36}$$

where ϵ_μ is called the **polarization vector**. Now, let us work with the first constraint $\Box A = 0$. We can verify (below) that this constraint requires $K^\mu K_\mu = 0$, i.e. K^μ is a null vector in flat Minkowskian spacetime, i.e. it is light-like.

$$\Box A_\mu = \partial_\mu \partial^\mu A_\mu \tag{5.3.37}$$

$$= \partial_\mu \left(\eta^{\mu\nu} \partial_\nu \epsilon_\mu e^{-i\mathbf{k}\cdot\mathbf{x}} \right) \tag{5.3.38}$$

$$= \partial_\mu \left[\epsilon_\mu \eta^{\mu\nu} \left(\partial_\nu e^{-i\mathbf{k}\cdot\mathbf{x}} \right) \right] \tag{5.3.39}$$

$$= \partial_\mu \left[\epsilon_\mu \eta^{\mu\nu} (-iK_\mu) \left(e^{-i\mathbf{k}\cdot\mathbf{x}} \partial_\nu X^\mu \right) \right] \tag{5.3.40}$$

$$= \partial_\mu \left[\epsilon_\mu \eta^{\mu\nu} (-iK_\mu) \left(e^{-i\mathbf{k}\cdot\mathbf{x}} \delta_\nu^\mu \right) \right] \tag{5.3.41}$$

$$= (-iK^\nu)\partial_\mu A_\mu \tag{5.3.42}$$

$$= 4(-iK^\nu)(-iK_\mu)A_\mu \tag{5.3.43}$$

$$= 4(-i)^2 K_\mu K^\mu A_\mu \tag{5.3.44}$$

$$= 0 \text{ if } K_\mu K^\mu = 0. \tag{5.3.45}$$

This requirement implies that the vector field A_μ is massless ($K^\mu K_\mu = 0$ rather than m^2 like in the scalar field example), which is a good thing, since we are working with electromagnetism and these mathematical objects ultimately describe electromagnetic waves - photons.

Next, the other constraint $\partial^\nu A_\nu = 0$ motivates us to choose Λ such that

$$\Box\Lambda = -\partial^\nu A_\nu \tag{5.3.46}$$

so that $\partial^\nu A'_\nu = 0$. However, this isn't enough information to choose Λ, because $\Box\Lambda = -\partial^\nu A_\nu$ alone cannot fix Λ. We also have to look at the gauge transformation of A_μ and make sure that the auxiliary part A_0 vanishes, then we have enough information to pick Λ. Let us start with the form of A_μ:

$$A_\mu = \epsilon_\mu e^{-i\mathbf{k}\cdot\mathbf{x}}. \tag{5.3.47}$$

We know that $\Box A_\mu = 0$, so if we pick

$$\Lambda = \lambda e^{-i\mathbf{k}\cdot\mathbf{x}} \tag{5.3.48}$$

then we are guaranteed $\Box\Lambda = 0 = \partial^\nu A_\nu$. So that is quite nice. But what is λ? We can only set λ if we fix the value of A_0. It makes sense to set $A_0 = 0$ so that it vanishes. With this, we can perform a gauge transformation on A_0, set it to zero, and find λ:

$$A_0 \to A_0 + \partial_0\Lambda = 0 \tag{5.3.49}$$

$$\to e^{-i\mathbf{k}\cdot\mathbf{x}}(\epsilon_0 - i\lambda K_0) = 0 \tag{5.3.50}$$

$$\to \lambda = \frac{\epsilon_0}{iK_0}. \tag{5.3.51}$$

So, if we pick

$$\Lambda = \frac{\epsilon}{iK_0}e^{-i\mathbf{k}\cdot\mathbf{x}} \tag{5.3.52}$$

then $A_0 = 0$ implies

$$\partial^\nu A_\nu = \partial^0 A_0 + \vec{\nabla}\cdot\vec{A} = \vec{\nabla}\cdot\vec{A} = 0. \tag{5.3.53}$$

So, it turns out that we can simply complete gauge fixing by setting two additional constraints on A_μ:

$$\begin{cases} A_0 = 0 \\ \vec{\nabla}\cdot\vec{A} = 0, \end{cases} \tag{5.3.54}$$

i.e., we require that the vector potential A_μ be divergence-free and to have a vanishing auxiliary mode. With this, we can the degree of freedom of A_μ from 4

from 2, making it a physical description. We shall show that can be done now. If we revisit the polarization vector

$$\epsilon_\mu = (\epsilon_0, \epsilon_1, \epsilon_2, \epsilon_3) \tag{5.3.55}$$

and require that $A_0 = 0$, then obviously $\epsilon_0 = 0$. But since we also require $\vec{\nabla} \cdot \vec{A} = 0$, we have

$$\vec{\nabla} \cdot \vec{A} = \partial_j A^j \tag{5.3.56}$$
$$= \partial_j \left(\epsilon^j e^{-i\mathbf{k}\cdot\mathbf{x}}\right) \tag{5.3.57}$$
$$\propto A^j(-iK_j) \tag{5.3.58}$$
$$= 0. \tag{5.3.59}$$

So, $\vec{k} \cdot \vec{A} = 0$, i.e., they are orthogonal vectors. Now, consider a photon traveling in the z-direction. Because $K_\mu k^\mu = 0$, we have

$$K^\mu = (K, 0, 0, K). \tag{5.3.60}$$

Because $\vec{k} \cdot \vec{A} \propto \vec{k} \cdot \vec{\epsilon}$, it follows that $K\epsilon^3 = 0$, i.e., there is no longitudinal component in the polarization vector. Therefore,

$$\epsilon_\mu = (0, \epsilon_1, \epsilon_2, 0). \tag{5.3.61}$$

So, the vector potential A_μ has the form

$$A_\mu = \epsilon_\mu e^{-i\mathbf{k}\cdot\mathbf{x}} = \begin{pmatrix} 0 \\ \epsilon_1 \\ \epsilon_2 \\ 0 \end{pmatrix} e^{-i\mathbf{k}\cdot\mathbf{x}}. \tag{5.3.62}$$

We can construct the independent modes of propagation as

$$\begin{cases} \epsilon_\mu^{(1)} = (0, 1, 0, 0)^\top \\ \epsilon_\mu^{(2)} = (0, 0, 1, 0)^\top, \end{cases} \tag{5.3.63}$$

which give two physical transverse modes of the vector potential

$$\begin{cases} A_\mu^{(1)} = (0, 1, 0, 0)^\top e^{-i\mathbf{k}\cdot\mathbf{x}} = \epsilon_\mu^{(1)} e^{-i\mathbf{k}\cdot\mathbf{x}} \\ A_\mu^{(2)} = (0, 0, 1, 0)^\top e^{-i\mathbf{k}\cdot\mathbf{x}} = \epsilon_\mu^{(2)} e^{-i\mathbf{k}\cdot\mathbf{x}}, \end{cases} \tag{5.3.64}$$

All is good, but we might wonder why photons are massless. The answer depends on who we ask, but mathematically, it is the gauge symmetry requirement that "makes" photons massless. Suppose that the Lagrangian is of the form

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} m^2 A_\mu A^\mu, \tag{5.3.65}$$

which suggests that the field is massive (hence the mass term m). Under gauge transformation,

$$A_\mu \to A_\mu + \partial_\mu \Lambda, \tag{5.3.66}$$

we have

$$\begin{cases} F_{\mu\nu} \to F_{\mu\nu} \\ A_\mu A^\mu = (A_\mu + \partial_\mu \Lambda)(A^\mu + \partial^\mu \Lambda) \not\to A_\mu A^\mu \end{cases} \tag{5.3.67}$$

So, the Lagrangian is no longer gauge-invariant. Therefore we claim that massive fields do not have gauge symmetry, i.e., we cannot have the mass term if we require mass invariance:

$$\text{Massless field} \iff \text{Gauge invariance.} \tag{5.3.68}$$

This also hints to us that mass comes from some mechanism that breaks gauge symmetry, which we will explore later on.

Chapter 6

Symmetries and Conservation Laws in Field Theory

6.1 Hamiltonian formalism (primer)

6.2 Overview of Noether's Theorem: A Consequence of Variational Principle

So far, we have seen quite a lot of *miraculous* coincidences, such as the fact that the Lagrangian somehow gives the Klein-Gordon equation and so on. We have also made a leap of faith from our traditional point-like description of particles x^μ to field-like descriptions ϕ and somehow the physics hasn't changed, i.e. we recognize that if the action is unchanged by a re-parameterization of x^μ and ϕ, then there exist one or more conserved quantities. This is the idea of Noether's theorem. In this subsection we will get an overview of Noether's theorem and apply it to illustrate conservation rules.

Let us go back and redefine the Lagrangian such that it also depends on x^μ - so that we take into account the interaction of ϕ with the space x^μ:

$$\mathcal{L} = \mathcal{L}(\phi, \partial_\mu \phi, x^\mu). \tag{6.2.1}$$

Next, recall our earlier definition of the variation:

$$\phi'(x) = \phi(x) + \delta\phi(x). \tag{6.2.2}$$

This definition merely compares ϕ' and ϕ at the same location in spacetime. To get the full, total variation, we define:

$$\Delta\phi = \phi'(x') - \phi(x) \tag{6.2.3}$$

$$= [\phi'(x') - \phi(x')] + [\phi(x') - \phi(x)] \tag{6.2.4}$$

$$\approx \delta\phi + (\partial_\mu\phi)\delta x^\mu. \tag{6.2.5}$$

So, the variation is now

$$\delta S = \int \mathcal{L}(\phi', \partial_\mu\phi', x'^\mu) \, d^4x' - \int \mathcal{L}(\phi, \partial_\mu\phi, x^\mu) \, d^4x \tag{6.2.6}$$

$$= \int \mathcal{L}(\phi', \partial_\mu\phi', x'^\mu) J\left(\frac{x'}{x}\right) d^4x - \int \mathcal{L}(\phi, \partial_\mu\phi, x^\mu) \, d^4x, \tag{6.2.7}$$

where $J(x'/x)$ denotes the Jacobian - or the scaling factor:

$$J\left(\frac{x'}{x}\right) = \det\left(\frac{\partial x'^\mu}{\partial x^\lambda}\right) = \det\left(\frac{\partial(x^\mu + \delta x^\mu)}{\partial x^\lambda}\right). \tag{6.2.8}$$

So, the variation becomes

$$\delta S = \int \delta\mathcal{L} + \mathcal{L}\partial_\mu(\delta x^\mu) \, d^4x \tag{6.2.9}$$

where

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta(\partial_\mu\phi) + \frac{\partial\mathcal{L}}{\partial x^\mu}\delta x^\mu. \tag{6.2.10}$$

Now, because $\delta(\partial_\mu\phi) = \partial_\mu(\delta\phi)$, the action variation becomes

$$\delta S = \int \left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\mu(\delta\phi) + \frac{\partial\mathcal{L}}{\partial x^\mu}\delta x^\mu + \mathcal{L}\partial_\mu(\delta x^\mu)\right] d^4x \tag{6.2.11}$$

$$= \int \left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\mu(\delta\phi) + \left(\frac{\partial\mathcal{L}}{\partial x^\mu}\delta x^\mu + \mathcal{L}\partial_\mu(\delta x^\mu)\right)\right] d^4x \tag{6.2.12}$$

$$= \int \left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\mu(\delta\phi) + \partial_\mu(\mathcal{L}\delta x^\mu)\right] d^4x. \tag{6.2.13}$$

Next, let us rewrite the second term in terms of the reverse-product rule:

$$\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\mu(\delta\phi) = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\right) - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\delta\phi. \tag{6.2.14}$$

Assume that we are integrating over some region R in spacetime, the action variation becomes

$$\delta S = \int_R \left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\right) - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\delta\phi + \partial_\mu(\mathcal{L}\delta x^\mu)\right] d^4x \tag{6.2.15}$$

$$= \int_R \left\{\delta\phi\left[\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\right] + \partial_\mu\left[\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\right) + (\mathcal{L}\delta x^\mu)\right]\right\} d^4x. \tag{6.2.16}$$

Gauss' theorem says that integration over a divergence (recall that ∂_μ denotes divergence) of a field over a region is equal to the integration of that field over the boundary of that region, so

$$\delta S = \int_R \delta\phi \left[\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) \right] d^4x + \int_{\partial R} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi + \mathcal{L}\delta x^\mu \right] d\sigma_\mu.$$
(6.2.17)

At this point, there are two routes to take. (1) If we restrict the variation to zero at the boundaries, we will end up with the Euler-Lagrange equations, which comes from setting the integrand of the first integral to zero:

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) = 0.$$
(6.2.18)

(2) The other route we can take is not requiring the variation to be zero at the boundary. Doing a small "add and subtract" trick to the second integrand:

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi + \mathcal{L}\delta x^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} [\delta\phi + (\partial_\nu \phi)\delta x^\nu] + \mathcal{L}\delta x^\mu - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi \delta x^\nu$$
(6.2.19)

$$= \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} [\delta\phi + (\partial_\nu \phi)\delta x^\nu] - \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi - \mathcal{L}\delta^\mu_\nu \right] \delta x^\nu.$$
(6.2.20)

Recall that the total variation is defined as

$$\Delta\phi \approx \delta\phi + (\partial_\mu \phi)\delta x^\mu.$$
(6.2.21)

We define the second bracketed term as the *energy-momentum tensor* (we will justify this later):

$$\theta^\mu_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi - \mathcal{L}\delta^\mu_\nu.$$
(6.2.22)

So, once again, the action variation becomes:

$$\delta S = \int_R \delta\phi \left[\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) \right] d^4x + \int_{\partial R} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \Delta\phi - \theta^\mu_\nu \delta x^\nu \right] d\sigma_\mu.$$
(6.2.23)

Let the infinitesimal transformations be

$$\Delta\phi = \Phi_\nu \delta\omega^\nu$$
(6.2.24)

$$\Delta x^\mu = X^\mu_\nu \delta\omega^\nu \approx \delta x^\mu,$$
(6.2.25)

where Φ_ν^μ is a matrix and Φ_ν is just a row vector. By requiring that $\delta S = 0$ and requiring that the Euler-Lagrange equation hold true, we get

$$\int_{\partial R} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \Delta\phi - \theta_\nu^\mu \delta x^\nu \right] d\sigma_\mu = 0 \qquad (6.2.26)$$

$$\int_{\partial R} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \Phi_\nu \delta\omega^\nu - \theta_\nu^\mu X_\mu^\nu \delta\omega^\nu \right] d\sigma_\mu = 0 \qquad (6.2.27)$$

$$\int_{\partial R} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \Phi_\nu - \theta_\nu^\mu X_\mu^\nu \right] \delta\omega^\nu \, d\sigma_\mu = 0. \qquad (6.2.28)$$

Let us define

$$J_\nu^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \Phi_\nu - \theta_\nu^\mu X_\mu^\nu. \qquad (6.2.29)$$

Now, because

$$\int_{\partial R} J_\nu^\mu \delta\omega^\nu \, d\sigma_\mu = 0 \qquad (6.2.30)$$

must hold for any arbitrary $\delta\omega^\nu$, we require

$$\int_{\partial R} J_\nu^\mu \, d\sigma_\mu = 0. \qquad (6.2.31)$$

Now, recall Gauss' theorem one more time:

$$\int_{\partial R} J_\nu^\mu \, d\sigma_\mu = \int_R \partial_\mu J_\nu^\mu \, d^4x. \qquad (6.2.32)$$

This means J_ν^μ is *divergence-free*, i.e. J_ν^μ is a *conserved* quantity:

$$\partial_\mu J_\nu^\mu = 0. \qquad (6.2.33)$$

We can think of J_ν^μ as *current*, whose existence is invariant under the given transformations. We can also calculate another conversed quantity called the "charge"

$$Q_\nu = \int_\sigma J_\nu^\mu \, d\sigma_\mu. \qquad (6.2.34)$$

Let's look at $\mu = 0$, i.e. assuming t is constant:

$$Q_\nu = \int_V J_\nu^0 \, d^3x. \qquad (6.2.35)$$

Now, revisit Gauss' theorem:

$$\int_V \partial_0 J_\nu^0 \, d^3x + \int_V \partial_i J_\nu^i \, d^3x = 0 \qquad (6.2.36)$$

$$\int_V \partial_0 J_\nu^0 \, d^3x + 0 = 0 \qquad (6.2.37)$$

$$\frac{d}{dt} \int_V J_\nu^0 \, d^3x = \frac{dQ_\nu}{dt} = 0. \qquad (6.2.38)$$

So *charge* is conserved over time. This is the essence of *Noether's theorem*.

Finally, let us justify the definition of θ^μ_ν as the energy-momentum tensor. We require that the laws of physics remain the same translationally, and the same in all time. So, let see what we get if we make the transformations

$$X^\mu_\nu = \delta^\mu_\nu \tag{6.2.39}$$

$$\Phi_\mu = 0. \tag{6.2.40}$$

Now recall the definition of J^μ_ν, then apply the transformations to the definition:

$$J^\mu_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \Phi_\nu - \theta^\mu_\nu X^\nu_\mu \tag{6.2.41}$$

$$= -\theta^\mu_\nu \delta^\nu_\mu \tag{6.2.42}$$

$$= -\theta^\mu_\nu. \tag{6.2.43}$$

And so the conservation law, by taking $\mu = 0$, is

$$\frac{d}{dt} \int \theta^0_\nu \, d^3x = \frac{d}{dt} P_\nu = 0. \tag{6.2.44}$$

Let us calculate the first component P_0 from the definition of θ^μ_ν, to show (partly) that P_ν is the 4-momentum:

$$P_0 = \int \theta^0_0 \, d^3x = \int \left\{ \frac{\partial \mathcal{L}}{\partial(\partial_0 \phi)} - \mathcal{L} \right\} d^3x \tag{6.2.45}$$

$$= \int \left\{ \frac{\partial \mathcal{L}}{\partial \dot\phi} \dot\phi - \mathcal{L} \right\} d^3x. \tag{6.2.46}$$

The right-hand side the energy of the field. Next, we can show

$$\int \theta^0_i \, d^3x \tag{6.2.47}$$

is the momentum from the fact that $\partial \phi / \phi x^\mu$ is a 4-vector under Lorentz transformations.

Now, if we had assumed that the Lagrangian hadn't involved x^μ, then we would have ended at the Euler-Lagrange equation, i.e. the system does not exchange energy and momentum with the outside. We condense that with the following proposition:

Proposition 6.2.1. Conservation of energy and momentum holds for a system whose Lagrangian does not depend of x^μ.

Now, let us look at the relationship between the energy-momentum tensor θ^μ_ν and the x^μ-independent Lagrangian, which can be given by

$$\mathcal{L} = \frac{1}{2}g^{\mu\nu}(\partial_\mu\phi)(\partial_\nu\phi) - \frac{m^2}{2}\phi^2. \tag{6.2.48}$$

So $\theta^{\mu\nu}$ can be written as (some of the derivations can be found in the previous subsection):

$$\theta^{\mu\nu} = g^{\lambda\nu}\theta^\mu_\lambda \tag{6.2.49}$$

$$= g^{\lambda\nu}\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\lambda\phi - \delta^\mu_\lambda\mathcal{L}\right) \tag{6.2.50}$$

$$= g^{\lambda\nu}(g^{\mu\lambda}(\partial_\lambda\phi) - \delta^\mu_\lambda\mathcal{L}) \tag{6.2.51}$$

$$= g^{\lambda\nu}(\partial^\mu\phi\,\partial_\lambda\phi - \delta^\mu_\lambda\mathcal{L}) \tag{6.2.52}$$

$$= (\partial^\nu\phi)(\partial^\mu\phi) - g^{\mu\nu}\mathcal{L}. \tag{6.2.53}$$

We observe that μ and ν are exchangeable, hence $\theta^{\mu\nu}$ is symmetric. So, for a scalar field ϕ whose Lagrangian does not exchange energy and momentum with the external, then the energy-momentum tensor $\theta^{\mu\nu}$ is **symmetric**. However, in general $\theta^{\mu\nu}$ is not symmetric, in general, by definition. But, we can define the *canonical energy-momentum tensor* as

$$T^{\mu\nu} = \theta^{\mu\nu} + \partial_\lambda f^{\lambda\mu\nu} \tag{6.2.54}$$

such that $\partial_\mu T^{\mu\nu} = 0$ and $T^{\mu\nu}$ is symmetric. We will not go into detail about this, but the idea is similar to the three-dimensional case in vector calculus where the curl of a vector field is the same as the curl of that same vector field added to a gradient of some other scalar field. By adding the f term, what we wish to accomplish is have T be both divergence-free and symmetric. Now, why do we want the energy-momentum tensor to be symmetric? One of the reasons for this is that in general relativity, Einstein's field equation requires that the energy-momentum stress tensor be symmetric, because the Ricci tensor $R_{\mu\nu}$ and the metric tensor $g_{\mu\nu}$ are both symmetric.

6.3 Noether's Theorem on Symmetries and Conservation Laws

Informally, Noether's theorem states that every differentiable symmetry of the action of a physical system has a corresponding conservation law. In this subsection we will look at slightly different derivation of Noether's (first) theorem.

Consider a general action:

$$S = \int_\Omega d^x\,\mathcal{L}\left(\phi^A(x), \partial_\mu\phi^A(x)\right). \tag{6.3.1}$$

Next, consider infinitesimal spacetime and a transformation:

$$X^\mu \to X'^\mu = X^\mu + \delta X^\mu$$

$$\phi^A(x) \to \phi'^A(x') = \phi'^A(x) + \delta\phi^A(x). \tag{6.3.2}$$

Then

$$\delta S = \int_{\Omega'} d^4x' \, \mathcal{L}\left(\phi'^A(x'), \partial'_\mu \phi'^A(x')\right) - \int_\Omega d^4x \, \mathcal{L}\left(\phi^A(x), \partial_\mu \phi^A(x)\right). \tag{6.3.3}$$

Now we re-label $x' \to x$, since it is just a dummy variable in the integral:

$$\delta S = \int_{\Omega'} d^4x \, \mathcal{L}\left(\phi'^A(x), \partial'_\mu \phi'^A(x)\right) - \int_\Omega d^4x \, \mathcal{L}\left(\phi^A(x), \partial_\mu \phi^A(x)\right)$$

$$= \int_\Omega d^4x \left[\mathcal{L}\left(\phi'^A(x), \partial'_\mu \phi'^A(x)\right) - \mathcal{L}\left(\phi^A(x), \partial_\mu \phi^A(x)\right)\right]$$

$$+ \int_{\Omega'-\Omega} d^4x \, \mathcal{L}\left(\phi'^A(x), \partial_\mu \phi'^A(x)\right). \tag{6.3.4}$$

Note that

$$\int_{\Omega'-\Omega} d^4x = \int_{\delta\Omega} dS_\lambda \, \delta X^\lambda. \tag{6.3.5}$$

So,

$$\int_{\Omega'-\Omega} d^4x \, \mathcal{L}\left(\phi'^A(x), \partial_\mu \phi'^A(x)\right) = \int_{\delta\Omega} dS_\lambda \, \delta X^\lambda \mathcal{L}\left(\phi'^A, \partial_\mu \phi'^A\right), \tag{6.3.6}$$

where to leading order terms:

$$\delta X^\lambda \mathcal{L}\left(\phi'^A, \partial_\mu \phi'^A\right) = \delta X^\lambda \mathcal{L}\left(\phi^A, \partial_\mu \phi^A\right). \tag{6.3.7}$$

By Gauss' law,

$$\int_{\delta\Omega} dS_\lambda \, \delta X^\lambda \mathcal{L}\left(\phi^A, \partial_\mu \phi^A\right) = \int_\Omega d^4x \, \partial_\lambda \left[\delta X^\lambda \mathcal{L}\left(\phi^A, \partial_\mu \phi^A\right)\right], \tag{6.3.8}$$

where ∂_μ denotes the divergence.

So we have, for the last term of δS,

$$\int_{\Omega'-\Omega} d^4x \, \mathcal{L}\left(\phi'^A(x), \partial_\mu \phi'^A(x)\right) = \int_\Omega d^4x \, \partial_\mu \left[\delta X^\mu \mathcal{L}\right] \tag{6.3.9}$$

Next, to simplify the $\mathcal{L}' - \mathcal{L}$ term, we define:

$$\bar{\delta} f(x) = f'(x) - f(x)$$

$$= [f'(x') - f(x)] - [f'(x') - f'(x)]$$

$$= \delta f(x) - \partial_\mu f(x)\delta X^\mu, \tag{6.3.10}$$

where f' denotes a "new" f rather than the derivative of f and

$$\delta f(x) = f'(x') - f(x) \tag{6.3.11}$$

and

$$\begin{aligned}
f'(x') &= f'(x + \delta x) \\
&= f'(x) + \delta X^\mu \partial_\mu f'(x) \\
&\approx f'(x) + \delta X^\mu \partial_\mu f(x).
\end{aligned} \tag{6.3.12}$$

Then we have

$$\begin{aligned}
&\mathcal{L}\left(\phi'^A, \partial'_\mu \phi'^A(x)\right) - \mathcal{L}\left(\phi^A(x), \partial_\mu \phi^A(x)\right) \\
&= \mathcal{L}\left(\phi^A(x) + \bar{\delta}\phi^A(x), \partial_\mu \phi^A + \bar{\delta}\partial_\mu \phi^A\right) - \mathcal{L}\left(\phi^A(x), \partial_\mu \phi^A(x)\right).
\end{aligned} \tag{6.3.13}$$

Since $\bar{\delta}$ and ∂_μ commute, we have

$$\begin{aligned}
&\mathcal{L}\left(\phi'^A, \partial'_\mu \phi'^A(x)\right) - \mathcal{L}\left(\phi^A(x), \partial_\mu \phi^A(x)\right) \\
&= \mathcal{L}\left(\phi^A(x) + \bar{\delta}\phi^A(x), \partial_\mu \phi^A + \partial_\mu \bar{\delta}\phi^A\right) - \mathcal{L}\left(\phi^A(x), \partial_\mu \phi^A(x)\right) \\
&\approx \mathcal{L}\left(\phi^A, \partial_\mu \phi^A\right) + \frac{\partial \mathcal{L}}{\partial \phi^A}\bar{\delta}\phi^A + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\partial_\mu(\bar{\delta}\phi^A) - \mathcal{L}\left(\phi^A, \partial_\mu \phi^A\right) \\
&= \left(\frac{\partial \mathcal{L}}{\partial \phi^A} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\right)\bar{\delta}\phi^A + \left(\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\bar{\delta}\phi^A + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\partial_\mu \phi^A\right) \\
&= \left(\frac{\partial \mathcal{L}}{\partial \phi^A} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\right)\bar{\delta}\phi^A + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\bar{\delta}\phi^A\right),
\end{aligned} \tag{6.3.14}$$

where the last equality comes from doing reverse product rule.

Putting everything together,

$$\begin{aligned}
\delta S &= \int_\Omega d^4x \left[\left(\frac{\partial \mathcal{L}}{\partial \phi^A} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\right)\bar{\delta}\phi^A + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\bar{\delta}\phi^A\right) + \partial_\mu\left(\delta X^\mu \mathcal{L}\right)\right] \\
&= \int_\Omega d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi^A} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\right)\bar{\delta}\phi^A + \int_\Omega d^4x\, \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\bar{\delta}\phi^A + \mathcal{L}\delta X^\mu\right].
\end{aligned} \tag{6.3.15}$$

Now, use $\bar{\delta}\phi^A \equiv \delta\phi^A - \partial_\mu \phi^A \delta X^\mu$, then

$$\begin{aligned}
&\partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\bar{\delta}\phi^A + \mathcal{L}\delta X^\mu\right] \\
&= \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\delta\phi^A - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}(\partial_\nu \phi^A)(\delta X^\nu) + \mathcal{L}\delta X^\mu\right] \\
&= \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\delta\phi^A\right) - \partial_\mu \left[\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\partial^\nu \phi^A - \eta^{\mu\nu}\mathcal{L}\right)\delta X_\nu\right].
\end{aligned} \tag{6.3.16}$$

Let us call

$$\boxed{T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\partial^\nu \phi^A - \eta^{\mu\nu}\mathcal{L}} \tag{6.3.17}$$

the **energy-momentum stress tensor**. Then the variation in the action becomes

$$\delta S = \int_\Omega d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi^A} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)} \right)\bar{\delta}\phi^A + \int_\Omega d^4x\, \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\delta\phi^A - T^{\mu\nu}\delta X_\nu \right). \tag{6.3.18}$$

Let us define the **4-current**

$$\boxed{J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\delta\phi^A - T^{\mu\nu}\delta X_\nu} \tag{6.3.19}$$

Then we get

$$\boxed{\delta S = \int_\Omega d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi^A} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)} \right)\bar{\delta}\phi^A + \int_\Omega d^4x\, \partial_\mu J^\mu} \tag{6.3.20}$$

where $\partial_\mu J^\mu$ is the divergence of J.

If we require the action be invariant under transformations such as

$$\phi^A \to \phi^A + \delta\phi^A \tag{6.3.21}$$

and/or

$$X^\mu \to X^\mu + \delta X^\mu \tag{6.3.22}$$

then $\delta S = 0$. And if ϕ^A is on shell and obeys the equations of motion, then the integrand in the first term of the action is also zero (this is just the Euler-Lagrange equation)

$$\frac{\partial \mathcal{L}}{\partial \phi^A} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)} = 0. \tag{6.3.23}$$

These together give a conserved current, i.e. the four-current is divergence free:

$$\boxed{\partial_\mu J^\mu = 0} \tag{6.3.24}$$

This is a major stepping stone towards the result of Noether's first theorem:

Theorem 6.3.1. When a theory has a symmetry and the equations of motion hold, there exists a conserved quantity.

So what is this conserved quantity in the theory we are working with? Consider $\partial_\mu J^\mu = 0$, then

$$\int d^3x \, \partial_\mu J^\mu = 0$$

$$\int d^3x \left(\partial_0 J^0 + \partial_j J^j \right) = 0$$

$$\frac{d}{dt} \int d^3x \, J^0 + \int d^3x \, \partial_j J^j = 0$$

$$\frac{d}{dt} \int d^3x \, J^0 + \int d^3x \, \boldsymbol{\nabla} \cdot \vec{J} = 0. \tag{6.3.25}$$

Let $J^\mu = (\rho, \vec{J})$, where ρ is the **charge density**, then

$$\frac{d}{dt} \int d^3x \, \rho + \int d^3x \, \boldsymbol{\nabla} \cdot \vec{J} = 0$$

$$\frac{dQ}{dt} + \int_S d\vec{A} \cdot \vec{J} = 0, \tag{6.3.26}$$

where we have used Gauss' law on the second term. Now, since we want $\vec{J} \to \vec{0}$ on the boundary for the field to be physical, the second term on the left hand side is just going to be zero. This implies that

$$\boxed{\frac{dQ}{dt} = 0} \tag{6.3.27}$$

We have the conservation of charge.

In the following subsubsection, we shall consider a few illuminating examples to see how for every symmetry, we can obtain a conserved quantity, as Noether's theorem says.

6.3.1   Space-time translations

The action constructed from relativistic fields is invariant under Lorentz transformations as well as translations of the coordinates. Consider the translation transformation:

$$x^\mu \to x^\mu + a^\mu, \tag{6.3.28}$$

where a^μ is constant. Since the fields don't change at any point, we require that

$$\delta\phi^A = 0. \tag{6.3.29}$$

Since we have shown

$$\partial_\mu J^\mu = 0 \tag{6.3.30}$$

has to hold, the definition

$$J^\mu = \frac{\partial \mathcal{L}}{\partial \partial(\partial_\mu \phi^A)} \delta \phi^A - T^{\mu\nu} \delta x_\nu \tag{6.3.31}$$

gives

$$\partial_\mu T^{\mu\nu} = 0. \tag{6.3.32}$$

This gives

$$P^\nu = \frac{d}{dt} \int d^3x \, T^{0\nu} = \int d^3x \, \partial_0 T^{0\nu} = -\int d^3x \, \partial_i T^{i\nu} = 0, \tag{6.3.33}$$

i.e., we get a conservation law since we require $T^{i\nu}$ to vanish at $x \to \infty$:

$$\boxed{P^\mu = \int d^3x \, T^{0\mu}} \tag{6.3.34}$$

which is nothing but the conservation of the 4-momentum of the field.

6.3.2   Lorentz transformations

Consider the infinitesimal Lorentz transformations:

$$x'^\mu = x^\mu + \omega^{\mu\nu} x_\nu, \tag{6.3.35}$$

where $\omega^{\mu\nu}$ is independent of x^μ. To keep $x^\mu x_\mu$ invariant, we require that $\omega^{\mu\nu}$ antisymmetric, i.e.,

$$\omega^{\mu\nu} = -\omega^{\nu\mu}. \tag{6.3.36}$$

To express the field variation $\delta \phi^A$ under this infinitesimal transformation, we have to use the spin matrix:

$$\Sigma^{\lambda\rho}, \tag{6.3.37}$$

defined as

$$\textbf{INSERT DEFINITION HERE}. \tag{6.3.38}$$

THIS PART IS MISSING...

6.3.3   Internal symmetries

Internal symmetries are those which relate different fields at the same space-time point. **EXPLAIN MORE** In this case, $\delta x^\mu = 0$. Consider th infinitesimal transformation

$$\phi^A(x) \rightarrow \phi^A + f^A_\tau(x)\delta\epsilon_\tau, \qquad r = 1, 2, \ldots, p, \qquad (6.3.39)$$

under which the action is invariant, where $delta\epsilon_\tau$ are infinitesimal parameters independent of space-time, and $f^A_\tau(x)$ are specified functions of the fields ϕ^A and their derivatives. The index τ is not summed over, but rather it indicates the type of symmetry. There maybe several independent symmetries in a system. We can treat them separately by defining the conserved current for the rth symmetry as

$$J^\mu_r = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)} \frac{\delta\phi^A}{\delta\epsilon_r}. \qquad (6.3.40)$$

I STILL DON'T FULLY UNDERSTAND THESE CONCEPTS...

Chapter 7

Spontaneous Symmetry Breaking

7.1   Introduction

Spontaneous symmetry breaking is a mechanism where symmetry still holds dynamically but the solutions break symmetry. In other words, spontaneous symmetry breaking is a process in which a physical system in a symmetric state ends up in an asymmetric state. In particular, it can describe systems where the equations of motion or the Lagrangian obey symmetries, but the lowest-energy vacuum solutions do not exhibit that same symmetry. When the system goes to one of those vacuum solutions, the symmetry is broken for perturbations around that vacuum even though the entire Lagrangian retains that symmetry.

Consider the Lagrangian of a real scalar field ϕ:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - V(\phi). \tag{7.1.1}$$

Suppose that the potential $V(\phi)$ is invariant under party transformations, i.e.,

$$V(\phi) = V(-\phi), \tag{7.1.2}$$

Then by inspection the Lagrangian is also invariant under parity transformations:

$$\mathcal{L}' = \frac{1}{2}(-\partial_\mu \phi)(-\partial^\mu \phi) - V(-\phi) \tag{7.1.3}$$

$$= \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - V(\phi) \tag{7.1.4}$$

$$= \mathcal{L}. \tag{7.1.5}$$

So symmetry is still preserved. We can consider such a potential $V(\phi)$ such that $V(\phi) = V(-\phi)$:

$$V(\phi) = \frac{1}{2}m^2\phi^2 + \frac{1}{4}\lambda\phi^4 \tag{7.1.6}$$

where $m^2 > 0, \lambda > 0$. If we plot $V(\phi)$ versus ϕ, we see that there is a unique minimum at $\phi = 0$.
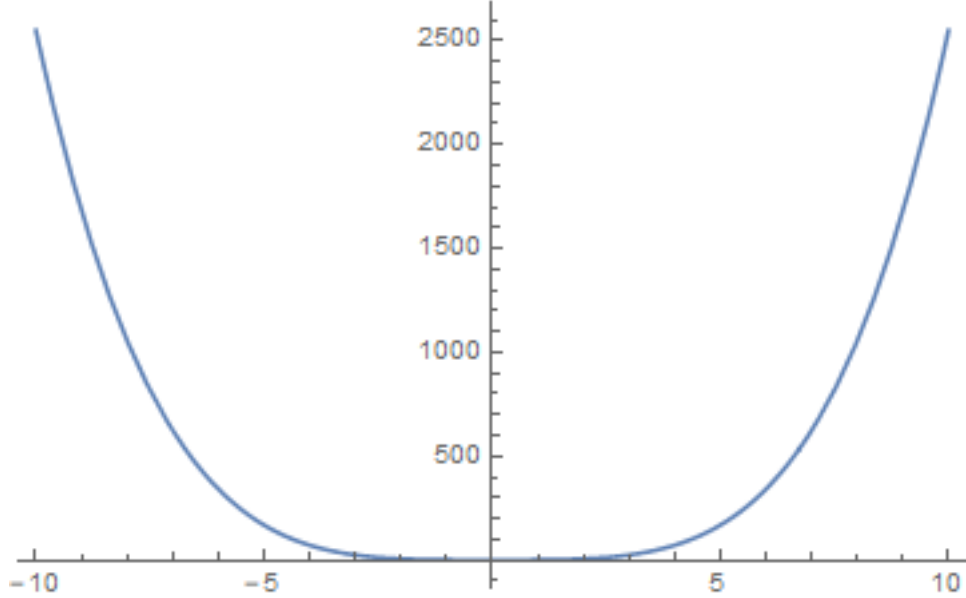
Figure 7.1: Plot of potential $V(\phi \cdot \phi)$ with $m^2 > 0$. We notice a unique global minimum.

The ground state, or the state of lowest energy, in field theory is called the **vacuum expectation value**, denoted $\langle\phi\rangle$. In our example the state of lowest energy has zero energy, so

$$\langle\phi\rangle = 0. \tag{7.1.7}$$

Now, we look at whether at $\phi = 0$ there is still symmetry. To do this, we once again use variational methods, i.e., we look at small excitations around the vacuum:

$$\phi = \langle\phi\rangle + \epsilon = 0 + \epsilon = \epsilon. \tag{7.1.8}$$

In which case, the Lagrangian becomes

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\epsilon)(\partial^\mu\epsilon) + \frac{1}{2}m^2\epsilon^2 + \dots \tag{7.1.9}$$

where the higher order terms are neglected because we're only concerned with infinitesimal variations. Now, by the look of the Lagrangian, we know that the theory is describing a massive particle in real scalar field. So, we have reasons to suspect that excitations in a field gives rise to a particle. We will explore this idea much more deeply as we move on.

Now, let us suppose that

$$V(\phi) = -\frac{1}{2}m^2\phi^2 + \frac{1}{4}\lambda\phi^4 \tag{7.1.10}$$

then the plot of $V(\phi)$ versus ϕ becomes Figure 7.2.
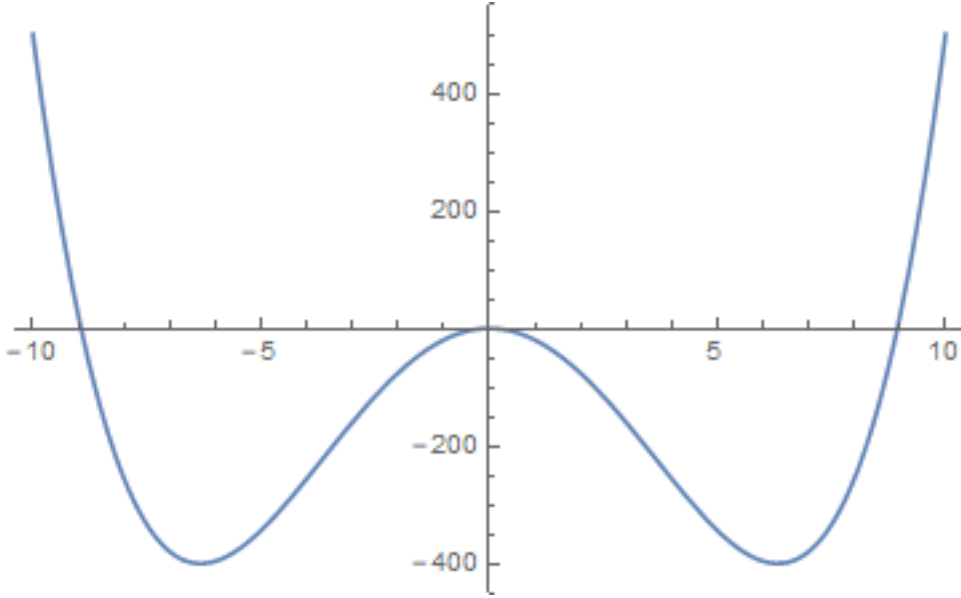
Figure 7.2: Plot of potential $V(\phi\cdot\phi)$ with $m^2 < 0$. We notice two global minima.

We observe that there are two possible vacuum solutions. So, we might wonder, well, which one is the vacuum solution? While either one *can* be a solution, nature spontaneously picks one. Let us look at what which ϕ's does $V(\phi)$ obtain minimal values:

$$\frac{dV}{d\phi} = m^2\phi + \lambda\phi^3 = 0, \tag{7.1.11}$$

i.e.,

$$\langle\phi\rangle = \pm\sqrt{\frac{-m^2}{\lambda}}. \tag{7.1.12}$$

Now suppose that nature picks the positive ϕ:

$$\langle\phi\rangle = \sqrt{\frac{-m^2}{\lambda}} \equiv \mathcal{V}. \tag{7.1.13}$$

Now, we can shift to a field defined with respect to the vacuum expectation value:

$$\phi' = \phi - \langle\phi\rangle = \phi - \mathcal{V}. \tag{7.1.14}$$

Then

$$\langle\phi'\rangle = 0. \tag{7.1.15}$$

In terms of the Lagrangian, the new Lagrangian is

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi')(\partial^\mu\phi') - (-m^2)\left[\frac{\phi'^4}{4\mathcal{V}^2} + \frac{\phi'^3}{\mathcal{V}} + \phi'^2 - \frac{\mathcal{V}^2}{4}\right], \tag{7.1.16}$$

which has no symmetry in terms of ϕ' (parity transformation). We say that the symmetry is hidden, since we always shift back to get symmetry.

Now, we can look at small excitations about $\langle\phi'\rangle$:

$$\phi' = \langle\phi'\rangle + \epsilon = 0 + \epsilon = \epsilon. \tag{7.1.17}$$

Plugging into the Lagrangian, we get

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\epsilon)(\partial^\mu\epsilon) - \frac{1}{2}(-2m^2)\epsilon^2, \tag{7.1.18}$$

which acts as a massive particle scalar field with mass (note that we are still assuming $-2m^2 > 0$). Again, we see that by breaking symmetry, we get a massive particle. We shall verify this. Recall the Lagrangian:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \frac{1}{2}m^2\phi^2 - \frac{1}{4}\lambda\phi^4. \tag{7.1.19}$$

Assume that

$$\langle\phi\rangle = \sqrt{\frac{-m^2}{\lambda}} = \mathcal{V}. \tag{7.1.20}$$

Now let

$$\phi = \langle\phi\rangle + \epsilon = \mathcal{V} + \epsilon, \tag{7.1.21}$$

where $\mathcal{V}$ is a constant, this gives

$$\partial_\mu\phi = \partial_\mu\epsilon. \tag{7.1.22}$$

Putting everything together, the new $V(\phi)$ is

$$V = \frac{1}{2}m^2(\mathcal{V} + \epsilon)^2 + \frac{1}{4}\lambda(\mathcal{V} + \epsilon)^2 \tag{7.1.23}$$

$$= \frac{1}{2}m^2(\mathcal{V}^2 + 2\mathcal{V}\epsilon + \epsilon^2) + \frac{1}{4}\lambda(\mathcal{V}^4 + 4\mathcal{V}^3\epsilon + 6\mathcal{V}^2\epsilon^2 + 4\mathcal{V}\epsilon^3 + \epsilon^4). \tag{7.1.24}$$

Keeping the linear terms

$$V \approx \epsilon(m^2 + \mathcal{V}) + \epsilon^2\left(\frac{1}{2}m^2 + \frac{3}{2}\lambda\mathcal{V}^2\right) + \dots \tag{7.1.25}$$

$$\approx \epsilon\mathcal{V}(m^2 + \lambda\mathcal{V}^2) + \epsilon^2(-m^2) \tag{7.1.26}$$

$$\approx \epsilon\mathcal{V}\left(m^2 - \lambda\frac{m^2}{\lambda}\right) + \epsilon^2(-m^2). \tag{7.1.27}$$

So,

$$V(\epsilon) \approx -\epsilon^2 m^2 = \frac{1}{2}(-2m^2)\epsilon^2. \tag{7.1.28}$$

So the new Lagrangian is verified, as desired. We observe that $V(\epsilon)$ is symmetric around 0.

We observe that under a parity transformation, we still get symmetry. We call this **discrete symmetry**, to distinguish from **continuous symmetry** under smooth/continuous transformation.

Theorem 7.1.1. Goldstone's Theorem: In a theory with a continuous symmetry that is spontaneously broken, then there will be a massless particle, referred to as **Nambu-Goldstone mode**.

To illustrate the essence of this theorem, we consider two scalar fields written as a vector

$$\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}, \tag{7.1.29}$$

with an associated Lagrangian defined as

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi) \cdot (\partial^\mu\phi) - V(\phi \cdot \phi). \tag{7.1.30}$$

This theory has a global $O(2)$ symmetry (continuous):

$$\phi' = R\phi \tag{7.1.31}$$

where R is just a rotation matrix, by a (continuous) angle θ:

$$R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{7.1.32}$$

Under R, the dot product is invariant:

$$\phi'^\top \cdot \phi' = (R\phi)^\top (R\phi) = \phi^\top R^\top R\phi = \phi^\top \phi. \tag{7.1.33}$$

So it is clear that the Lagrangian is invariant, as expected:

$$\mathcal{L}' = \mathcal{L}. \tag{7.1.34}$$

Now, suppose that

$$V(\phi \cdot \phi) = \frac{1}{2}m^2 \phi \cdot \phi + \frac{1}{4}\lambda(\phi \cdot \phi)^2. \tag{7.1.35}$$

If $m^2 > 0$, then we can plot the potential $V(\phi)$: Again, we see that there is a
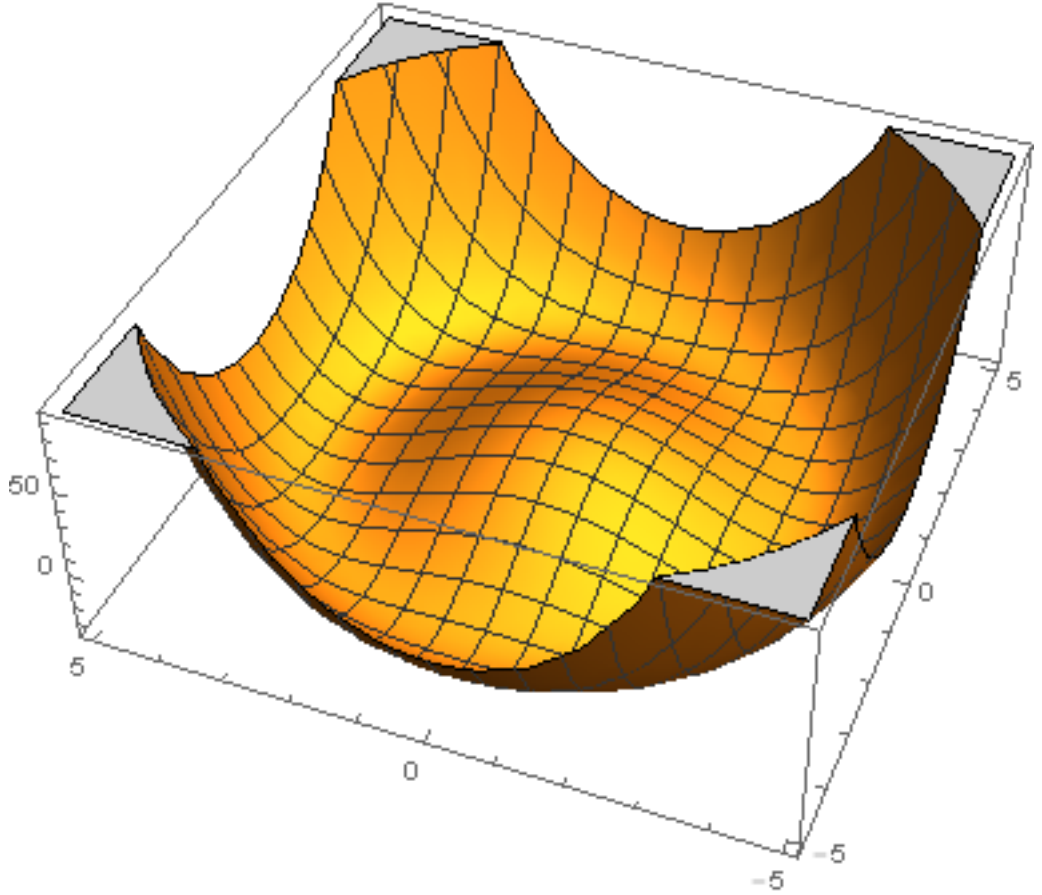
unique vacuum solution:

$$\langle \phi \rangle = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{7.1.36}$$

But what if $m^2 < 0$, we get a similar situation at last time, but now in three dimensions where we also have a circle of possible ground state. Now, let nature spontaneously picks a vacuum. We can pick

$$\langle \phi \rangle = \begin{pmatrix} \mathcal{V} \\ 0 \end{pmatrix}. \tag{7.1.37}$$

Again, let us look at an excitation about ϕ and shift:

$$\phi' = \phi - \langle \phi \rangle = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} - \begin{pmatrix} \mathcal{V} \\ 0 \end{pmatrix} = \begin{pmatrix} \phi_1 - \mathcal{V} \\ \phi_2 \end{pmatrix}. \tag{7.1.38}$$

For small excitations, we can approximate:

$$\phi' = \langle \phi' \rangle + \epsilon = \langle \phi' \rangle + \begin{pmatrix} \eta \\ \xi \end{pmatrix} = \begin{pmatrix} \eta \\ \xi \end{pmatrix}. \tag{7.1.39}$$

Note that the shift makes $\langle \phi \rangle = 0$.

We can express the Lagrangian in terms of these, and find that

1. One field is massless, corresponding to the Nambu-Goldstone mode.

2. The other is massive, corresponding to the Higgs particle.

7.2 Continuous Global Symmetry

Consider two scalar fields ϕ_1 and ϕ_2, written as

$$\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}. \tag{7.2.1}$$

And consider the transformation $\phi' \to R\phi$ where R is a rotation of angle θ where θ does not depend on x. We can look at the Lagrangian:

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi \cdot \partial^\mu \phi - V(\phi \cdot \phi). \tag{7.2.2}$$

For $m^2 < 0$, then

$$V(\phi \cdot \phi) = \frac{1}{2} m^2 \phi^2 + \frac{1}{4} \lambda \phi^2 \tag{7.2.3}$$

has minimal value at

$$\langle \phi \rangle^2 = \frac{-m^2}{\lambda} \equiv \mathcal{V}^2. \tag{7.2.4}$$

Let's pick

$$\langle \phi \rangle = \begin{pmatrix} \mathcal{V} \\ 0 \end{pmatrix} \tag{7.2.5}$$

and do a shift

$$\phi' = \phi - \langle \phi \rangle \tag{7.2.6}$$

so that

$$\langle \phi' \rangle = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{7.2.7}$$

Let us consider excitations around the vacuum:

$$\phi' = \langle \phi \rangle + \begin{pmatrix} \eta \\ \zeta \end{pmatrix} = \begin{pmatrix} \eta \\ \zeta \end{pmatrix} \tag{7.2.8}$$

and (again) a shift

$$\phi = \langle \phi \rangle + \phi' = \begin{pmatrix} \mathcal{V} + \eta \\ \zeta \end{pmatrix} \tag{7.2.9}$$

which gives

$$\phi \cdot \phi = (\mathcal{V} + \eta)^2 + \zeta^2 \tag{7.2.10}$$

and

$$\partial_\mu \phi = \partial_\mu \phi' = \begin{pmatrix} \partial_\nu \eta \\ \partial_\mu \zeta \end{pmatrix}. \tag{7.2.11}$$

We can now write the Lagrangian in terms of the excitations, dropping cubic and higher order terms. First, we find what the new potential looks like:

$$V(\phi \cdot \phi) = \frac{1}{2}m^2\phi^2 + \frac{1}{2}\lambda\phi^4 \tag{7.2.12}$$

$$= \frac{1}{2}m^2 \left[(\mathcal{V} + \eta)^2 + \zeta^2 \right] + \frac{1}{4}\lambda \left[(\mathcal{V} + \eta)^2 + \zeta^2 \right]^2 \tag{7.2.13}$$

$$= \frac{1}{2}m^2 \left[\mathcal{V}^2 + 2\mathcal{V}\eta + \eta^2 + \zeta^2 \right] + \frac{1}{4}\lambda \left[\mathcal{V}^2 + 2\mathcal{V}\eta + \eta^2 + \zeta^2 \right]^2 \tag{7.2.14}$$

$$= \frac{1}{2}m^2 \left[\mathcal{V}^2 + 2\mathcal{V}\eta + \eta^2 + \zeta^2 \right] + \frac{1}{4}\lambda \left[4\mathcal{V}^3\eta + 4\mathcal{V}^2\eta^2 + 2\mathcal{V}^2\zeta^2 + 2\mathcal{V}^2\eta^2 + \dots \right] \tag{7.2.15}$$

$$= \eta \left[\frac{1}{2}m^2 2\mathcal{V} + \frac{1}{4}\lambda 4\mathcal{V}^3 \right] + \eta^2 \left[\frac{1}{2}m^2 + \frac{3}{2}\lambda\mathcal{V}^2 \right] + \zeta^2 \left[\frac{1}{2}m^2 + \frac{\lambda}{2}\mathcal{V}^2 + \dots \right] + \dots \tag{7.2.16}$$

Recall that we have defined the minimum:

$$\mathcal{V}^2 = -\frac{m^2}{\lambda}. \tag{7.2.17}$$

This gives

$$V(\phi^2) = \mathcal{V}\eta(m^2 + \lambda\mathcal{V}^2) + \eta^2 \left[\frac{1}{2}m^2 - \frac{3}{2}m^2\right] + \frac{1}{2}\eta^2(m^2 - m^2). \qquad (7.2.18)$$

To second order, this becomes

$$V(\phi^2) = -m^2\eta^2 = \frac{1}{2}(-2m^2\eta^2). \qquad (7.2.19)$$

So, the Lagrangian is

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi \cdot \phi) \qquad (7.2.20)$$

$$= \frac{1}{2}\begin{pmatrix}\partial_\mu\eta & \partial^\mu\zeta\end{pmatrix}\begin{pmatrix}\partial^\mu\eta \\ \partial^\mu\zeta\end{pmatrix} - \left[\frac{1}{2}(-2m^2\eta^2)\right] + \dots \qquad (7.2.21)$$

$$= \frac{1}{2}\left[\partial_\mu\eta\partial^\mu\eta + \partial_\mu\zeta\partial^\mu\zeta\right] - \frac{1}{2}(-2m^2)\eta^2 + \dots \qquad (7.2.22)$$

Notice that we started with two scalars $\phi = \begin{pmatrix}\phi_1 & \phi_2\end{pmatrix}^\top$ and the wrong sign $m^2 < 0$. But after spontaneous symmetry breaking and a physical vacuum excitation, we have

$$\mathcal{L} = \boxed{\frac{1}{2}\partial_\mu\eta\partial^\mu\eta - \frac{1}{2}(-2m^2)\eta^2} + \boxed{\frac{1}{2}\partial_\mu\zeta\partial^\mu\zeta} \qquad (7.2.23)$$

Observe that we a get 1 massive scalar field η with mass $-2m^2 > 0$ and a massless scalar ζ. The massive scalar field is what we will ultimately describe as the **Higgs boson**, while the massless field describes the Nambu-Goldstone mode. Let us revisit the Goldstone theorem:

Theorem 7.2.1. Goldstone: For every continuous global symmetry that is spontaneously broken, there emerges a massless particle.

For example, in electromagnetism, there is a U(1) local gauge symmetry. We have shown earlier that this gives rise to the massless photon.

There are slight "exceptions" to the rule, though. Some issues might arise when we try to describe the weak interaction. We would like to describe the force-carrying particles in the weak interactions as a gauge theory. SU(2) gives us 3 gauge fields. However, since the weak interaction is too weak and short-ranged, it was suspected that the weak force is carried by 3 massive vector fields. But, the problem is that, according to Goldstone's theorem, we can't have both a gauge symmetry and massive terms for the gauge fields.

Remark 7.2.1. Any interacting massless particle is detectable because it's got a long-range interaction. This implies that particles in Nambu-Goldstone modes are detectable.

7.3   An O(2) example

So far, we have looked at spontaneous symmetry breaking of of a global gauge theory. Now, we will look at spontaneous symmetry breaking of a local gauge theory. This subsection is essentially about Higgs and others found a mechanism where the massless Nambu-Goldstone modes get "eaten" and the gauge fields acquire mass.

Consider a local O(2) transformation and consider (once again) 2 real scalar fields ϕ_1, ϕ_2, written as a vector.

$$\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}. \tag{7.3.1}$$

The transformation is essentially a rotation, we shall call R. For the sake of simplicity, we can take R as a 2×2 matrix. So, the transformation has the form:

$$\phi' = R(x)\phi = \begin{pmatrix} \cos\alpha(x) & -\sin\alpha(x) \\ \sin\alpha(x) & \cos\alpha(x) \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}, \tag{7.3.2}$$

where the angle of rotation α depends on the coordinates (hence local). Also note that since the determinant of R is 1, this transformation can also be classified as a special orthogonal transformation, belonging to SO(2), an abelian group.

To greatly simplify our notation, we can write the transformation $\phi \to \phi'$ as

$$R = e^{i\alpha(x)T} \tag{7.3.3}$$

where T is a 2×2 generator. For this particular transformation, T is Hermitian

$$T = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}. \tag{7.3.4}$$

It might be useful later to observe that

$$T^2 = T \tag{7.3.5}$$

$$T^\top = -T \tag{7.3.6}$$

$$T^\dagger = T. \tag{7.3.7}$$

For small $\alpha(x)$, the transformation reduces into a much simpler form:

$$R \approx I + i\alpha(x)T \tag{7.3.8}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + i\alpha(x) \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} = \begin{pmatrix} 1 & -\alpha \\ \alpha & 1 \end{pmatrix}. \tag{7.3.9}$$

Let us return to our original Lagrangian:

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi \cdot \partial^\mu\phi - V(\phi \cdot \phi). \tag{7.3.10}$$

Let $\phi \to \phi' = R(x)\phi$, we get

$$\partial_\mu\phi = \partial_\mu\phi' = R\partial_\mu\phi + (\partial_\mu R)\phi. \tag{7.3.11}$$

Notice that there is no local symmetry here, which means we ought to fix this by changing the derivative (as before), i.e., we need to define a gauge-covariant derivative:

$$D_\mu = \partial_\mu + igA_\mu \tag{7.3.12}$$

where g is a coupling constant and A_μ TA_μ is some (matrix) function (note that we are still working with matrices here). Formally, in order for D_μ to act on ϕ, it has to be a 2×2 matrix. So formally,

$$\boxed{D_\mu = I\partial_\mu + TigA_\mu} \tag{7.3.13}$$

where I is the identity matrix, and T is the generator. Now, to obtain gauge invariance, we want $D_\mu\phi \to D'_\mu\phi' = RD_\mu\phi$ so that

$$D_\mu\phi \cdot D^\mu\phi \to (RD_\mu\phi)^\top (RD^\mu\phi) = (D_\mu\phi)^\top R^\top R(D^\mu\phi) \tag{7.3.14}$$
$$= (D_\mu\phi)^\top (D^\mu\phi) \tag{7.3.15}$$
$$= (D_\mu\phi) \cdot (D^\mu\phi). \tag{7.3.16}$$

The question is: how must A_μ transform in order for this to hold? Recall what we have done in the previous sections when we required that A_μ transforms as

$$A'_\mu = A_\mu + \partial_\mu\Lambda \tag{7.3.17}$$

in order to get gauge invariance in electromagnetism. To find how A_μ has to transform here, we follow sort of a similar path, while keeping in mind that A_μ is actually a matrix. To do this, we first look at

$$D'_\mu\phi' = (\partial_\mu + iqA'_\mu)R\phi \tag{7.3.18}$$
$$= R\partial_\mu\phi + (\partial_\mu R)\phi + iqA'_\mu R\phi. \tag{7.3.19}$$

Since we require invariance, we set

$$D'_\mu\phi' = RD_\mu\phi, \tag{7.3.20}$$

i.e., we set

$$R\partial_\mu\phi + (\partial_\mu R)\phi + iqA'_\mu R\phi = R(\partial_\mu iqA_\mu)\phi \tag{7.3.21}$$

which says

$$R\partial_\mu\phi + (\partial_\mu R)\phi + iqA'_\mu R\phi = R\partial_\mu\phi RiqA_\mu\phi, \tag{7.3.22}$$

which is satisfied if

$$igA'_\mu R\phi = igRA_\mu - (\partial_\mu R)\phi. \tag{7.3.23}$$

So,

$$\frac{ig}{ig}A'_\mu R = RA_\mu - \frac{1}{ig}\partial_\mu R \tag{7.3.24}$$

$$A'_\mu RR^{-1} = RA_\mu R^{-1} + \frac{i}{g}(\partial_\mu R)R^{-1}. \tag{7.3.25}$$

But since $R^{-1}R = RR^{-1} = 1$,

$$\boxed{A'_\mu = RA_\mu R^{-1} + \frac{i}{g}(\partial_\mu R)R^{-1}} \tag{7.3.26}$$

This is the transformation rule for A_μ under a local O(2) gauge transformation.

Let us consider a quick educated example. With

$$R = e^{i\alpha(x)T} \tag{7.3.27}$$

$$R^{-1} = e^{-i\alpha(x)T}, \tag{7.3.28}$$

we get

$$A_\mu \rightarrow A'_\mu = RA_\mu R^{-1} + \frac{i}{g}(\partial_\mu R)R^{-1} \tag{7.3.29}$$

$$= e^{i\alpha(x)T}A_\mu e^{-i\alpha(x)T} + \frac{i}{g}\left(i(\partial_\mu\alpha)TR\right)R^{-1}. \tag{7.3.30}$$

Recall that A_μ is a matrix. So, actually A_μ is written as

$$A_\mu \equiv A_\mu T. \tag{7.3.31}$$

So, we actually have

$$A'_\mu T = A_\mu T e^{i\alpha(x)T}e^{-i\alpha(x)T} - \frac{1}{g}(\partial_\mu\alpha(x))T. \tag{7.3.32}$$

So, the function A_μ must obey

$$A'_\mu = A_\mu - \frac{1}{g}(\partial_\mu\alpha). \tag{7.3.33}$$

Looks familiar? You're right, because we have seen this before in studying the Maxwell's equations. If we define

$$\Lambda(x) = -\frac{\alpha(x)}{g} \tag{7.3.34}$$

then we get the familiar transformation rule for A_μ:

$$A'_\mu = A_\mu + \partial_\mu \lambda(x). \tag{7.3.35}$$

Note that, as a bonus we also get

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{7.3.36}$$

which implies an anti-symmetry in the electromagnetic field strength tensor

$$F_{\mu\nu} = -F_{\nu\mu}, \tag{7.3.37}$$

which carries a gauge invariance. Now that we know the transformation rule for A_μ and have defined $F_{\mu\nu}$ in such a way that it is gauge invariant, we can add it to the Lagrangian to make the theory (local $O(2)$ gauge theory, that is) dynamical. The complete Lagrangian for this theory is then

$$\mathcal{L} = \frac{1}{2}(D_\mu \phi) \cdot (D^\mu \phi) - V(\phi \cdot \phi) - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{7.3.38}$$

But note that this is only true for the simple case of a $O(2)$ transformation. To find how A_μ transforms in general, we have to use the general, boxed equation that involves R and R^{-1}.

7.4   Introduction to the Higgs Mechanism

To summarize what we have been doing so far, consider (once again) the potential

$$V(\phi \cdot \phi) = \frac{1}{2}m^2\phi^2 + \frac{1}{4}\lambda(\phi \cdot \phi)^2. \tag{7.4.1}$$

If $m^2 > 0$, then there is spontaneous symmetry breaking, i.e. we get massless A_μ with 2 modes of motion (we have shown this in the worked problems) and 2 massive scalar fields ϕ_1 and ϕ_2, i.e., two Nambu-Goldston modes and two massive modes.

On the other hand, if $m^2 < 0$, we get spontaneous symmetry breaking, and Goldstone theorem says that for a global symmetry, we get one massive Higgs scalar and one massless Nambu-Goldstone mode, totaling two rather than 4.

(Note that this is not true for local symmetry)

Motivated by our recent example, now we will show how spontaneous symmetry breaking in O(2) can give rise to the Higgs mechanism. We will show that in this symmetry breaking process, the Nambu-Goldstone mode gets "eaten" (this is actually a technical term) and $A_\mu \to A'_\mu$, which is massive. Ultimately, we will show that we are left with a massive A'_μ with three modes and a massive Higgs scalar without any Nambu-Goldstone modes, totaling (once again) 4 modes.

Recall the Lagrangian of the form

$$\mathcal{L} = \frac{1}{2}(D_\mu \phi) \cdot (D^\mu \phi) - V(\phi \cdot \phi) - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{7.4.2}$$

with

$$\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \tag{7.4.3}$$

and

$$D_\mu \phi = (\partial_\mu \mathbb{I} - igTA_\mu). \tag{7.4.4}$$

We recall that this has an O(2) invariance with

$$\begin{cases} \phi \to \phi' = R\phi \\ A_\mu \to A_{-\mu} = RA_\mu R^{-1} + \frac{i}{g}(\partial_\mu R)R^{-1}. \end{cases} \tag{7.4.5}$$

To see how the Higgs mechanism comes about, we shall consider an illustrating example in O(2) with a re-parameterization of ϕ. We shall write

$$\phi = R^{-1}\phi = R^{-1}\begin{pmatrix} 0 \\ \mathcal{V} + \mathcal{E} \end{pmatrix} \tag{7.4.6}$$

and let

$$\alpha = \frac{\zeta}{\mathcal{V}}, \tag{7.4.7}$$

where these terms are defined in the previous subsection. With this, we get

$$R^{-1} = e^{-i\alpha T} \approx \begin{pmatrix} 1 & \alpha \\ -\alpha & 1 \end{pmatrix} = \begin{pmatrix} 1 & \zeta/\mathcal{V} \\ -\zeta/\mathcal{V} & 1 \end{pmatrix}. \tag{7.4.8}$$

This gives us a new expression for ϕ:

$$\phi = R^{-1}\phi' = \begin{pmatrix} 1 & \zeta/\mathcal{V} \\ -\zeta/\mathcal{V} & 1 \end{pmatrix}\begin{pmatrix} 0 \\ \mathcal{V} + \mathcal{E} \end{pmatrix} \approx \begin{pmatrix} \zeta \\ \mathcal{V} + \mathcal{E} \end{pmatrix}. \tag{7.4.9}$$

This is in fact the expression we obtained earlier in the previous subsection, where $\mathcal{E}$ denotes a small excitation. This re-parameterization is called the "Unitary Gauge."

Now, let us put this new ϕ into the Lagrangian:

$$\mathcal{L} = \frac{1}{2} D_\mu (R^{-1}\phi') \cdot D^\mu (R^{-1}\phi') - V(R^{-1}\phi' \cdot R^{-1}\phi') - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}. \quad (7.4.10)$$

Since this is gauge invariant, we can perform a gauge transformation:

$$\begin{cases} \phi \to R\phi = R \circ R^{-1}\phi' = \phi' \\ A_\mu \to A'_\mu, \end{cases} \quad (7.4.11)$$

which gives

$$\mathcal{L} = \frac{1}{2} D'_\mu \phi' \cdot D'^\mu \phi' - V(\phi' \cdot \phi') - \frac{1}{4} F'_{\mu\nu} F'^{\mu\nu}. \quad (7.4.12)$$

Now, we note that

$$\phi' \cdot \phi' = (\mathcal{V} + \mathcal{E})^2, \quad (7.4.13)$$

so, keeping only the quadratic terms

$$V(\phi' \cdot \phi') = \frac{1}{2} m^2 (\mathcal{V} + \mathcal{E})^2 + \frac{1}{4} \lambda (\mathcal{V} + \mathcal{E})^4 \quad (7.4.14)$$

$$= \mathcal{E}(m^2 \mathcal{V} + \lambda \mathcal{V}^3) + \mathcal{E}^2 \left(\frac{1}{2} m^2 + \frac{3}{2} \lambda \mathcal{V}^2 + \dots \right). \quad (7.4.15)$$

But because

$$\mathcal{V}^2 = \frac{-m^2}{\lambda}, \quad (7.4.16)$$

$V(\phi' \cdot \phi')$ simplifies to

$$V(\phi' \cdot \phi') \approx \mathcal{E}(-\lambda \mathcal{V}^3 + \lambda \mathcal{V}^3) + \mathcal{E}^2 \left(\frac{1}{2} m^2 - \frac{3}{2} m^2 + \dots \right) \quad (7.4.17)$$

$$\approx -\mathcal{E}^2 m^2 \quad (7.4.18)$$

$$= \frac{1}{2} (-2m^2) \mathcal{E}^2. \quad (7.4.19)$$

We also need to look at how D_μ changes:

$$D'_\mu = (\mathbb{I} \partial_\mu + igTA'_\mu) \quad (7.4.20)$$

$$= \begin{pmatrix} \partial_\mu & 0 \\ 0 & \partial_\mu \end{pmatrix} + ig \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} A'_\mu \quad (7.4.21)$$

$$= \begin{pmatrix} \partial_\mu & -igA'_\mu \\ igA'_\mu & \partial_\mu \end{pmatrix}. \quad (7.4.22)$$

So,

$$D'_\mu \phi' = \begin{pmatrix} \partial_\mu & -igA'_\mu \\ igA'_\mu & \partial_\mu \end{pmatrix} \begin{pmatrix} 0 \\ \mathcal{V} + \mathcal{E} \end{pmatrix} \tag{7.4.23}$$

$$= \begin{pmatrix} -gA'_\mu(\mathcal{V} + \mathcal{E}) \\ \partial_\mu \mathcal{E} \end{pmatrix}. \tag{7.4.24}$$

Therefore,

$$D'_\mu \phi' \cdot D'^\mu \phi' = \begin{pmatrix} -gA'_\mu(\mathcal{V} + \mathcal{E}) \\ \partial_\mu \mathcal{E} \end{pmatrix}^\top \begin{pmatrix} -gA'^\mu(\mathcal{V} + \mathcal{E}) \\ \partial^\mu \mathcal{E} \end{pmatrix} \tag{7.4.25}$$

$$= g^2 A'_\mu A'^\mu (\mathcal{V} + \mathcal{E})^2 + \partial_\mu \mathcal{E} \partial^\mu \mathcal{E}. \tag{7.4.26}$$

Thus the Lagrangian becomes:

$$\mathcal{L} = \frac{1}{2}\partial_\mu \mathcal{E} \partial^\mu \mathcal{E} - \frac{1}{2}(-2m^2)\mathcal{E}^2 - \frac{1}{4}F'_{\mu\nu}F'^{\mu\nu} \tag{7.4.27}$$

$$+ \frac{g^2 \mathcal{V}^2}{2}A'_\mu A'^\mu + \frac{g^2}{2}(2\mathcal{E}\mathcal{V} + \mathcal{E}^2)A'_\mu A'^\mu + \ldots \tag{7.4.28}$$

What does this theory describe? We shall break this down term-by-term:

1. The two terms:

$$\frac{1}{2}\partial_\mu \mathcal{E} \partial^\mu \mathcal{E} - \frac{1}{2}(-2m^2)\mathcal{E}^2 \tag{7.4.29}$$

which has $-m^2 > 0$ describe a massive scalar, spin 0 particle. This particle is called **the Higgs boson**.

2. The next two terms:

$$-\frac{1}{4}F'_{\mu\nu}F'^{\mu\nu} + \frac{g^2 \mathcal{V}^2}{2}A'_\mu A'^\mu \tag{7.4.30}$$

describe a massive vector gauge field (recall what we have derived in the electromagnetism section)

3. And the last (significant) term:

$$\frac{g^2}{2}(2\mathcal{E}\mathcal{V} + \mathcal{E}^2)A'_\mu A'^\mu \tag{7.4.31}$$

is the interaction between $\mathcal{E}$ and A'_μ.

Now, notice that we do not have a massless Nambu-Goldstone mode (the ζ term is long gone). We say that this mode has got "eaten", resulting in A_μ becoming a massive A'_μ. We can also count the degrees of freedom again to see what have changed. Recall that before spontaneous symmetry breaking,

we have 2 modes for the massless vector field A_μ (the photon) and 2 massive modes for $\phi = \begin{pmatrix} \phi_1 & \phi_2 \end{pmatrix}^\top$. So we had 4 in total. After spontaneous symmetry breaking, we have 1 degree of freedom for the massive scalar field $\mathcal{E}$, and 3 for the massive gauge field A'_μ (we will show why there are 3 degrees of freedom in the Worked Problems section). So, we also end up with 4 degrees of freedom, except without a massless mode.

Note that since O(2) is U(1) are very similar, we will not work out the details of spontaneous symmetry breaking of U(1) gauge theory here. This will serve as an exercise that will be covered in its entirety (theory derived from scratch) in the Worked Problems section.

Chapter 8

Gravitation and Lagrangian Formulation of General Relativity

8.1 Review of General Relativity & Curved Spacetime

8.1.1 General Relativity

Taking the speed of light, c, to be 1, the Einstein equation for general relativity is

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi G T_{\mu\nu}. \tag{8.1.1}$$

Sometimes, it is more useful to write the Einstein's equation as

$$R_{\mu\nu} = 8\pi G \left(T_{\mu\nu} - \frac{1}{2} T g_{\mu\nu} \right). \tag{8.1.2}$$

In vacuum, where all components of the energy-momentum stress tensor are zero, the Einstein's equation becomes

$$R_{\mu\nu} = 0. \tag{8.1.3}$$

8.1.2 Curved Spacetime

Perhaps the most important change as we go from flat to general curved spacetime is the notion of the "derivative." In introduction to general relativity, we

have introduced the **covariant derivative** and **absolute derivative**. We shall revisit the covariant derivative, as it will be important in the derivation of Lagrangian formulation of gravitation. The covariant derivative has the form - or forms, should I say:

$$D_\mu w_\nu = \partial_\mu w_\nu - \Gamma^\lambda_{\mu\nu} w_\lambda \tag{8.1.4}$$

$$D_\mu v^\nu = \partial_\mu v^\nu + \Gamma^\nu_{\lambda\nu} v^\lambda. \tag{8.1.5}$$

Note the sign difference in the definitions. The covariant derivative is defined *differently* for covariant and contravariant vectors. In fact, this follows if we define the covariant derivative for one kind of vectors.

Properties 8.1.1. Strictly speaking, some of theses properties are required for the definition of the covariant derivative to make sense and work.

1. Linearity: $D(U + V) = D(U) + D(V)$

2. Product rule: $D(U \otimes V) = D(U) \otimes D(V)$

3. Commutes with contractions: $D_\mu(T^\lambda{}_{\lambda\rho}) = (DT)_\mu{}^\lambda{}_{\lambda\rho}$

Next, recall the definition of the **Christoffel symbol**:

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\rho} \left(\partial_\mu g_{\rho\nu} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu} \right). \tag{8.1.6}$$

Now, consider a contravariant vector V^μ, we can define the **divergence** in curved spacetime based on the flat 3-space definition

$$\mathrm{div}\vec{V} = \partial_i V^i \tag{8.1.7}$$

by "contracting" an index and adding the Christoffel symbols:

$$D_\mu V^\mu = \partial_\mu V^\mu + \Gamma^\mu_{\mu\lambda} V^\lambda. \tag{8.1.8}$$

Next, invoking the definition of the Christoffel symbols, we can compute

$$\Gamma^\mu_{\mu\lambda} = \frac{1}{2} g^{\mu\rho}(\partial_\mu g_{\rho\lambda} + \partial_\lambda g_{\rho\mu} - \partial_\rho g_{\mu\lambda}) \tag{8.1.9}$$

$$= a \tag{8.1.10}$$

8.2   Lagrangian Formulation

8.3   Overview

The action is

$$S = \int \mathcal{L}(\Phi^i, D_\mu \Phi^i) \, d^n x. \tag{8.3.1}$$

$\mathcal{L}$ is a density.

$$\mathcal{L} = \sqrt{-g}\hat{\mathcal{L}} \tag{8.3.2}$$

where $\hat{\mathcal{L}}$ is a scalar and

$$g = \det(g_{ij}) \tag{8.3.3}$$

is the determinant of the metric tensor. The associated Euler-Lagrange equation is

$$\frac{\partial \hat{\mathcal{L}}}{\partial \Phi} - D_\mu \left(\frac{\partial \hat{\mathcal{L}}}{\partial(D_\mu \Phi)} \right) = 0. \tag{8.3.4}$$

Recall Stokes' theorem:

$$\int_\Sigma \nabla_\mu V^\mu \sqrt{|g|}\, d^n x = \int_{\partial\Sigma} n_\mu V^\mu \sqrt{|\gamma|}, d^{n-1}x. \tag{8.3.5}$$

Setting the variation equal to zero at the boundary and integrate by parts give

$$\int A^\mu (D_\mu B) \sqrt{-g}\, d^n x = - \int (D_\mu A^\mu) B \sqrt{-g}\, d^n x + \text{ boundary terms.} \tag{8.3.6}$$

$$S_\phi = \int \left[-\frac{1}{2} g^{\mu\nu}(D_\mu \phi)(D_\nu \phi) - V(\phi) \right] \sqrt{-g}\, d^n x, \tag{8.3.7}$$

The equation of motion:

$$\Box \phi - \frac{dV}{d\phi} = 0. \tag{8.3.8}$$

The **covariant d'Alembertian** now becomes

$$\Box \equiv D^\mu D_\mu \equiv g^{\mu\nu} D_\mu D_\nu. \tag{8.3.9}$$

The Hilbert action

$$S_H = \int \sqrt{-g}\, d^n x. \tag{8.3.10}$$

8.3.1 Introduction to the Lagrangian Formulation of General Relativity

Recall that the action in flat spacetime has the form:

$$S = \int \mathcal{L}\, d^4 x. \tag{8.3.11}$$

This is due to the fact that the metric is just $\eta_{\mu\nu} = \text{diag}(1, -1, -1, -1)$. But in curved spacetime, however, the metric is $g_{\mu\nu} \neq \eta_{\mu\nu}$. In some sense we can find how the form of the action changes as we move from flat to general curved spacetime. Consider the transformation $x^\mu \to x^{\mu'}$, then $dx^\mu \to X^{\mu'}_\nu dx^\nu$. The Jacobian matrix for this transformation is

$$\left[X^{\mu'}_\nu \right] = \left[\frac{\partial x^{\mu'}}{\partial x^\nu} \right] \tag{8.3.12}$$

and

$$g_{\mu'\nu'} = X^\alpha_{\mu'} X^\beta_{\nu'} g_{\alpha\beta}. \tag{8.3.13}$$

So the "volume" element becomes:

$$d^4 x \to d^4 x' = \left| \frac{\partial x^{\mu'}}{\partial x^\nu} \right| d^4 x. \tag{8.3.14}$$

Since we want our volume element to be invariant, we have to compensate with a factor of

$$g = \det(g_{\mu\nu}) = |g_{\mu\nu}|. \tag{8.3.15}$$

But we are not quite there yet. Notice that the metric tensor has the form

$$[g_{\mu\nu}] = \begin{pmatrix} + & & & \\ & - & & \\ & & - & \\ & & & - \end{pmatrix}, \tag{8.3.16}$$

which means $\det(g_{\mu\nu}) < 0$. So, we use $-g > 0$. From Eq. (8.3.15), we get that

$$g' = \left| \frac{\partial x}{\partial x'} \right|^2 g. \tag{8.3.17}$$

So,

$$g' = \left| \frac{\partial x'}{\partial x} \right|^{-2} g, \tag{8.3.18}$$

which means

$$\sqrt{-g'} = \left| \frac{\partial x'}{\partial x} \right|^{-1} \sqrt{-g}, \tag{8.3.19}$$

since $-g, -g' > 0$. Therefore,

$$d^4 x \sqrt{-g} = d^4 x' \left| \frac{\partial x'}{\partial x} \right|^{-1} \left| \frac{\partial x'}{\partial x} \right| \sqrt{-g'} = d^4 x' \sqrt{-g'}. \tag{8.3.20}$$

We have

$$d^4x\sqrt{-g} = d^4x'\sqrt{-g'} \tag{8.3.21}$$

as desired. Now, just for sanity check, we can return to flat spacetime where $g'_{\mu\nu} = g_{\mu\nu} = \eta_{\mu\nu} = \text{diag}(1, -1, -1, -1)$. Then, $-g = 1$ hence $\sqrt{-g} = 1$, which says $d^4x\sqrt{-g} = d^4x = d^4x' = d^4x'\sqrt{-g'}$. So, for curved spaces the action is given by

$$\boxed{S = \int d^4x\sqrt{-g}\mathcal{L}} \tag{8.3.22}$$

For pure gravity (no matter), the action is called the **Einstein-Hilbert action**

$$\boxed{S = \int d^4x\sqrt{-g}\frac{1}{16\pi G}R} \tag{8.3.23}$$

where R is the **Ricci scalar**, defined by a contraction

$$R = R^\mu{}_\mu = R^{\mu\nu}{}_{\mu\nu} = g^{\mu\nu}R_{\mu\nu} \tag{8.3.24}$$

where $R_{\mu\nu}$ is defined in terms of other contractions

$$R_{\mu\nu} = R^\lambda{}_{\mu\lambda\nu} = g^{\lambda\rho}R_{\rho\mu\lambda\nu}. \tag{8.3.25}$$

We also recall that

$$R^\rho{}_{\sigma\mu\nu} = \partial_\mu\Gamma^\rho_{\nu\sigma} - \partial_\nu\Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma} \tag{8.3.26}$$

where the **Christoffel symbols** are defined in terms of the metric tensor:

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2}g^{\lambda\sigma}\left(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu}\right). \tag{8.3.27}$$

8.3.2 Variation on the metric tensor, without matter: $T_{\mu\nu} = 0$

To find the physical laws or the equations of motion, we need to vary the action with respect to $g_{\mu\nu}$ and $g^{\mu\nu}$ since it is written in terms of the metric (the field of interest is the metric). To start, we must pick $g_{\mu\nu}$ or $g^{\mu\nu}$ as the fundamental field. The choice shouldn't matter, but for certain conveniences, we shall pick $g^{\mu\nu}$ as the fundamental field.

Now, $g_{\mu\nu}$ (the metric tensor) and $g^{\mu\nu}$ (its inverse) obey

$$g^{\mu\nu}g_{\nu\sigma} = \delta^\nu_\sigma, \tag{8.3.28}$$

which is covariantly constant. So,

$$(\delta g^{\mu\nu})g_{\nu\sigma} + g^{\mu\nu}(\delta g_{\nu\sigma}) = 0, \tag{8.3.29}$$

i.e.,

$$(\delta g^{\mu\nu})g_{\nu\sigma} = -g^{\mu\nu}(\delta g_{\nu\sigma}) \tag{8.3.30}$$

Multiplying with $g^{\sigma\rho}$ we get

$$(\delta g^{\mu\nu})g_{\nu\sigma}g^{\sigma\rho} = -g^{\mu\nu}g^{\sigma\rho}(\delta g_{\nu\sigma}) \tag{8.3.31}$$

$$(\delta g^{\mu\nu})\delta^{\nu}_{\rho} = -g^{\mu\nu}g^{\sigma\rho}(\delta g_{\nu\sigma}) \tag{8.3.32}$$

i.e.,

$$\delta g^{\mu\rho} = -g^{\mu\nu}g^{\sigma\rho}\delta g_{\nu\sigma}. \tag{8.3.33}$$

We an rewrite this as

$$\delta g^{\mu\nu} = -g^{\mu\alpha}g^{\nu\beta}\delta g_{\alpha\beta}. \tag{8.3.34}$$

Likewise,

$$\delta g_{\mu\nu} = -g_{\mu\alpha}g_{\nu\beta}\delta g^{\alpha\beta}. \tag{8.3.35}$$

Now, consider

$$\sqrt{-g}\mathcal{L} \sim \sqrt{-g}R = \sqrt{-g}g^{\mu\nu}R_{\mu\nu} \tag{8.3.36}$$

We have,

$$\begin{aligned}\delta(\sqrt{-g}\mathcal{L}) &\sim \delta(\sqrt{-g}R)\\ &= \delta\left[\sqrt{-g}g^{\mu\nu}R_{\mu\nu}\right]\\ &= (\delta\sqrt{-g})g^{\mu\nu}R_{\mu\nu} + \sqrt{-g}(\delta g^{\mu\nu})R_{\mu\nu} + (\sqrt{-g})g^{\mu\nu}(\delta R_{\mu\nu}). \end{aligned} \tag{8.3.37}$$

First, we want to find $\delta\sqrt{-g}$ in terms of $\delta g^{\mu\nu}$. We already know that $g = \det(g_{\mu\nu})$, so we use the following identity for matrices

$$\ln(\det(M)) = \text{Tr}(\ln(M)), \tag{8.3.38}$$

which can be a fun exercise to verify. Varying this identity, we get

$$\delta\ln(\det(M)) = \delta\,\text{Tr}(\ln(M)), \tag{8.3.39}$$

i.e.,

$$\frac{1}{\det(M)}\delta\det(M) = \text{Tr}(\delta\ln(M)) = \text{Tr}(M^{-1}\delta M). \tag{8.3.40}$$

If we let $M = [g_{\mu\nu}]$ and $M^{-1} = [g^{\mu\nu}]$ then we have

$$\ln(g) = \text{Tr}(\ln[g_{\mu\nu}]). \tag{8.3.41}$$

Varying with respect to the metric, we get

$$\frac{1}{g}(\delta g) = \text{Tr}(g^{\mu\nu}\delta g_{\mu\nu}). \tag{8.3.42}$$

But since $g^{\mu\nu}\delta g_{\mu\nu}$ is just a number, $\text{Tr}(g^{\mu\nu}\delta g_{\mu\nu}) = g^{\mu\nu}\delta g_{\mu\nu}$. Hence,

$$\boxed{\delta g = gg^{\mu\nu}\delta g_{\mu\nu} = -gg_{\mu\nu}\delta g^{\mu\nu}}. \tag{8.3.43}$$

Then

$$\begin{aligned}
\delta\sqrt{-g} &= \delta(-g)^{1/2} \\
&= \frac{1}{2}(-g)^{-1/2}\delta(-g) \\
&= \frac{-1}{2}\frac{1}{\sqrt{-g}}\delta g \\
&= \frac{-1}{2}\frac{1}{\sqrt{-g}}(-g)g_{\mu\nu}\delta g^{\mu\nu} \\
&= \frac{-1}{2}\sqrt{-g}g_{\mu\nu}\delta g^{\mu\nu}
\end{aligned} \tag{8.3.44}$$

i.e.,

$$\boxed{\delta\sqrt{-g} = \frac{-1}{2}\sqrt{-g}g_{\mu\nu}\delta g^{\mu\nu} = \frac{1}{2}\sqrt{-g}g^{\mu\nu}\delta g_{\mu\nu}} \tag{8.3.45}$$

Now, what about $\delta R = \delta R_{\mu\nu}$? Recall that

$$\begin{aligned}
\delta(\sqrt{-g}R) &= (\delta\sqrt{-g})(g^{\mu\nu}R_{\mu\nu}) + \sqrt{-g}(\delta g^{\mu\nu})R_{\mu\nu} + (\sqrt{-g})g^{\mu\nu}(\delta R_{\mu\nu}) \\
&= (\delta\sqrt{-g})R + \sqrt{-g}(\delta g^{\mu\nu})R_{\mu\nu} + (\sqrt{-g})g^{\mu\nu}(\delta R_{\mu\nu}) \\
&= \frac{-1}{2}\sqrt{-g}g_{\mu\nu}(\delta g^{\mu\nu})R + \sqrt{-g}(\delta g^{\mu\nu})R_{\mu\nu} + (\sqrt{-g})g^{\mu\nu}(\delta R_{\mu\nu})
\end{aligned}$$
$$\tag{8.3.46}$$

Since we have found $\delta\sqrt{-g}$ in terms of $\delta g^{\mu\nu}$,

$$\delta(\sqrt{-g}R) = \sqrt{-g}\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right)\delta g^{\mu\nu} + \sqrt{-g}g^{\mu\nu}\delta R_{\mu\nu}. \tag{8.3.47}$$

We immediately recognize the Einstein tensor $G_{\mu\nu}$

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R. \tag{8.3.48}$$

From the Einstein equations, we know that

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}. \tag{8.3.49}$$

But since there is no matter, $T_{\mu\nu} = 0$, i.e., $G_{\mu\nu} = 0$. So, if we do the variations correctly, we should get

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 0. \tag{8.3.50}$$

It turns out that after some simplification

$$\boxed{\int d^4x\sqrt{-g}\, g^{\mu\nu}\delta R_{\mu\nu} = 0} \tag{8.3.51}$$

which means we don't have to worry about the third term in the expansion of $\delta\sqrt{-g}R$ that involves $\delta R_{\mu\nu}$. Therefore, with

$$S = \int d^4x\sqrt{-g}\,\frac{1}{16\pi G}R, \tag{8.3.52}$$

we have

$$16\pi G\,\delta S = \int d^4x\sqrt{-g}\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right)\delta g^{\mu\nu} = 0, \tag{8.3.53}$$

since we require varying the action gives zero, which implies

$$\boxed{R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 0} \tag{8.3.54}$$

as expected.

Exercise 8.3.1. Verify Eq. (8.3.51).

Solution 8.3.1. We want to show that

$$\int d^4x\sqrt{-g}\, g^{\mu\nu}\delta R_{\mu\nu} = 0. \tag{8.3.55}$$

Recall that the Ricci tensor is a contraction of the Riemann tensor, which is given by

$$R^{\rho}{}_{\sigma\mu\nu} = \partial_\mu\Gamma^\rho_{\nu\sigma} - \partial_\nu\Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma}. \tag{8.3.56}$$

Consider arbitrary variations of the connection:

$$\Gamma^\rho_{\nu\mu} \to \Gamma^\rho_{\nu\mu} + \delta\Gamma^\rho_{\nu\mu}. \tag{8.3.57}$$

Since $\delta\Gamma^\rho_{\nu\mu}$ is a difference between the two connections, it is a tensor. Hence we take its covariant derivative:

$$D_\lambda(\delta\Gamma^\rho_{\nu\mu}) = \partial_\lambda(\delta\Gamma^\rho_{\nu\mu}) + \Gamma^\rho_{\lambda\sigma}\delta\Gamma^\sigma_{\nu\mu} - \Gamma^\sigma_{\lambda\nu}\delta\Gamma^\rho_{\sigma\mu} - \Gamma^\sigma_{\lambda\mu}\delta\Gamma^\rho_{\nu\sigma}. \tag{8.3.58}$$

We can then show to first order (in 8.3.3) that

$$\delta R^\rho{}_{\mu\lambda\nu} \approx D_\lambda(\delta\Gamma^\rho_{\nu\mu}) - D_\nu(\delta\Gamma^\rho_{\lambda\mu}), \tag{8.3.59}$$

so, variation on the contraction is

$$\delta R_{\mu\nu} = \delta R^\lambda{}_{\mu\lambda\nu} \approx D_\lambda(\delta\Gamma^\lambda_{\nu\mu}) - D_\nu(\delta\Gamma^\lambda_{\lambda\mu}). \tag{8.3.60}$$

So we can write

$$\int d^4x\sqrt{-g}g^{\mu\nu}\delta R_{\mu\nu} = \int d^4x\sqrt{-g}g^{\mu\nu}\left[D_\lambda(\delta\Gamma^\lambda_{\nu\mu}) - D_\nu(\delta\Gamma^\lambda_{\lambda\mu})\right] \tag{8.3.61}$$

$$= \int d^4x\sqrt{-g}D_\sigma\left[g^{\mu\nu}(\delta\Gamma^\sigma_{\nu\mu}) - g^{\mu\sigma}(\delta\Gamma^\lambda_{\lambda\mu})\right]. \tag{8.3.62}$$

We will verify the second equality in 8.3.2. Now,

$$\delta\Gamma^\sigma_{\mu\nu} = -\frac{1}{2}\left[g_{\lambda\mu}D_\nu(\delta g^{\lambda\sigma}) + g_{\lambda\nu}D_\mu(\delta g^{\lambda\sigma}) - g_{\mu\alpha}g_{\nu\beta}D^\sigma(\delta g^{\alpha\beta})\right]. \tag{8.3.63}$$

Therefore,

$$\int d^4x\sqrt{-g}g^{\mu\nu}\delta R_{\mu\nu} = \int d^4x\sqrt{-g}D_\sigma\left[g_{\mu\nu}D^\sigma(\delta g^{\mu\nu}) - D_\lambda(\delta g^{\sigma\lambda})\right]. \tag{8.3.64}$$

But notice that this result is an integral with respect to the natural volume element of the covariant divergence of a vector. So, by Stokes' theorem, this is equal to a boundary contribution at infinity, which we can (in a hand-wavy manner) set to zero by making the variation vanish at infinity, i.e.,

$$\int d^4x\sqrt{-g}g^{\mu\nu}\delta R_{\mu\nu} = 0, \tag{8.3.65}$$

as we wanted. So, this term does not contribute the action variation. $\qquad\square$

Exercise 8.3.2. Verify the following approximation in Eq. (8.3.59)

$$\delta R^\rho{}_{\mu\lambda\nu} \approx D_\lambda(\delta\Gamma^\rho_{\nu\mu}) - D_\nu(\delta\Gamma^\rho_{\lambda\mu}) \tag{8.3.66}$$

to first order.

Solution 8.3.2. Do this $\qquad\square$

Exercise 8.3.3. Verify the equality used in Eq. (8.3.61)

$$g^{\mu\nu}\left[D_\lambda(\delta\Gamma^\lambda_{\nu\mu}) - D_\nu(\delta\Gamma^\lambda_{\lambda\mu})\right] = D_\sigma\left[g^{\mu\nu}(\delta\Gamma^\sigma_{\nu\mu}) - g^{\mu\sigma}(\delta\Gamma^\lambda_{\lambda\mu})\right] \tag{8.3.67}$$

Solution 8.3.3. Do this ◻

With non-zero cosmological constant Λ and no matter ($T_{\mu\nu} = 0$), the action simply becomes:

$$S = \frac{1}{16\pi G} \int d^4x \sqrt{-g}(R - 2\Lambda). \qquad (8.3.68)$$

Then,

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu} = 0. \qquad (8.3.69)$$

Exercise 8.3.4. Verify Eq. (8.3.69).

Solution 8.3.4. We simply repeat what we have done before:

$$
\begin{aligned}
16\pi G(\delta S) &= \int d^4x\, \delta\left(\sqrt{-g}(R - 2\Lambda)\right) \\
&= \int d^4x\, \delta(\sqrt{-g}R) - 2\Lambda\delta\sqrt{-g} \\
&= \int d^4x\, \delta\left(\sqrt{-g}g^{\mu\nu}R_{\mu\nu}\right) - 2\Lambda\delta\sqrt{-g} \\
&= \int d^4x\, \delta\left(\sqrt{-g}g^{\mu\nu}R_{\mu\nu}\right) - 2\Lambda\delta\sqrt{-g} \\
&= \int d^4x\sqrt{-g}\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right)\delta g^{\mu\nu} - 2\Lambda\left(-\frac{1}{2}\sqrt{-g}g_{\mu\nu}\right)\delta g^{\mu\nu} \\
&= \int d^4x\sqrt{-g}\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu}\right)\delta g^{\mu\nu} \\
&= 0. \qquad (8.3.70)
\end{aligned}
$$

Since we require $\delta S = 0$,

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu} = 0 \qquad (8.3.71)$$

must hold. ◻

8.3.3 Variations on the metric tensor, with matter: $T_{\mu\nu} \neq 0$

A natural question is how we can add matter into the action. It turns out that we simply add a matter Lagrangian to the Lagrangian density:

$$S = \int d^4x\sqrt{-g}\left(\frac{1}{16\pi G}(R - 2\Lambda) + \mathcal{L}_{matter}\right), \qquad (8.3.72)$$

where the matter Lagrangian term $\mathcal{L}_{matter}$ also involves the metric:

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu \phi)g^{\mu\nu}(\partial_\nu \phi) - \frac{1}{2}m^2\phi^2, \tag{8.3.73}$$

where ϕ is a scalar field. One thing to notice is that there is an interaction between the field and the metric. This makes things a little more complicated.

If we want to vary this action, we will have to include the term

$$\delta\left(\sqrt{-g}\left[-\frac{1}{2}(\partial_\mu \phi)g^{\mu\nu}(\partial_\nu \phi) - \frac{1}{2}m^2\phi^2\right]\right). \tag{8.3.74}$$

These extra terms will contribute to $T_{\mu\nu}$. So we simply define

$$\delta S = \delta S_g + \delta S_M = \delta\left(\int d^4x \sqrt{-g}\frac{1}{16\pi G}(R - 2\Lambda) + \int d^4x \sqrt{-g}\mathcal{L}_M\right) \tag{8.3.75}$$

and

$$\boxed{\delta\left(\sqrt{-g}\mathcal{L}_M\right) = -\frac{1}{2}\sqrt{-g}T_{\mu\nu}\delta g^{\mu\nu}} \tag{8.3.76}$$

or

$$\boxed{T_{\mu\nu} = \frac{-2}{\sqrt{-g}}\frac{\delta}{\delta g^{\mu\nu}}\left[\sqrt{-g}\mathcal{L}_M\right]} \tag{8.3.77}$$

With that, for $\Lambda = 0$,

$$\begin{aligned}
\delta S &= \int d^4x \frac{\sqrt{-g}}{16\pi G}\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right)\delta g^{\mu\nu} + \delta(\sqrt{-g}\mathcal{L}_M) \\
&= \int d^4x \left[\frac{1}{16\pi G}\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right) - \frac{1}{2}T_{\mu\nu}\right]\sqrt{-g}\delta g^{\mu\nu} \\
&= 0.
\end{aligned} \tag{8.3.78}$$

Therefore,

$$\frac{1}{16\pi G}\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right) = \frac{1}{2}T_{\mu\nu}. \tag{8.3.79}$$

So, we get the Einstein equations

$$\boxed{R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G T_{\mu\nu}} \tag{8.3.80}$$

Now, if

$$\mathcal{L}_M = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{8.3.81}$$

then

$$\sqrt{-g}\mathcal{L}_M = -\frac{1}{4}\sqrt{-g}F_{\mu\nu}g^{\mu\alpha}g^{\nu\beta}F_{\alpha\beta} \qquad (8.3.82)$$

and we can show that

$$F_{\mu\nu} = D_\mu A_\nu - D_\nu A_\mu = \partial_\mu A_\nu - \partial_\nu A_\mu. \qquad (8.3.83)$$

We can vary this with respect to $g^{\mu\nu}$ to get $T^{\mu\nu}$ (which we can readily verify)

$$\boxed{T_{\mu\nu} = F_{\mu\lambda}F_\nu{}^\lambda - \frac{1}{4}g_{\mu\nu}F_{\alpha\beta}F^{\alpha\beta}} \qquad (8.3.84)$$

We can actually check and see that this is actual energy-momentum stress tensor. For instance,

$$T_{00} \sim (\vec{E}^2 + \vec{B}^2) \sim \text{ Energy density} \qquad (8.3.85)$$

$$T_{0j} \sim \text{Poynting vectors.} \qquad (8.3.86)$$

Exercise 8.3.5. Verify Eq. (8.3.84).

Solution 8.3.5. We start from the definition of the energy-momentum stress tensor:

$$T_{\mu\nu} = \frac{-2}{\sqrt{-g}}\frac{\delta}{\delta g^{\mu\nu}}\left[\sqrt{-g}\mathcal{L}_M\right]. \qquad (8.3.87)$$

To make the simplification a little easier, we shall assume δ is with respect to $g^{\mu\nu}$ implicitly, so we just write:

$$\frac{-1}{2}\sqrt{-g}T_{\mu\nu}\delta g^{\mu\nu} = \delta\left[\sqrt{-g}\mathcal{L}_M\right] \qquad (8.3.88)$$

instead. By (8.3.82)

$$\delta\left[\sqrt{-g}\mathcal{L}_M\right] = \delta\left(-\frac{1}{4}\sqrt{-g}F_{\mu\nu}g^{\mu\alpha}g^{\nu\beta}F_{\alpha\beta}\right)$$

$$= -\frac{1}{4}(\delta\sqrt{-g})F_{\alpha\beta}F^{\alpha\beta} - \frac{1}{4}\sqrt{-g}\delta\left(F_{\mu\nu}g^{\mu\alpha}g^{\nu\beta}F_{\alpha\beta}\right). \qquad (8.3.89)$$

We know that

$$\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g}g_{\mu\nu}\delta g^{\mu\nu}. \qquad (8.3.90)$$

And,

$$\delta\left(F_{\mu\nu}g^{\mu\alpha}g^{\nu\beta}F_{\alpha\beta}\right) = g^{\nu\beta}F_{\mu\nu}F_{\alpha\beta}\delta g^{\mu\alpha} + g^{\mu\alpha}F_{\mu\nu}F_{\alpha\beta}\delta g^{\nu\beta}. \qquad (8.3.91)$$

With respect to $g^{\mu\nu}$:

$$\delta\left(F_{\mu\nu}g^{\mu\alpha}g^{\nu\beta}F_{\alpha\beta}\right) = g^{\nu\beta}F_{\mu\nu}F_{\alpha\beta}\delta g^{\mu\alpha} + g^{\mu\alpha}F_{\mu\nu}F_{\alpha\beta}\delta g^{\nu\beta}$$

$$= F_{\mu\nu}F_{\alpha\beta}\left[\delta_\alpha^\nu g^{\nu\beta} + \delta_\beta^\mu g^{\mu\alpha}\right]\delta g^{\mu\nu}$$

$$= \left[F_{\mu\nu}F_{\alpha\beta}\delta_\alpha^\nu g^{\nu\beta} - F_{\nu\mu}F_{\alpha\beta}\delta_\beta^\mu g^{\mu\alpha}\right]\delta g^{\mu\nu}$$

$$= \left[F_\mu{}^\beta F_{\nu\beta} - F_\nu{}^\alpha F_{\alpha\mu}\right]\delta g^{\mu\nu}$$

$$= \left[F_\mu{}^\beta F_{\nu\beta} + F_\nu{}^\alpha F_{\mu\alpha}\right]\delta g^{\mu\nu}$$

$$= 2\left(F_{\mu\lambda}F_\nu{}^\lambda\right)\delta g^{\mu\nu}, \tag{8.3.92}$$

where we have used the anti-symmetric property of $F_{\mu\nu}$ (and $F^{\mu\nu}$) and λ is a dummy variable replacing α, β in the contraction. So, putting everything together,

$$T_{\mu\nu} = \frac{-2}{\sqrt{-g}}\frac{\delta}{\delta g^{\mu\nu}}\left[\sqrt{-g}\mathcal{L}_M\right]$$

$$= \frac{-2}{\sqrt{-g}}\left[\left(\frac{1}{8}\sqrt{-g}g_{\mu\nu}F_{\alpha\beta}F^{\alpha\beta}\right) - \frac{1}{4}\sqrt{-g}\,(???)\right]$$

$$= -\frac{1}{4}g_{\mu\nu}F_{\alpha\beta}F^{\alpha\beta} + \frac{1}{2}\left[2\left(F_{\mu\lambda}F_\nu{}^\lambda\right)\right]$$

$$= -\frac{1}{4}g_{\mu\nu}F_{\alpha\beta}F^{\alpha\beta} + F_{\mu\lambda}F_\nu{}^\lambda. \tag{8.3.93}$$

Thus we have shown

$$T_{\mu\nu} = F_{\mu\lambda}F_\nu{}^\lambda - \frac{1}{4}g_{\mu\nu}F_{\alpha\beta}F^{\alpha\beta}. \tag{8.3.94}$$

$\square$

8.4 Properties of Einstein Equations

Chapter 9

Diffeomorphism, Vierbein Formalism, and general spacetime symmetries

9.1 Overview of Diffeomorphisms and the Lie Derivative

Loosely speaking, a diffeomorphism is a mapping of one manifold to another. It is an isomorphism of smooth manifolds. It is an invertible function that maps one differentiable manifold to another such that both the function and its inverse are smooth.

In general relativity, a diffeomorphism is a mapping of spacetime to itself. Before getting rigorous, we can think of a simple diffeomorphism as a mapping of point P with X^μ to point Q with $X^\mu + \xi^\mu$. We would like to know how scalars,vectors, and tensors change under diffeomorphisms. It turns out that diffeomorphisms are tightly related to Lie algebra, and thus quite naturally changes in tensors are given by the Lie derivative.

General coordinate transformations are a type of diffeomorphism. Within general coordinate transformations, there are two kinds: **passive** and **active**. Passive transformations only change "perspective" roughly speaking, i.e., we only change the coordinates describing a tensor, while active transformations *actively* alternate tensors.

An example of a passive transformation is rotation in coordinates by some angle θ - tilting our frame of reference. An example of an active transformation is a rotation of a vector $\vec{A}$ by an angle $-\theta$.

insert figure here

Here we see that the new components of $\vec{A}$ in X' frame under a passive transformation are the same as the components of $A^{i'}$ of $\vec{A}'$ under an active transformation.

Symmetries always involve active transformations. With unbroken symmetries passive and active transformations are inverses of each other.

In general relativity, a diffeomorphism is an active transformation $A^{\mu} \to A'^{\mu}$ under moving or translating $X^{\mu} \to X^{\mu}\xi^{\mu}$. The passive version is a general transformation. Since this is invertible, the corresponding passive transformation is $X^{\mu} \to X^{\mu} - \xi^{\mu}$. We can use the inverse general coordinate transformation to find the form of the **Lie derivative**. Let $\mathcal{L}_{\xi}$ denote a Lie derivative using ξ^{μ}. Then under diffeomorphism,

$$A_{\mu} \to A_{\mu} + \mathcal{L}_{\xi} A_{\mu}$$
$$A^{\mu} \to A^{\mu} + \mathcal{L}_{\xi} A^{\mu}.$$

If $\delta S = 0$ under diffeomorphism, then the theory is diffeomorphism-invariant. General relativity *is* diffeomorphism-invariant.

Let us consider an infinitesimal general coordinate transformation:

$$X^{\mu'} = X^{\mu} - \xi^{\mu}. \tag{9.1.1}$$

Now, a vector under general coordinate transformation obeys

$$A^{\mu'}(x') = J^{\mu'}_{\nu} A^{\nu}(x)$$

where $J^{\mu'}_{\nu}$ denotes the Jacobian matrix whose elements are

$$\boxed{J^{\mu'}_{\nu} = \frac{\partial X^{\mu'}}{\partial X^{\nu}} = \frac{\partial}{\partial X^{\nu}} \left(X^{\mu} - \xi^{\mu} \right) = \delta^{\mu}_{\nu} - \partial_{\nu} \xi^{\mu}} \tag{9.1.2}$$

It follows that

$$A^{\mu'}(x') = \left(\delta^{\mu}_{\nu} - \partial_{\nu} \xi^{\mu} \right) A^{\nu}(x) \tag{9.1.3}$$

where we can Taylor expand $A^{\mu'}(x')$ at $x' = x - \xi$ to get

$$A^{\mu'}(x - \xi) \approx A^{\mu'}(x) - \xi^{\nu} \partial_{\nu} A^{\mu'}(x) + \ldots \tag{9.1.4}$$

Then use the fact that

$$\xi^\nu \partial_\nu A^{\mu'}(x) \approx \xi^\nu \partial_\nu A^\mu(x) + \dots \tag{9.1.5}$$

we get

$$A^{\mu'}(x') = A^{\mu'}(x) - \xi^\nu \partial_\nu A^\mu(x) = A^\mu(x) - (\partial_\nu \xi^\mu) A^\nu(x). \tag{9.1.6}$$

And therefore

$$\boxed{A^{\mu'}(x) = A^\mu(x) - (\partial_\nu \xi^\mu) A^\nu(x) + \xi^\nu (\partial_\nu A^\mu(x))} \tag{9.1.7}$$

We require that this gives the same result as for the active diffeomorphism: $A^\mu(x) \to A'^\mu(x) = A^\mu(x) + \mathcal{L}_\xi A^\mu(x)$. So the Lie derivative of a contravariant vector must be defined such that

$$\boxed{\mathcal{L}_\xi A^\mu(x) = -(\partial_\nu \xi^\mu) A^\nu(x) + \xi^\nu (\partial_\nu A^\mu(x))} \tag{9.1.8}$$

We note that this definition does not involve parallel transport, which means the Christoffel symbols also don't appear.

For a scalar $\phi(x)$, under the diffeomorphism $X^\mu \to X^{\mu'} = X^\mu - \xi^\mu$, we know that

$$\phi(x) = \phi'(x') = \phi'(x - \xi) \approx \phi(x) - \xi^\nu \partial_\nu \phi'(x)(x) \approx \phi'(x) - \xi^\nu \partial_\nu \phi(x). \tag{9.1.9}$$

Therefore,

$$\boxed{\phi'(x) = \phi(x) + \xi^\nu \partial_\nu \phi(x)} \tag{9.1.10}$$

Under a diffeomorphism $X^\mu \to X^\mu + \xi^\mu$, we want

$$\phi(x) \to \phi'(x') = \phi(x) + \mathcal{L}_\xi \phi(x). \tag{9.1.11}$$

Thus, the Lie derivative of a scalar must be defined such that

$$\boxed{\mathcal{L}_\xi \phi(x) = \xi^\nu \partial_\nu \phi(x)} \tag{9.1.12}$$

Next, we can find the form of the Lie derivative of a covariant vector, $A_\mu(x)$ by considering

$$A_{\mu'}(x') = J^\nu_{\mu'} A_\nu(x). \tag{9.1.13}$$

We can verify that

$$\boxed{X^\nu_{\mu'} = \delta^\nu_\mu + \partial_\mu \xi^\nu} \tag{9.1.14}$$

by multiplying the Jacobian matrix and its inverse and show this is the identity matrix:

$$J^{\nu}_{\mu'}J^{\mu'}_{\gamma} = (\delta^{\nu}_{\mu} + \partial_{\mu}\xi^{\nu})(\delta^{\mu}_{\gamma} + \partial_{\gamma}\xi^{\mu}) \tag{9.1.15}$$

$$= \delta^{\nu}_{\gamma} - \partial_{\gamma}\xi^{\nu} + \partial_{\gamma}\xi^{\nu} \tag{9.1.16}$$

$$= \delta^{\nu}_{\gamma} \tag{9.1.17}$$

$$\square$$

So then

$$A_{\mu'}(x') = J^{\nu}_{\mu'}A_{\nu}(x) \tag{9.1.18}$$

$$= (\delta^{\nu}_{\mu} + \partial_{\mu}\xi^{\nu})A_{\nu}(x) \tag{9.1.19}$$

$$= A_{\mu}(x) + (\partial_{\mu}\xi^{\nu})A_{\nu}(x). \tag{9.1.20}$$

Once again, we Taylor expand at $x' = x - \xi$ to get

$$A_{\mu'}(x') = A_{\mu'}(x - \xi) \tag{9.1.21}$$

$$\approx A_{\mu'}(x) - \xi^{\nu}\partial_{\nu}A_{\mu'}(x) \tag{9.1.22}$$

$$\approx A_{\mu'}(x) - \xi^{\nu}\partial_{\nu}A_{\mu}(x). \tag{9.1.23}$$

Then,

$$A_{\mu'}(x) = A_{\mu}(x) + (\partial_{\mu}\xi^{\nu})A_{\nu}(x) - \xi^{\nu}(\partial_{\nu}A_{\mu}(x)). \tag{9.1.24}$$

We claim that (and this is totally legitimate as we will show later when we look at the Lie derivative more rigorously) this result is the same as how $A_{\mu}(x) \to A_{\mu'}(x) = A_{\mu}(x) + \mathcal{L}_{\xi}A_{\mu}(x)$ under the diffeomorphism $X^{\mu} \to X^{\mu} + \xi^{\mu}$. Thus, we get the Lie derivative of a covariant derivative:

$$\boxed{\mathcal{L}_{\xi}A_{\mu}(x) = (\partial_{\mu}\xi^{\nu})A_{\nu}(x) + \xi^{\nu}\partial_{\nu}A_{\mu}(x)} \tag{9.1.25}$$

Given these, we can now guess the form of the Lie derivative for a general tensor, say $\tau^{\mu\nu}{}_{\sigma}$:

$$\boxed{\mathcal{L}_{\xi}\tau^{\mu\nu}{}_{\sigma} = -(\partial_{\alpha}\xi^{\mu})\tau^{\alpha\nu}{}_{\sigma} - (\partial_{\alpha}\xi^{\nu})\tau^{\mu\alpha}{}_{\sigma} + (\partial_{\sigma}\xi^{\alpha})\tau^{\mu\nu}{}_{\alpha} + \xi^{\alpha}\partial_{\alpha}\tau^{\mu\nu}{}_{\sigma}} \tag{9.1.26}$$

We will of course look at the most general case later when we look at the mathematics more carefully in Sean Carroll's book.

Next, let us look at the most important tensor the **metric tensor** $g_{\mu\nu}$. Under a diffeomorphism $g_{\mu\nu} \to g_{\mu\nu} + \mathcal{L}_{\xi}g_{\mu\nu}$, we can show that the Lie derivative of the metric tensor has the following form, from definition

$$\mathcal{L}_{\xi}g_{\mu\nu} = (D_{\mu}\xi^{\alpha})g_{\alpha\nu} + (D_{\nu}\xi^{\alpha})g_{\mu\alpha} + \xi^{\alpha}D_{\alpha}g_{\mu\nu}. \tag{9.1.27}$$

As an aside: the Lie derivative can be written completely in covariant derivatives (and not partial derivatives) and still remain equivalent, as we will show in an exercise. As a consequence, Lie derivatives of tensors and tensors. This fact comes in handy right now when we work with the metric tensor, which is covariantly constant, i.e.,

$$D_\alpha g_{\mu\nu} = 0, \qquad (9.1.28)$$

so not only can we drop the last term but we can also pull the metric tensor into the covariant derivative. And so the Lie derivative of $g_{\mu\nu}$ reduces to

$$\mathcal{L}_\xi g_{\mu\nu} = D_\mu \left(g_{\alpha\nu} \xi^\alpha \right) + D_\nu \left(g_{\mu\alpha} \xi^\alpha \right). \qquad (9.1.29)$$

But now the metric tensor acts on ξ^α and lower the index to a corresponding one, and thus the Lie derivative of the metric tensor has this very nice, symmetric form

$$\boxed{\mathcal{L}_\xi g_{\mu\nu} = D_\mu \xi_\nu + D_\nu \xi_\mu} \qquad (9.1.30)$$

And so under a diffeomorphism the metric tensor transforms as

$$g_{\mu\nu} \to g_{\mu\nu} + \mathcal{L}_\xi g_{\mu\nu} = g_{\mu\nu} + D_\mu \xi_\nu + D_\nu \xi_\mu. \qquad (9.1.31)$$

We notice that if we have not invoked the covariant derivatives, the we will not get this nicely simplified formula involving only partial derivatives.

Here is the following layout for what we will do next:

- Look at the mathematics of diffeomorphism

- Look at some basic mathematics of Lie algebra

- See how diffeomorphisms are a symmetry of general relativity, and what it means to break diffeomorphism.

9.2 Spacetime Symmetry

In this section we want to consider

1. Global Lorentz Transformation in Minkowski spacetime (no gravity)

2. Diffeomorphism in curved spacetime (with gravity)

3. Local Lorentz Transformations in curved spacetime (with gravity)

9.2.1 Global Lorentz Transformations in Minkowski Space-time

We know that Lorentz Transformations are coordinate transformations $X^\mu \to X^{\mu'}$, given by

$$X^{\mu'} = \Lambda^{\mu'}_\nu X^\nu, \tag{9.2.1}$$

where the $\Lambda^{\mu'}_\nu$ are constants. For vectors,

$$V_{\mu'} = \Lambda^\nu_{\mu'} V_\nu \qquad V^{\mu'} = \Lambda^{\mu'} = \Lambda^{\mu'}_\nu V^\nu. \tag{9.2.2}$$

The transformations also have inverses:

$$\Lambda^{\mu'}_\nu A^\nu_{\alpha'} = \delta^{\mu'}_{\alpha'} = \delta^\mu_\alpha = \Lambda^\mu_{\nu'} \Lambda^{\nu'}_\alpha. \tag{9.2.3}$$

Also,

$$\eta_{\mu'\nu'} = \Lambda^\alpha_{\mu'} \Lambda^\beta_{\nu'} \eta_{\alpha\beta} \tag{9.2.4}$$

However, all this is from the passive point of view. Now, we want to look at the active Lorentz Transformations where X^μ doesn't change (i.e., no primed indices). Notice that we have been writing $\Lambda^{\mu'}_\nu$ with one index on top of another. Now, distinguish $\Lambda_\mu{}^\nu$ versus $\Lambda^{\nu\mu}$. But we keep the "inverse" identities:

$$\Lambda_\mu{}^\alpha \Lambda^\mu{}_\beta = \delta^\alpha_\beta = \Lambda^\alpha{}_\mu \Lambda_\beta{}^\mu. \tag{9.2.5}$$

We must also keep

$$\eta_{\mu\nu} \to \Lambda_\mu{}^\alpha \Lambda_\nu{}^\beta \eta_{\alpha\beta} = \eta_{\mu\nu}, \tag{9.2.6}$$

i.e., the Minkowski metric is unchanged, as expected, since we haven't changed anything major yet.

Now, consider infinitesimal Lorentz Transform

$$\begin{cases} \Lambda_\mu{}^\nu = \delta^\nu_\mu + \epsilon_\mu{}^\nu \\ \Lambda^\mu{}_\nu = \delta^\mu_\nu + \epsilon^\mu{}_\nu \end{cases} \tag{9.2.7}$$

where $\epsilon_\mu{}^\nu$ is a small and constant, and we neglect squared terms. We have that

$$\Lambda_\mu{}^\alpha \Lambda^\mu{}_\beta = (\delta^\alpha_\mu + \epsilon_\mu{}^\alpha)(\delta^\mu_\nu + \epsilon^\mu{}_\nu) \tag{9.2.8}$$

$$= \delta^\alpha_\nu + \epsilon^\alpha{}_\nu + \epsilon_\nu{}^\alpha + \mathcal{O}, \qquad = \delta^\alpha_\nu, \tag{9.2.9}$$

provided that $\epsilon^\alpha{}_\beta = -\epsilon_\beta{}^\alpha$. We also have that

$$\Lambda^\mu{}_\alpha \Lambda_\nu{}^\alpha = (\delta^\mu_\alpha + \epsilon^\mu{}_\alpha)(\delta^\alpha_\nu + \epsilon_\nu{}^\alpha)$$
$$= \delta^\mu_\nu + \epsilon^\mu{}_\nu + \epsilon_\nu{}^\mu + \mathcal{O}$$
$$= \delta^\mu_\nu,$$

provided that $\epsilon^{\mu}{}_{\nu} = -\epsilon_{\nu}{}^{\mu}$. Furthermore,

$$\begin{aligned}
\eta_{\mu\nu} &= \Lambda_{\mu}{}^{\alpha}\Lambda_{\nu}{}^{\beta}\eta_{\alpha\beta} \\
&= \eta_{\mu\nu} + \epsilon_{\mu}{}^{\alpha}\eta_{\alpha\nu} + \epsilon_{\nu}{}^{\beta}\eta_{\mu\beta} \\
&= \eta_{\mu\nu} + \epsilon_{\mu\nu} + \epsilon_{\nu\mu} \\
&= \eta_{\mu\nu},
\end{aligned}$$

since

$$\boxed{\epsilon_{\mu\nu} = -\epsilon_{\nu\mu}}$$

So, we conclude that the Minkowski is unchanged if the above equality holds. The parameters $\epsilon_{\mu\nu}$ are anti-symmetric and 4-dimensional, and hence has 6 independent components.

With this, we can summarize how things transform under infinitesimal Lorentz transformations:

1. Scalars: $\phi \to \phi$, i.e., invariant.

2. Coordinates: X^{μ}'s don't change.

3. d^4x doesn't change.

4. Minkowski metric is invariant: $\eta_{\mu\nu} \to \eta_{\mu\nu}$.

However, all dynamical vectors and tensors change:

$$\begin{cases} A_{\mu} \to A_{\mu} + \epsilon_{\mu}{}^{\nu}A_{\nu} \\ A^{\mu} \to A^{\mu} + \epsilon^{\mu}{}_{\nu}A^{\nu} \end{cases} . \tag{9.2.10}$$

For a tensor:

$$\tau^{\mu\nu}{}_{\lambda} \to \tau^{\mu\nu}{}_{\lambda} + \epsilon^{\mu}{}_{\alpha}\tau^{\alpha\nu}{}_{\lambda} + \epsilon^{\nu}{}_{\alpha}\tau^{\mu\alpha}{}_{\lambda} + \epsilon_{\lambda}{}^{\alpha}\tau^{\mu\nu}{}_{\alpha}. \tag{9.2.11}$$

Now, when will

$$S = \int d^4x\,\mathcal{L} \tag{9.2.12}$$

be invariant, i.e., $(\delta S = 0)$ under global Lorentz transformations? If $\mathcal{L}$ is a scalar function then under a global Lorentz transformation, $\mathcal{L} \to \mathcal{L}$. So, then we must have $S \to S$, or $\delta S = 0$, which says we have a symmetry.

Exercise 9.2.1. Is

$$\mathcal{L} = \frac{1}{2}(\partial_{\mu}\phi)(\partial^{\mu}\phi) - \frac{1}{2}m^2\phi^2 \tag{9.2.13}$$

a scalar under Lorentz transformations?

Solution 9.2.1. Under Lorentz transformation, $\phi \to \phi$, $\phi^2 \to \phi^2$, and so $\frac{1}{2}m^2\phi^2 \to \frac{1}{2}m^2\phi^2$. Now,

$$\begin{cases} \partial_\mu\phi \to \partial_\mu\phi + \epsilon_\mu{}^\alpha \partial_\alpha\phi \\ \partial^\mu\phi \to \partial^\mu\phi + \epsilon^\mu{}_\alpha \partial^\alpha\phi \end{cases}. \tag{9.2.14}$$

So,

$$(\partial_\mu\phi)(\partial^\mu\phi) = (\epsilon_\mu{}^\alpha \partial_\alpha\phi + \partial_\mu\phi)(\epsilon^\mu{}_\alpha \partial^\alpha\phi + \partial^\mu\phi) \tag{9.2.15}$$

$$= \mathcal{O} + (\partial_\mu\phi)(p^\mu\phi) + (\partial_\mu\phi)(\epsilon^\mu{}_\alpha \partial^\alpha\phi) + (\partial^\mu\phi)(\epsilon_\mu{}^\alpha \partial_\alpha\phi) \tag{9.2.16}$$

$$= (\partial_\mu\phi)(\partial^\mu\phi) + \epsilon^{\mu\alpha}(\partial_\alpha\phi)(\partial_\mu\phi) + \epsilon^{\mu\alpha}(\partial_\alpha\phi)(\partial_\mu\phi). \tag{9.2.17}$$

Now,

$$\epsilon^{\mu\alpha}(\partial_\alpha\phi)(\partial_\mu\phi) = -\epsilon^{\alpha\mu}(\partial_\mu\phi)(\partial_\alpha\phi) \tag{9.2.18}$$

$$= -\epsilon^{\alpha\mu}(\partial_\alpha\phi)(\partial_\mu\phi). \tag{9.2.19}$$

So,

$$\epsilon^{\mu\nu}(\partial_\alpha\phi)(\partial_\mu\phi) = 0. \tag{9.2.20}$$

And thus,

$$(\partial_\mu\phi)(\partial^\mu\phi) \to (\partial_\mu\phi)(\partial^\mu\phi). \tag{9.2.21}$$

Exercise 9.2.2. Show that

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + m^2 A_\mu A^\mu \tag{9.2.22}$$

is a scalar under global Lorentz transformation.

Solution 9.2.2. Use the fact that $F_{\mu\nu}$ is a tensor, so that

$$F_{\mu\nu} \to F_{\mu\nu} + \epsilon_\mu{}^\alpha F_{\alpha\nu} + \epsilon_\nu{}^\alpha F_{\mu\alpha}. \tag{9.2.23}$$

Likewise, for $A_\mu A^\mu$. Next, show that $\mathcal{L} \to \mathcal{L}$.

$\square$

9.2.2 Diffeomorphism in Curved Spacetime

$$S = \int d^4x \sqrt{-g}\mathcal{L}. \tag{9.2.24}$$

Under diffeomorphism with ξ^μ, scalars and tensors transform with changes given by the Lie derivatives.

1. Scalars:

$$\phi \to \phi + \xi^\alpha \partial_\alpha \phi \qquad (9.2.25)$$

$$\phi \to \phi + \xi^\alpha D_\alpha \phi. \qquad (9.2.26)$$

2. Covariant vectors:

$$A_\mu \to A_\mu + (\partial_\mu \xi^\alpha) A_\alpha + \xi^\alpha \partial_\alpha A_\mu \qquad (9.2.27)$$

$$A_\mu \to A_\mu + (D_\mu \xi^\alpha) A_\alpha + \xi^\alpha D_\alpha A_\mu. \qquad (9.2.28)$$

3. Contravariant vectors:

$$A^\mu \to A^\mu - (\partial_\alpha \xi^\mu) A^\alpha + \xi^\alpha \partial_\alpha A^\mu \qquad (9.2.29)$$

$$A^\mu \to A^\mu - (D_\alpha \xi^\mu) A^\alpha + \xi^\alpha D_\alpha A^\mu. \qquad (9.2.30)$$

4. Tensors:

$$\tau^\mu{}_\nu \to \tau^\mu{}_\nu - (D_\alpha \xi^\mu) \tau^\alpha{}_\nu + (D_\nu \xi^\alpha) \tau^\mu{}_\alpha + \xi^\alpha D_\alpha \tau^\mu{}_\nu. \qquad (9.2.31)$$

Now,

$$g_{\mu\nu} \to g_{\mu\nu} + D_\mu \xi_\nu + D_\nu \xi_\mu. \qquad (9.2.32)$$

So this means $\sqrt{-g}$ also transforms. In fact, we can find the identities:

$$\boxed{\Gamma^\mu_{\mu\nu} = \frac{1}{\sqrt{-g}} \partial_\nu \sqrt{-g}} \qquad (9.2.33)$$

DERIVATION IS LEFT AS EXERCISE. And with this we can show that

$$\boxed{D_\mu V^\mu = \frac{1}{\sqrt{-g}} \partial_\mu \left(\sqrt{-g} V^\mu \right)} \qquad (9.2.34)$$

DERIVATION IS LEFT AS EXERCISE. With these identities, the Lie derivative of $\sqrt{-g}$ can be found.

$$\sqrt{-g} \to \sqrt{-g} + \partial_\alpha \left(\sqrt{-g} \xi^\alpha \right). \qquad (9.2.35)$$

And thus this says

$$\boxed{\mathcal{L}_\xi \sqrt{-g} = \partial_\alpha \left(\sqrt{-g} \xi^\alpha \right) = \sqrt{-g} D_\mu \xi^\mu} \qquad (9.2.36)$$

If $\mathcal{L}$ is a scalar under diffeomorphism, then

$$\mathcal{L} \to \mathcal{L} + \xi^\alpha \partial_\alpha \mathcal{L}. \qquad (9.2.37)$$

But what about the action S? Under diffeomorphism, since d^4x doesn't change,

$$S \to \int d^4x \left[\sqrt{-g} + \partial_\alpha \left(\sqrt{-g}\xi^\alpha\right)\right] \left(\mathcal{L} + \xi^\alpha \partial_\alpha \mathcal{L}\right) \tag{9.2.38}$$

$$= \int d^4x \sqrt{-g}\mathcal{L} + \int d^4x \left[\sqrt{-g}\xi^\alpha \partial_\alpha \mathcal{L} + \partial_\alpha \left(\sqrt{-g}\xi^\alpha\right)\mathcal{L}\right] \tag{9.2.39}$$

$$= S + \int d^4x \, \partial_\alpha \left(\sqrt{-g}\xi^\alpha \mathcal{L}\right). \tag{9.2.40}$$

Thus, using Gauss' theorem in four dimensions

$$S \to S + \int_{\partial\Omega} d^3x \, \hat{n}_\alpha \left(\sqrt{-g}\xi^\alpha \mathcal{L}\right). \tag{9.2.41}$$

As we push the 3-dimensional surface to infinity where $\xi^\alpha = 0$, $S \to S$, and thus $\delta S = 0$. The action is unchanged under diffeomorphism, and hence we also have a symmetry in general relativity.

9.2.3 Local Lorentz Transformation & Vierbein Formalism

Lorentz symmetry are local symmetries in local frames. Local Lorentz frames exist at every point, so we can always coordinate transform into a local Lorentz frame at a point P, so that the metric is Minkowskian, i.e.,

$$g_{\mu\nu} \to g_{\mu'\nu'} = X^\alpha_{\mu'} X^\beta_{\nu'} g_{\alpha\beta} = \eta_{\mu'\nu'} \tag{9.2.42}$$

at point P. However, if there is curvature, then $R^{\lambda'}{}_{\mu'\nu'\sigma'} \neq 0$ at P still holds, since the Riemann curvature tensor depends on $\partial\Gamma$, the partial derivatives of the Christoffel symbols.

There is another way, though, to go without using a coordinate transformation. That is to use **Vierbeins**.

Roughly speaking, **Vierbeins** are local vectors in a tangent space that performs "a change of basis."

$$\boxed{g_{\mu\nu} = e^a{}_\mu e^b{}_\nu \eta_{ab}} \tag{9.2.43}$$

With Vierbeins, we can redefined general relativity by making them dynamical. Any tensor can be written as

$$A_\mu = e_\mu{}^a A_a \tag{9.2.44}$$

where A_μ is a component in local Lorentz basis. There's also the concept of an inverse Vierbein:

$$e^\mu{}_a e_\nu{}^a = \delta^\mu_\nu \tag{9.2.45}$$

$$e^\mu{}_a e_\mu{}^b = \delta^a_b. \tag{9.2.46}$$

Then,

$$g^{\mu\nu} = e^\mu{}_a e^\nu{}_b \eta^{ab}. \tag{9.2.47}$$

It is easy to verify that $g^{\mu\alpha} g_{\alpha\nu} = \delta^\mu_\nu$:

$$g^{\mu\alpha} g_{\alpha\nu} = e^\mu{}_c e^\alpha{}_d \eta^{cd} e_\alpha{}^f e_\nu{}^g \eta_{fg} \tag{9.2.48}$$

$$= e^\mu{}_c \delta^f_d e_\nu{}^g \eta^{cd} \eta_{fg} \tag{9.2.49}$$

$$= e^\mu{}_c e_\nu{}^g \eta^{cf} \eta_{fg} \tag{9.2.50}$$

$$= e^\mu{}_c e_\nu{}^g \delta^c_g \tag{9.2.51}$$

$$= e^\mu{}_c e_\nu{}^c \tag{9.2.52}$$

$$= \delta^\mu_\nu. \tag{9.2.53}$$

$$\square$$

Next, we can also look at the determinant of $g_{\mu\nu}$, where $g_{\mu\nu} = e_\mu{}^a e_\nu{}^b \eta_{ab} = e_\mu{}^a \eta_{ab} e_\nu{}^b$:

$$g = \det(g) \tag{9.2.54}$$

$$= e\,(\det(\eta)e) \tag{9.2.55}$$

$$= e(-1)e \tag{9.2.56}$$

$$= e^2. \tag{9.2.57}$$

Therefore,

$$\boxed{\sqrt{-g} = e} \tag{9.2.58}$$

So, the action under Vierbein formalism is

$$\boxed{S = \int \sqrt{-g}\,d^4x\mathcal{L}(g_{\mu\nu}, A_\mu, \dots) = \int e\,d^4x\mathcal{L}(e_\mu{}^a, A_a, \dots)} \tag{9.2.59}$$

One of the advantages of the Vierbein formalism is the ability to incorportate **spinors** that deescribe Fermions into the action and Lagrangian. Spinors are neither vectors nor tensors nor scalars. In other words, there are no vector/tensor representation for Fermions. They are just spinors - a completely different mathematical object. There is no problem dealing with Fermions in special relativity, since special relativity has spinor representation under the Lorentz group. However, general relatibity has no representation for these under diffeomorphisms if we don't use Vierbeins.

In particular, in the relativistic quantum theory developed by Paul Dirac, the Dirac Lagrangian is given by

$$\mathcal{L} = i(\partial_\mu \bar{\psi})\gamma^\mu \psi, \tag{9.2.60}$$

which is a mixture of spinors and vectors, where γ^μ is the four-component spinor. The theory actually makes use of a four components, with two describing ordinary Fermions, while the other two describes anti-particles.

Back to general relativity. Vierbeins allow us to include Fermions into general relativity, making them vectors in spacetime. Since there is no unique local Lorentz basis at any point P, we can rotate or boost. And so, these vectors in spacetime must follows transformation rules:

$$e_\mu{}^a \to e_\mu{}^a + \epsilon^a{}_b e_\mu{}^b. \tag{9.2.61}$$

But Vierbeins are also vectors under diffeomorphisms:

$$e_\mu{}^a \to e_\mu{}^a + (D_\mu \xi^\alpha) e_\alpha{}^a + \xi^\alpha D_\alpha e_\mu{}^a. \tag{9.2.62}$$

With these, we can also look at how the metric transforms. First, under diffeomorphism:

$$g_{\mu\nu} = e_\mu{}^a e_\nu{}^b \eta_{ab} \tag{9.2.63}$$
$$\to \left(e_\mu{}^a + (D_\mu \xi^\alpha) e_\alpha{}^a + \xi^\alpha D_\alpha e_\mu{}^a\right) \left(e_\nu{}^b + (D_\nu \xi^\alpha) e_\alpha{}^b + \xi^\alpha D_\alpha e_\mu{}^b\right) \eta_{ab}. \tag{9.2.64}$$

Expand this out to first order, we get

$$g_{\mu\nu} \to g_{\mu\nu} + (D_\mu \xi^\alpha) e_\alpha{}^a e_\nu{}^b \eta_{ab} + \xi^\alpha D_\alpha e_\mu{}^a e_\nu{}^b \eta_{ab}$$
$$+ e_\mu{}^a (D_\nu \xi^\alpha) e_\alpha{}^b \eta_{ab} + e_\mu{}^a \xi^\alpha D_\alpha e_\nu{}^b \eta_{ab} \tag{9.2.65}$$
$$= g_{\mu\nu} + (D_\mu \xi^\alpha) g_{\alpha\nu} + (D_\nu \xi^\alpha)_{\mu\alpha} + \xi^\alpha D_\alpha \left(e_\mu{}^a e_\nu{}^b \eta_{ab}\right) \tag{9.2.66}$$
$$= g_{\mu\nu} + (D_\mu \xi^\alpha) g_{\alpha\nu} + (D_\nu \xi^\alpha)_{\mu\alpha} + \xi^\alpha D_\alpha g_{\mu\nu} \tag{9.2.67}$$
$$= g_{\mu\nu} + (D_\mu \xi^\alpha) g_{\alpha\nu} + (D_\nu \xi^\alpha)_{\mu\alpha} \tag{9.2.68}$$
$$= g_{\mu\nu} + \mathcal{L}_\xi g_{\mu\nu} \tag{9.2.69}$$

as expected, where we have used the fact that the metric is covariantly constant. On other hand, what about under local Lorentz transformations?

$$g_{\mu\nu} = e_\mu{}^a e_\nu{}^b \eta_{ab} \tag{9.2.70}$$
$$\to \left(e_\mu{}^a + \epsilon^a{}_c e_\mu{}^b\right) \left(e_\nu{}^b + \epsilon^b{}_c e_\nu{}^c\right) \eta_{ab} \tag{9.2.71}$$
$$= g_{\mu\nu} + \epsilon^a{}_c e_\mu{}^b e_\nu{}^b \eta_{ab} + e_\mu{}^a \epsilon^b{}_c e_\nu{}^c \eta_{ab} + \dots \tag{9.2.72}$$
$$= g_{\mu\nu} + \epsilon_{bc} e_\mu{}^b e_\nu{}^b + e_{ac} \epsilon^b{}_c e_\nu{}^c \tag{9.2.73}$$
$$= g_{\mu\nu} + \epsilon_{bc} e_\mu{}^b e_\nu{}^b + e_{bc} \epsilon^b{}_c e_\nu{}^c \tag{9.2.74}$$
$$= g_{\mu\nu} + \epsilon_{bc} \left(e_\mu{}^b e_\nu{}^b + \epsilon^b{}_c e_\nu{}^c\right) \tag{9.2.75}$$
$$= g_{\mu\nu} + \epsilon_{cb} \left(e_\mu{}^c e_\nu{}^c + \epsilon^c{}_b e_\nu{}^b\right). \tag{9.2.76}$$

But due the anti-symmetry in Lorentz transformations:

$$\epsilon_{bc} = -\epsilon_{cb} \tag{9.2.77}$$

it must hold that

$$-\left(e_\mu{}^b e_\nu{}^b + \epsilon^b{}_c e_\nu{}^c\right) = \left(e_\mu{}^c e_\nu{}^c + \epsilon^c{}_b e_\nu{}^b\right) \tag{9.2.78}$$

which means

$$e_\mu{}^c e_\nu{}^c + \epsilon^c{}_b e_\nu{}^b = 0, \tag{9.2.79}$$

and so

$$g_{\mu\nu} \to g_{\mu\nu} \tag{9.2.80}$$

under local Lorentz transformations. Therefore, the metric is invariant under local Lorentz transformation, as we wanted. Now, note that the metric has 10 independent components due to its symmetry: $g_{\mu\nu} = g_{\nu\mu}$. Meanwhile, the Vierbeins have 16 independent components. These 6 extra components are Lorentz degrees of freedom. This means we can make 6 local Lorentz transformations with $\epsilon_{ab} = \epsilon_{ba}$, i.e., we can gauge away 6 of these components using local Lorentz transformations, giving 10 effectively independent components.

Now, if we consider how a Lagrangian made entire out of scalars, how would it transform under local Lorentz transformations using the Vierbeins? Consider $\mathcal{L} = A_\mu A^\mu$, then under the Vierbein formalism:

$$A_\mu A^\mu = \left(e_\mu{}^a A_a\right)\left(e^\mu{}_b A^b\right). \tag{9.2.81}$$

Under local Lorentz transformations,

$$\begin{aligned}
A_\mu A^\mu &\to \left(e_\mu{}^a + \epsilon^a{}_c e_\mu{}^c\right)\left(A_a + \epsilon_a{}^c A_c\right)\left(e^\mu{}_b + \epsilon_b{}^c e^\mu{}_c\right)\left(A^b + \epsilon^b{}_c A^c\right) \\
&= A_\mu A^\mu + \epsilon^a{}_c e_\mu{}^c A_a e^\mu{}_b A^b + e_\mu{}^a \epsilon_a{}^c A_c e^\mu{}_b A^b \\
&\quad + e_\mu{}^a A_a \epsilon_b{}^c e^\mu{}_c A^b + e_\mu{}^a A_a e^\mu{}_b \epsilon^b{}_c A^c.
\end{aligned} \tag{9.2.82}$$

Now, notice that

$$\begin{aligned}
\epsilon^a{}_c e_\mu{}^c A_a e^\mu{}_b A^b + e_\mu{}^a \epsilon_a{}^c A_c e^\mu{}_b A^b &= \epsilon_{ac}\left(e_\mu{}^c A^c e^\mu{}_b A^b + \epsilon_a{}^c A^a e^\mu{}_b A^b\right) \tag{9.2.83} \\
&= \epsilon_{ca}\left(e_\mu{}^a A^a e^\mu{}_b A^b + \epsilon_c{}^a A^a e^\mu{}_b A^b\right) \tag{9.2.84} \\
&= -\epsilon_{ac}\left(e_\mu{}^c A^c e^\mu{}_b A^b + \epsilon_a{}^c A^a e^\mu{}_b A^b\right) \tag{9.2.85} \\
&= 0. \tag{9.2.86}
\end{aligned}$$

Therefore, if the action is given in terms of scalars:

$$S = \int e\, d^4x\, \mathcal{L}(g_{\mu\nu}, \phi, A_\mu, \dots) \tag{9.2.87}$$

where the Lagrangian is a scalar, then the action is invariant under local Lorentz transformations. But without Vierbeins, Fermions are not allowed in the theory. Under the Vierbein formalism, we not only have a scalar action that is invariant under local Lorentz transformations, but also Fermions incorporated into the

theory.

But there is more to the story. We can ask about what happens to derivatives in the Vierbein formalism. In the general relativity that we are familiar with, we have seen the notion of the covariant derivative:

$$D_\mu A_\nu = \partial_\mu A_\nu - \Gamma^\lambda_{\mu\nu} A_\lambda. \tag{9.2.88}$$

But what about $D_\nu e_\mu{}^a$? It turns out that this is not a tensor under both local Lorentz transformations and diffeomorphism unless we change the definition of the derivative once again. To make this work, we need a second type of connection:

$$D_\nu e_\mu{}^a = \partial_\nu e_\mu{}^a - \Gamma^\lambda_{\nu\mu} e_\lambda{}^a + \omega_\nu{}^a{}_b e_\mu{}^b \tag{9.2.89}$$

where ω is called the **spin connection**, which makes $D_\nu e_\mu{}^a$ a tensor. It also turns out that

$$\omega_\nu{}^{ab} = -\omega_\nu{}^{ba}, \tag{9.2.90}$$

and so there are 24 independent components for the spin connection. We can then bring the definition of the spin connection to the Riemann curvature tensor, which has the following form

$$R^\lambda{}_{\mu\nu\sigma} = \Gamma\Gamma - \Gamma\Gamma + \partial\Gamma - \partial\Gamma. \tag{9.2.91}$$

Adding the spin connection gives us **torsion**, and the Riemann space becomes Riemann-Cartan space.

9.2.4   Riemann-Cartan spacetime

Note: This subsubsection is taken from Robert Bluhm's notes.

In graviattional theories that include fields with intrinsic spin, such as spin-$\frac{1}{2}$ fermions, a new geometrical quantity known as torsion $T^\lambda{}_{\mu\nu}$ can be introduced. It is defined as

$$\boxed{T^\lambda{}_{\mu\nu} = \Gamma^\lambda{}_{\mu\nu} - \Gamma^\lambda{}_{\nu\mu}} \tag{9.2.92}$$

In general relativity, where the connection $\Gamma^\lambda{}_{\mu\nu}$ is symmetric, the torsion $T^\lambda{}_{\mu\nu}$ simply vanishes. So in which case(s) does the torsion not vanish? The curvature tensor is the relevant geometrical quantity in general relativity, which is sourced by energy-momentum. The spacetime in general relativity that we have been dealing with so far is (pseudo)-Riemannian. However, as we have pointed out earlier, general relativity theory in this space cannot include fermions. When spin is present, this acts as a source of spacetime torsion, and the resulting spacetime generalizes to hat is called a Riemann-Cartan spacetime.

In a gravitational theory in Riemann-Cartan spacetime using vierbein formalism, both the vierbein and the spin connection are treated as independent fields. Thus a generic action can be written as

$$S = \int d^4x \, e\mathcal{L}(e_\mu{}^a, \omega_\mu{}^a{}_b, \phi, \psi, A^\mu, D_\mu\psi, \dots), \qquad (9.2.93)$$

where $e = \sqrt{-g}$ and $\phi, \psi, A^\mu, D_\mu\psi, \dots$ denote conventional matter fields and their derivatives. If this action is varied with respect to the vierbein, the result is Einstein's equations relating the curvature to the energy momentum density. Alternatively, if the action is varied with respect to the spin connection, the result is the equations relation the torsion with the spin density.

Using a vierbein treatment and allowing a Riemann-Cartan geometry, it becomes straightforward to introduce fermions. The starting point is the usual description in terms of Dirac theory in special relativity. Quantities such as the Dirac matrices γ^a are defined in a local Lorentz frame where the matrix is Minkowskian η_{ab}. The vierbein is then used to define the corresponding quantity in the spacetime frame, for example, the Dirac matrices become $e^\mu{}_a\gamma^a$. Partial derivatives acting on fermions fields become covariant derivatives involving the spin connection,

$$\partial_\mu\psi \to D_\mu\psi = \partial_\mu\psi + \frac{i}{4}\omega_{\mu ab}\sigma^{ab}\psi, \qquad (9.2.94)$$

where σ^{ab} is the commutation $\frac{i}{2}[\gamma^a, \gamma^b]$.

The Lagrangian in the vierbein formalism must be a scalar with respect to both diffeomorphism and local Lorentz transformations. For tensor quantities, scalar combinations of tensors can be combined using either spacetime components or local components. For example, a quadratic mass term for a vector field can be written as

$$\mathcal{L} = A^\mu g_{\mu\nu} A^\nu \qquad (9.2.95)$$
$$= e^\mu{}_a A^a e_\mu{}^c e_\nu{}^d \eta_{cd} e^\nu{}_b A^b \qquad (9.2.96)$$
$$= \delta^c_a A^a \eta_{cd} \delta^d_b A^b \qquad (9.2.97)$$
$$= A^a \eta_{ab} A^b, \qquad (9.2.98)$$

which we have shown before. So, the vierbein drops out of scalar combinations that do not involve derivatives, so that the combination can be witten in terms of the local components A^a and the local metric η_{ab}.

In deriving the equations of motion of the a gravitational theory involving tensor quantities, such as a vector, it must there be decided whether the basic quantities to vary in the scalar Lagrangian are the spacetime components A^μ or the local quantities A^a. This choice affects how the vierbein appears. It therefore also affects the definition of the energy-momentum in a vierbein formalism,

which is obtained inside the action as

$$T_e^{\mu\mu} = \frac{1}{e}e^{\nu a}\frac{\delta(e\mathcal{L})}{\delta e_\mu{}^a}.\qquad(9.2.99)$$

The result of varying with respect to th vierbein is Einstein's equations (with $\Lambda = 0$),

$$G^{\mu\nu} = 8\pi G T_e^{\mu\nu}.\qquad(9.2.100)$$

In general, he form of the energy-momentum tensor in the vierbein formalism $T_e^{\mu\nu}$ will differ from that of the energy-momentum obtained using the metric. This is especially the case when there is torsion. However, for consistency, the energy-momentum tensor $T_e^{\mu\nu}$ is a vierbein treatment must be divergence-free,

$$D_\mu T_e^{\mu\nu} = 0,\qquad(9.2.101)$$

meaning that it is still a conserved quantity.

Chapter 10

Quantization

Chapter 11

Worked Problems

11.1 Massive Vector Fields

The goal of this problem set is to familiarize ourselves with varying the Lagrangian, as well as showing that there are 3 degrees of freedom to a massive vector gauge field.

Consider the following Lagrangian for problems 1-6.

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m^2 A_\mu A^\mu \tag{11.1.1}$$

which describes a massive particle with a vector field.

Problem 11.1.1. Verify that this Lagrangian does not have gauge symmetry under the gauge transformation $A_\mu \to A_\mu + \partial_\mu \Lambda$, i.e. $\mathcal{L} \not\to \mathcal{L}$.

Solution 11.1.1. We simply brute force through the problem and show that $\mathcal{L}' \neq \mathcal{L}$, where:

$$\mathcal{L}' = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m^2 A'_\mu A'^\mu. \tag{11.1.2}$$

Note that even though $F^{\mu\nu}$ is defined in terms of A_μ, $F_{\mu\nu}$ is invariant under the gauge transformation because it is the 4-dimensional curl of A_μ. We can verify this fact later. Focusing only on the massive term in the new Lagrangian:

$$\mathcal{L}' = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m^2 A'_\mu A'^\mu \tag{11.1.3}$$

$$= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m^2 (A_\mu + \partial_\mu\Lambda)(A^\mu + \partial^\mu\Lambda) \tag{11.1.4}$$

$$= \mathcal{L} + \frac{1}{2}m^2 (A_\mu\partial^\mu\Lambda + A^\mu\partial_\mu\Lambda + \partial_\mu\Lambda\partial^\mu\Lambda) \tag{11.1.5}$$

It is clear that our choice of the scalar field Λ affects the form of the new Lagrangian. So, $\mathcal{L} \neq \mathcal{L}'$ in general. Therefore, the massive vector Lagrangian does not have gauge symmetry under the given gauge transformation.

As an aside, we can verify that $F_{\mu\nu} \to F_{\mu\nu}$. The $F^{\mu\nu} \to F^{\mu\nu}$ identity will follow accordingly.

$$F'_{\mu\nu} = \partial_\mu A'_\nu - \partial_\nu A'_\mu \tag{11.1.6}$$

$$= \partial_\mu(A_\nu + \partial_\nu \Lambda) - \partial_\nu(A_\mu + \partial_\mu \Lambda) \tag{11.1.7}$$

$$= F_{\mu\nu} + \partial_\mu \partial_\nu \Lambda - \partial_\nu \partial_\mu \Lambda \tag{11.1.8}$$

$$= F_{\mu\nu}, \tag{11.1.9}$$

where $\partial_\mu \partial_\nu \Lambda - \partial_\nu \partial_\mu \Lambda = 0$ because Λ is a scalar (this is also true by index-swapping). So the electromagnetic field tensor has gauge symmetry under the given gauge transformation.

Problem 11.1.2. Vary the action

$$S = \int \mathcal{L} \, d^4 x \qquad (11.1.10)$$

with respect to A_μ and show that the equations of motion are

$$\Box A - \partial_\mu \partial^\nu A_\nu + m^2 A_\mu = 0. \qquad (11.1.11)$$

Solution 11.1.2. The Lagrangian is given by

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} m^2 A_\mu A^\mu. \qquad (11.1.12)$$

Varying the term with the electromagnetic field tensor gives

$$\delta \left(F^{\mu\nu} F_{\mu\nu} \right) = 2\delta \left[\left(\partial^\mu A^\nu - \partial^\nu A^\mu \right) \left(\partial_\mu A_\nu - \partial_\nu A_\mu \right) \right] \qquad (11.1.13)$$

$$= 2\delta \left(\partial^\mu A^\nu \, \partial_\mu A_\nu - \partial_\mu A_\nu \, \partial^\mu A^\nu \right) \qquad (11.1.14)$$

$$= 2 \left(\partial^\mu \delta A^\nu \, \partial_\mu A_\nu + \partial^\mu A^\nu \, \partial_\mu \delta A_\nu - \partial_\mu \delta A_\nu \, \partial^\mu A^\nu - \partial_\mu A_\nu \, \partial^\mu \delta A^\nu \right) \qquad (11.1.15)$$

$$= 2 \left(\partial^\mu \delta A^\nu \, \partial_\mu A_\nu + \partial_\nu A_\mu \, \partial^\nu \delta A^\mu - \partial_\mu \delta A_\nu \, \partial^\mu A^\nu - \partial^\nu A^\mu \, \partial_\nu \delta A_\mu \right) \qquad (11.1.16)$$

$$= 4 \left(\partial^\nu \delta A^\mu \, \partial_\nu A_\mu - \partial_\mu \delta A_\nu \, \partial^\mu A^\nu \right). \qquad (11.1.17)$$

Next, we integrate by parts to get the boundary terms (which will vanish when we vary the action):

$$\partial^\nu \delta A^\mu \, \partial_\nu A_\mu = \partial^\nu (\partial_\mu A_\nu \, \delta A^\mu) - (\partial^\nu \partial_\mu A_\nu) \delta A^\mu = \partial_\mu (\partial^\nu A^\mu \, \delta A_\nu) - (\partial^\nu \partial_\mu A_\nu) \delta A^\mu \qquad (11.1.18)$$

$$\partial_\mu \delta A_\nu \partial^\mu A^\nu = \partial_\mu (\partial^\mu A^\nu \, \delta A_\nu) - \partial_\mu (\partial^\mu A^\nu) \, \delta A_\nu = \partial_\mu (\partial^\mu A^\nu \, \delta A_\nu) - \partial_\mu (\partial^\mu A^\nu) \, \delta A_\nu. \qquad (11.1.19)$$

This gives

$$\delta \left(-\frac{1}{4} F^{\mu\nu} F_{\mu\nu} \right) = \partial_\mu (\partial^\mu A^\nu) \, \delta A_\nu - (\partial^\nu \partial_\mu A_\nu) \delta A^\mu \qquad (11.1.20)$$

$$= \left(\Box A^\mu - \partial^\mu \partial^\nu A_\nu \right) \delta A_\mu \qquad (11.1.21)$$

$$= \left(\Box A_\mu - \partial_\mu \partial^\nu A_\nu \right) \delta A^\mu \qquad (11.1.22)$$

Finally, we vary the term with the vector potential A_μ with respect to A_μ:

$$\delta \left(\frac{1}{2} m^2 A_\mu A^\mu \right) = \frac{1}{2} m^2 \left(\delta A_\mu A^\mu + A_\mu \delta A^\mu \right) \qquad (11.1.23)$$

$$= \frac{1}{2} m^2 \left(\delta A_\mu A^\mu + A^\mu \delta A_\mu \right) \qquad (11.1.24)$$

$$= \left(m^2 A^\mu \right) \delta A_\mu \qquad (11.1.25)$$

$$= \left(m^2 A_\mu \right) \delta A^\mu. \qquad (11.1.26)$$

Putting everything together,

$$\delta\mathcal{L} = \left(\Box A_\mu - \partial_\mu\partial^\nu A_\nu m^2 A_\mu\right)\delta A^\mu. \tag{11.1.27}$$

Varying the action and requiring $\delta S = 0$ for all δA^μ forces $\delta\mathcal{L} = 0$. We arrive at the equations of motion

$$\Box A_\mu - \partial_\mu\partial^\nu A_\nu + m^2 A_\mu = 0 \tag{11.1.28}$$

as desired.

Problem 11.1.3. Take ∂^μ of the Lagrangian and show that it gives a constraint: $\partial^\mu A_\mu = 0$. Use this to simplify the equations of motion to

$$\Box A_\mu - m^2 A_\mu = 0. \tag{11.1.29}$$

Solution 11.1.3. The given Lagrangian has no x^μ dependence, so we expect $\partial^\mu \mathcal{L} = 0$. We will show that this adds a constraint to A^μ.

$$\partial^\mu \mathcal{L} = \partial^\mu \left(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} m^2 A_\mu A^\mu \right) \tag{11.1.30}$$

$$= -\frac{1}{4} \partial^\mu \left(F_{\mu\nu} F^{\mu\nu} \right) + \frac{1}{2} m^2 \left(\partial^\mu A_\mu A^\mu + A_\mu \partial^\mu A^\mu \right) \tag{11.1.31}$$

$$= -\frac{1}{2} \partial^\mu \left(\partial^\mu A^\nu \, \partial_\mu A_\nu - \partial_\mu A_\nu \, \partial^\mu A^\nu \right) + m^2 A^\mu \partial^\mu A_\mu \tag{11.1.32}$$

$$= -\frac{1}{2} \left[(\partial^\mu \partial^\mu A^\nu)(\partial_\mu A_\nu) - (\partial^\mu \partial^\mu A^\nu)(\partial_\mu A_\nu) \right. \tag{11.1.33}$$

$$\left. + (\Box A_\nu)(\partial^\mu A^\nu) - (\Box A_\nu)(\partial^\mu A^\nu) \right] + m^2 A^\mu \partial^\mu A_\mu. \tag{11.1.34}$$

$$= m^2 A^\mu \partial^\mu A_\mu. \tag{11.1.35}$$

Since $\partial^\mu \mathcal{L} = 0$, a massive $m \neq 0$ field is assumed, and A^μ is arbitrary, $\partial^\mu A_\mu = 0$ must hold as desired. With this, the equations of motion become:

$$\Box A_\mu + m^2 A_\mu = 0. \tag{11.1.36}$$

Notice that the form of the equations of motion is exactly the same as that of the Klein-Gordon equation.

Problem 11.1.4. Look for a solution of the form

$$A_\mu = \epsilon_\mu e^{-i\mathbf{k}\cdot\mathbf{x}}, \tag{11.1.37}$$

where ϵ_μ is the polarization vector, and show that it's a solution provided that

$$K^\mu K_\mu = m^2, \tag{11.1.38}$$

i.e., the vector field is massive.

Solution 11.1.4. We use the simplified equation of motion from the previous problem:

$$\Box A_\mu + m^2 A_\mu = \Box \epsilon_\mu e^{-i\mathbf{k}\cdot\mathbf{x}} + m^2 \epsilon_\mu e^{-i\mathbf{k}\cdot\mathbf{x}} \tag{11.1.39}$$

$$= \epsilon_\mu \Box e^{-iK_\alpha X^\alpha} + m^2 \epsilon_\mu e^{-iK_\alpha X^\alpha} \tag{11.1.40}$$

$$= \epsilon_\mu \partial^\beta (-iK_\alpha) e^{-iK_\alpha X^\alpha} \partial_\beta X^\alpha + m^2 \epsilon_\mu e^{-iK_\alpha X^\alpha} \tag{11.1.41}$$

$$= \epsilon_\mu (-iK_\beta) \eta^{\beta\gamma} \partial_\gamma e^{-iK_\beta X^\beta} + m^2 \epsilon_\mu e^{-iK_\alpha X^\alpha} \tag{11.1.42}$$

$$= A_\mu \left[(-iK_\beta)(-iK_\beta) \eta^{\beta\gamma} \delta_\gamma^\beta + m^2 \right] \tag{11.1.43}$$

$$= A_\mu \left(-K_\beta K^\beta + m^2 \right) \tag{11.1.44}$$

$$= 0. \tag{11.1.45}$$

This means A_μ is a solution for any arbitrary ϵ_μ if $K^\beta K_\beta = 0$, i.e., the vector field is massive.

Problem 11.1.5. Suppose that $K^\mu = (K^0, 0, 0, K^3)$, which denotes moving in the z-direction. Show that $\partial^\mu A_\mu = 0$ implies

$$\epsilon_0 = \frac{-K^3}{K^0} \epsilon_3 \tag{11.1.46}$$

or

$$\epsilon_\mu = \left(\frac{K^3}{K^0} \epsilon_3, \epsilon_1, \epsilon_2, \epsilon_3 \right)^\top. \tag{11.1.47}$$

Solution 11.1.5. We simply the requirement $\partial^\mu A_\mu = 0$:

$$\partial^\mu A_\mu = \eta^{\mu\nu} \partial_\nu A_\mu \tag{11.1.48}$$

$$= \eta^{\mu\nu} \epsilon_\mu (-iK_\alpha) e^{-iK_\alpha X^\alpha} \partial_\nu X^\alpha \tag{11.1.49}$$

$$= \eta^{\mu\nu} \epsilon_\mu (-iK_\alpha) e^{-iK_\alpha X^\alpha} \delta_\alpha^\nu \tag{11.1.50}$$

$$= -iK_\alpha \epsilon^\alpha e^{-iK_\alpha X^\alpha} \tag{11.1.51}$$

$$= 0. \tag{11.1.52}$$

This means $K_\alpha \epsilon^\alpha = 0$. But since $K^\alpha = (K^0, 0, 0, K^3)$, we require that

$$\epsilon_0 K^0 = \epsilon_3 K^3, \tag{11.1.53}$$

i.e.,

$$\epsilon_0 = \frac{-K^3}{K^0} \epsilon_3. \tag{11.1.54}$$

So,

$$\epsilon_\mu = \left(\frac{K^3}{K^0} \epsilon_3, \epsilon_1, \epsilon_2, \epsilon_3 \right)^\top. \tag{11.1.55}$$

Problem 11.1.6. Show that this gives 3 independent modes of motion

$$\begin{cases} \epsilon_\mu^{(1)} = (0,1,0,0)^\top \\ \epsilon_\mu^{(2)} = (0,0,1,0)^\top \\ \epsilon_\mu^{(3)} = \left(\frac{K^3}{K^0},0,0,1\right)^\top \end{cases} \tag{11.1.56}$$

where the first two modes are transverse modes and the third is longitudinal.

As expected for a massive vector, it has spin 1, $j = 1$ with 3 polarizations $m_j = -1,0,1$. This is in contrast to a massless vector that only has 2 transverse polarizations.

Solution 11.1.6. Since the polarization vector ϵ_μ has the form

$$\epsilon_\mu = \left(\frac{K^3}{K^0}\epsilon_3, \epsilon_1, \epsilon_2, \epsilon_3\right)^\top, \tag{11.1.57}$$

there are only 3 "degrees of freedom." More precisely, if $\mathcal{E}$ is a vector space spanned by all ϵ_ν, then $\dim(\mathcal{E}) = 3$. To decompose $\mathcal{E}$, we simply choose a simple basis set of ϵ_μ for $\mathcal{E}$ with three linearly independent ϵ_μ's:

$$\begin{cases} \epsilon_\mu^{(1)} = (0,1,0,0)^\top \\ \epsilon_\mu^{(2)} = (0,0,1,0)^\top \\ \epsilon_\mu^{(3)} = \left(\frac{K^3}{K^0},0,0,1\right)^\top \end{cases} . \tag{11.1.58}$$

11.2 Spontaneous Symmetry Breaking of a U(1) Gauge Theory

Problem 11.2.1. Global U(1) symmetry

Consider the following Lagrangian:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi^\dagger \phi) \tag{11.2.1}$$

for the complex scalar field ϕ:

$$\phi = \phi_1 + i\phi_2 \tag{11.2.2}$$
$$\phi^\dagger = \phi_1 - i\phi_2 \tag{11.2.3}$$
$$\phi^\dagger \phi = \phi_1^2 + \phi_2^2 \tag{11.2.4}$$

where ϕ_1 and ϕ_2 are real. Also, consider the potential

$$V(\phi^\dagger \phi) = \frac{1}{2}m^2 \phi^\dagger \phi + \frac{1}{4}\lambda(\phi^\dagger \phi)^2. \tag{11.2.5}$$

1. Show that $\mathcal{L}$ is invariant under a global U(1) transformation:

$$\phi \to \phi' = e^{i\alpha}\phi, \tag{11.2.6}$$

   where α is a real constant.

2. Suppose that $m^2 < 0$. Let

$$X = \phi^\dagger \phi, \tag{11.2.7}$$

   then

$$V(X) = \frac{1}{2}m^2 X + \frac{1}{4}\lambda X^2. \tag{11.2.8}$$

   Find the minimum of $V(X)$. Let

$$\mathcal{V} = \sqrt{\frac{-m^2}{\lambda}}. \tag{11.2.9}$$

   Show that the minimum is degenerate and find the equation in ϕ_1, ϕ_2 describing the space of possible vacua.

3. Let us introduce "polar" variables:

$$\phi = \rho e^{i\theta} \tag{11.2.10}$$
$$\phi^\dagger = \rho e^{-i\theta} \tag{11.2.11}$$

where $\rho(x), \theta(x)$ are fields that replace the expression of ϕ in terms of ϕ_1 and ϕ_2. What is the space of vacuum in terms of ρ and θ?

4. Spontaneously pick a vacuum state.

$$\langle \phi \rangle = \mathcal{V} \tag{11.2.12}$$

$$\langle \theta \rangle = 0 \tag{11.2.13}$$

so that

$$\langle \phi \rangle = \mathcal{V}. \tag{11.2.14}$$

Then consider small excitations about the vacuum:

$$\rho = \mathcal{V} + h \tag{11.2.15}$$

$$\theta = 0 + \frac{\mathcal{E}}{\mathcal{V}} \tag{11.2.16}$$

to get

$$\phi \approx (\mathcal{V} + h)e^{i\mathcal{E}/\mathcal{V}}. \tag{11.2.17}$$

Find the Lagrangian in terms of $h, \mathcal{E}, \mathcal{V}$. Truncate the Lagrangian to second-order terms in $h, \mathcal{E}$.

5. Identify the Higgs and Nambu-Goldstone fields.

6. Vary the Lagrangian with respect to h and $\mathcal{E}$. What equations of motion do you get? What are the freely propagating solutions for h and $\mathcal{E}$ at linear order?

Solution 11.2.1.

1. Under a U(1) transformation, $\phi' = \phi e^{i\alpha}\phi$ where α is constant. Notice that $\phi'^\dagger \phi' = \phi^\dagger \phi$. So, the new Lagrangian is

$$\mathcal{L}' = \frac{1}{2}(\partial_\mu \phi')^\dagger (\partial^\mu \phi') - V(\phi^\dagger \phi) \tag{11.2.18}$$

$$= \frac{1}{2}\left(e^{i\alpha}\partial_\mu \phi\right)^\dagger \left(e^{i\alpha}\partial^\mu \phi\right) - V(\phi^\dagger \phi) \tag{11.2.19}$$

$$= \frac{1}{2}e^{i\alpha - i\alpha}(\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi^\dagger \phi) \tag{11.2.20}$$

$$= \frac{1}{2}(\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi^\dagger \phi) \tag{11.2.21}$$

$$= \mathcal{L}. \tag{11.2.22}$$

So, this Lagrangian is invariant under a global U(1) transformation.

2. Let $X = \phi^\dagger \phi$. Then, we can rewrite $V(\phi^\dagger \phi)$ as

$$V(\phi^\dagger \phi) = V(X) = \frac{1}{2}m^2 X + \frac{1}{4}\lambda X^2. \qquad (11.2.23)$$
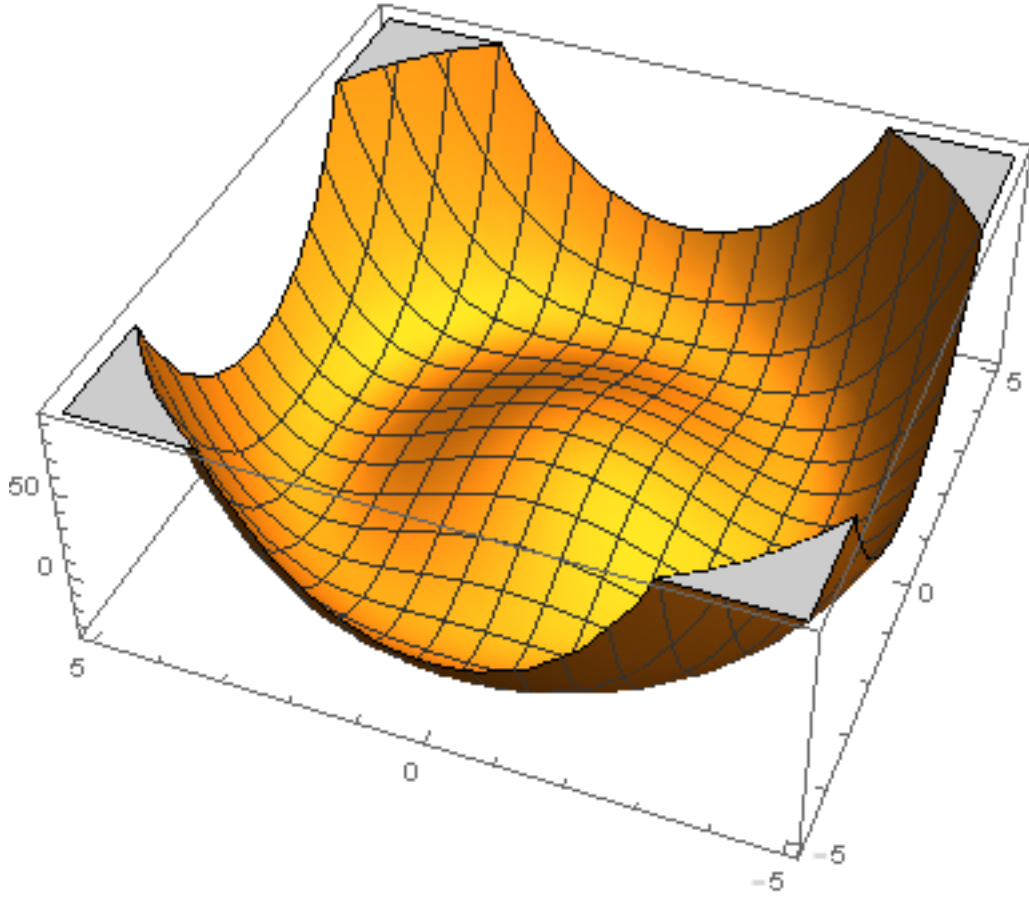
To find X that gives an extremal value of $V(X)$, we take dV/dX and set to zero:

$$0 = \frac{dV}{dX} = \frac{1}{2}m^2 + \frac{1}{2}\lambda X, \qquad (11.2.24)$$

which gives

$$X = \frac{-m^2}{\lambda}. \qquad (11.2.25)$$

Since we know V has a Mexican hat shape, at this $X \neq 0$, V is minimal.

Call

$$\mathcal{V} = \sqrt{\frac{-m^2}{\lambda}} = \sqrt{X} = \sqrt{\phi^\dagger \phi} = \sqrt{\phi_1^2 + \phi_2^2}. \qquad (11.2.26)$$

This shows there is a degeneracy at minimum V (a circle of radius $\mathcal{V}$). The equation in ϕ_1, ϕ_2 describing the space of possible vacua is just

$$\phi_1^2 + \phi_2^2 = \mathcal{V}^2. \qquad (11.2.27)$$

3. Let the polar variables be given. In terms of ρ and θ, the space of vacuum is just

$$\phi_1^2 + \phi_2^2 = \phi^\dagger \phi = \rho e^{i\theta} \rho e^{-i\theta} = \rho^2 = \mathcal{V}^2. \qquad (11.2.28)$$

We notice that the expression is independent of θ, which makes sense since excitations in θ does not change the energy level (i.e. not escaping degeneracy).

4. We spontaneously pick a vacuum state such that $\langle \rho \rangle = \mathcal{V}$, $\langle \theta \rangle = 0$, so that $\langle \phi \rangle = \mathcal{V}$. Consider small excitations about this vacuum:

$$\rho \to \mathcal{V} + h \qquad (11.2.29)$$
$$\theta \to 0 + \mathcal{E}/\mathcal{V} \qquad (11.2.30)$$

then

$$\phi = \rho e^{i\theta} \approx (\mathcal{V} + h)e^{i\mathcal{E}/\mathcal{V}}, \qquad (11.2.31)$$

so

$$\begin{aligned}
V(\phi^\dagger \phi) &= V((\mathcal{V} + h)e^{i\mathcal{E}/\mathcal{V}}(\mathcal{V} + h)e^{-i\mathcal{E}/\mathcal{V}}) \\
&= \frac{1}{2}m^2(\mathcal{V} + h)^2 + \frac{1}{4}\lambda(\mathcal{V} + h)^4 \\
&= \frac{1}{2}m^2(\mathcal{V}^2 + 2\mathcal{V}h + h^2) + \frac{1}{4}\lambda(\mathcal{V}^4 + 4\mathcal{V}^3 h \\
&\quad + 6\mathcal{V}^2 h^2 + 4\mathcal{V}h^3 + h^4). \qquad (11.2.32)
\end{aligned}$$

Before we get the new Lagrangian, we can differentiate the ϕ's first to simplify our simplification later:

$$\begin{aligned}
\partial_\mu \phi &= \partial_\mu(\mathcal{V} + h)e^{i\mathcal{E}/\mathcal{V}} \\
&= \partial_\mu h e^{i\mathcal{E}/\mathcal{V}} + (\mathcal{V} + h)\frac{i}{\mathcal{V}}e^{i\mathcal{E}/\mathcal{V}}\partial_\mu \mathcal{E} \qquad (11.2.33)
\end{aligned}$$

The ∂^μ differentiation is similar to this, differing only in the position of the index μ. So, the new Lagrangian (truncated to the second-order terms) is

then

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

$$= \frac{1}{2}\left((\partial_\mu h)e^{-i\mathcal{E}/\mathcal{V}} + (\mathcal{V}+h)\frac{-i}{\mathcal{V}}e^{-i\mathcal{E}/\mathcal{V}}\partial_\mu \mathcal{E}\right)$$

$$\times \left(\partial^\mu h e^{i\mathcal{E}/\mathcal{V}} + (\mathcal{V}+h)\frac{i}{\mathcal{V}}e^{i\mathcal{E}/\mathcal{V}}\partial^\mu \mathcal{E}\right)$$

$$- V(\phi^\dagger \phi). \tag{11.2.34}$$

By multiplying in, we get canceling cross-terms. For the potential $V(\phi^\dagger \phi)$, we eliminate terms with degrees higher than 2 in h. So, the Lagrangian becomes:

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu h\right)\left(\partial^\mu h\right) + \frac{1}{2}\left(\partial_\mu \mathcal{E}\right)\left(\partial^\mu \mathcal{E}\right) - h\left(\lambda \mathcal{V}^3 + m^2 \mathcal{V}\right)$$

$$- h^2 \left(\frac{1}{2}m^2 + \frac{3}{2}\lambda \mathcal{V}^2\right) + \dots$$

$$= \frac{1}{2}\left(\partial_\mu h\right)\left(\partial^\mu h\right) + \frac{1}{2}\left(\partial_\mu \mathcal{E}\right)\left(\partial^\mu \mathcal{E}\right) - h\mathcal{V}\left(\lambda \frac{-m^2}{\lambda} + m^2\right)$$

$$- \frac{1}{2}h^2 \left(m^2 + \lambda \frac{-3m^2}{\lambda}\right)$$

$$= \frac{1}{2}\left(\partial_\mu h\right)\left(\partial^\mu h\right) + \frac{1}{2}\left(\partial_\mu \mathcal{E}\right)\left(\partial^\mu \mathcal{E}\right) - \frac{1}{2}h^2(-2m^2). \tag{11.2.35}$$

So,

$$\mathcal{L} = \boxed{\frac{1}{2}\left(\partial_\mu h\right)\left(\partial^\mu h\right) - \frac{1}{2}h^2(-2m^2)} + \boxed{\frac{1}{2}\left(\partial_\mu \mathcal{E}\right)\left(\partial^\mu \mathcal{E}\right)} \tag{11.2.36}$$

5. The first box

$$\frac{1}{2}\left(\partial_\mu h\right)\left(\partial^\mu h\right) - \frac{1}{2}h^2(-2m^2) \tag{11.2.37}$$

describes a massive, spin 0, real scalar field - **the Higgs field**.

The second box

$$\frac{1}{2}\left(\partial_\mu \mathcal{E}\right)\left(\partial^\mu \mathcal{E}\right) \tag{11.2.38}$$

describes the massless **Nambu-Goldstone field**.

6. With the current Lagrangian:

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu h\right)\left(\partial^\mu h\right) + \frac{1}{2}\left(\partial_\mu \mathcal{E}\right)\left(\partial^\mu \mathcal{E}\right) - \frac{1}{2}h^2(-2m^2), \tag{11.2.39}$$

varying the Lagrangian with respect to h and $\mathcal{E}$ (which we have done many times before) gives

$$\delta\mathcal{L} = \frac{1}{2}\delta\left[(\partial^\mu h)(\partial_\mu h)\right] - \frac{1}{2}(-2m^2)\delta h^2 + \frac{1}{2}\delta\left[(\partial^\mu\mathcal{E})(\partial_\mu\mathcal{E})\right]$$
$$= (\partial_\mu\delta h)(\partial^\mu h) + 2m^2 h\delta h + (\partial_\mu\delta\mathcal{E})(\partial^\mu\mathcal{E}). \qquad (11.2.40)$$

Integrating by parts we get,

$$(\partial^\mu h)(\partial_\mu h) = \partial^\mu\left((\partial_\mu h)\delta h\right) - (\partial_\mu\partial^\mu h)\,\delta h \qquad (11.2.41)$$
$$(\partial^\mu\mathcal{E})(\partial_\mu\mathcal{E}) = \partial^\mu\left((\partial_\mu\mathcal{E})\delta\mathcal{E}\right) - (\partial_\mu\partial^\mu\mathcal{E})\,\delta\mathcal{E}. \qquad (11.2.42)$$

The boundary terms go to zero by the variation constraints. By requiring that varying the action gives zero, $\delta\mathcal{L} = 0$ as well. So we are left with

$$\delta\mathcal{L} = \left[(\partial_\mu\partial^\mu h + 2m^2)h\right]\delta h + \partial_\mu\partial^\mu\mathcal{E} = 0, \qquad (11.2.43)$$

which holds if and only if

$$\begin{cases} (\Box + 2m^2)h = 0 \\ \Box\mathcal{E} = 0, \end{cases} \qquad (11.2.44)$$

where $\Box \equiv \partial_\mu\partial^\mu$ is the d'Alembertian. So, the equations of motion we obtained from varying the Lagrangian are the Klein-Gordon equations. The first one,

$$\boxed{\left(\Box + 2m^2\right)h = 0} \qquad (11.2.45)$$

associated with the Higgs field, is consistent with the equation of motion for the massive scalar field.

The second equation,

$$\boxed{\Box\mathcal{E} = 0} \qquad (11.2.46)$$

associated with the Nambu-Goldstone mode, is consistent with the equation of motion for the massless scalar field.

Problem 11.2.2. Local U(1) symmetry

Consider the Lagrangian:

$$\mathcal{L} = \frac{1}{2}(D_\mu\phi)^\dagger(D^\mu\phi) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - V(\phi^\dagger\phi) \qquad (11.2.47)$$

where

$$D_\mu\phi = (\partial_\mu + iqA_\mu)\phi \qquad (11.2.48)$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \qquad (11.2.49)$$

$$V(\phi^\dagger\phi) = \frac{1}{2}m^2\phi^\dagger\phi + \frac{1}{4}\lambda(\phi^\dagger\phi)^2. \qquad (11.2.50)$$

1. Show that the Lagrangian is invariant under local U(1) transformations:

$$\phi \to \phi' = e^{i\alpha(x)}\phi \qquad (11.2.51)$$

$$A_\mu \to A'_\mu = A_\mu - \frac{1}{q}\partial_\mu\alpha(x) \qquad (11.2.52)$$

2. Again, suppose that $m^2 < 0$. Introduce polar variables (ρ, θ) again. Show that the vacuum solutions are degenerate and find the equation in ρ, θ describing the space of possible vacua.

3. Spontaneously pick the vacuum:

$$\langle\rho\rangle = \mathcal{V} \qquad (11.2.53)$$

$$\langle\theta\rangle = 0 \qquad (11.2.54)$$

$$\langle A_\mu\rangle = 0 \qquad (11.2.55)$$

and consider small excitations about the vacuum, $\rho \approx \mathcal{V} + h$ and $\theta \approx \mathcal{E}/\mathcal{V}$:

$$\phi \approx (\mathcal{V} + h)e^{i\mathcal{E}/\mathcal{V}} \qquad (11.2.56)$$

$$A_\mu = 0 + A_\mu. \qquad (11.2.57)$$

Which excitation (h or $\mathcal{E}$) stays in the degenerate space of possible vacua (minimum of V) and which involves an excitation of V away from the minimum?

4. Parameterize ϕ as:

$$\phi = e^{-i\alpha(x)}\phi' \qquad (11.2.58)$$

with $\phi' = \rho = \mathcal{V} + h$ by picking $\alpha(x) = -\mathcal{E}/\mathcal{V}$, called the "unitary gauge." Then perform the gauge transformation with this $\alpha(x)$:

$$\phi \to e^{i\alpha}\phi = \phi' = \mathcal{V} + h \qquad (11.2.59)$$

$$A_\mu \to A'_\mu \qquad (11.2.60)$$

Write out the Lagrangian in terms of $h, \mathcal{V}, A'_\mu, \mathcal{E}, \alpha, \ldots$

5. Keep up to quadratic terms in the Lagrangian in the small quantities $h, \mathcal{E}, A'_\mu$. Describe the types of fields in the theory (massive versus massless, scalar, vectors, etc.). Did any field disappear?

6. Vary the quadratic form of the Lagrangian with respect to the fields it contains. What are the field equations of motion?

7. When we started with

$$\mathcal{L} = \frac{1}{2}(D_\mu \phi)^\dagger (D^\mu \phi) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - V(\phi^\dagger \phi), \qquad (11.2.61)$$

how many degrees of freedom were there? What are they?

When we ended with $\mathcal{L}(h, \mathcal{E}, A'_\mu)$, how many degrees of freedom are there? What types of modes are they?

Solution 11.2.2.

1. Under a local U(1) transformation, the new Lagrangian is

$$\mathcal{L}' = \frac{1}{2}(D'_\mu \phi')^\dagger (D'^\mu \phi') - \frac{1}{4}F'_{\mu\nu}F'^{\mu\nu} - V(\phi'^\dagger \phi'). \qquad (11.2.62)$$

where

$$\phi' = \phi e^{i\alpha(x)} \qquad (11.2.63)$$

and

$$\begin{aligned} D'_\mu &= \partial_\mu + iqA'_\mu \\ &= \partial_\mu + iq\left(A_\mu - \frac{1}{q}\partial_\mu \alpha(x)\right) \\ &= \partial_\mu + iqA_\mu - i\partial_\mu \alpha(x) \end{aligned} \qquad (11.2.64)$$

which give

$$\phi'^\dagger \phi' = \phi^\dagger \phi \qquad (11.2.65)$$

so the potential is invariant

$$V(\phi'^\dagger \phi') = V(\phi^\dagger \phi). \qquad (11.2.66)$$

We also have

$$\begin{aligned} D'_\mu \phi' &= (\partial_\mu + iqA_\mu - i\partial_\mu \alpha(x))\phi e^{i\alpha(x)} \\ &= e^{i\alpha(x)}\partial_\mu \phi + i\phi e^{i\alpha(x)}\partial_\mu \alpha(x) + iqA_\mu \phi e^{i\alpha(x)} - i\phi e^{i\alpha(x)}\partial_\mu \alpha(x) \\ &= (D_\mu \phi)e^{i\alpha(x)}, \end{aligned} \qquad (11.2.67)$$

which we could have guessed since D_μ and A_μ are designed to accomplish just this.

We also know that the electromagnetic field strength tensor is also invariant, since the "new" electromagnetic field strength tensor is just the curl of A_μ plus a gradient:

$$
\begin{aligned}
F'_{\mu\nu}F'^{\mu\nu} &= (\partial_\mu A'_\nu - \partial_\nu A'_\mu)(\partial^\mu A'^\nu - \partial^\nu A'^\mu) \\
&= \left(F_{\mu\nu} - \frac{1}{q}\partial_\mu\partial_\nu\alpha(x) + \frac{1}{q}\partial_\nu\partial_\mu\alpha(x) \right) \\
&\quad \times \left(F^{\mu\nu} - \frac{1}{q}\partial^\mu\partial^\nu\alpha(x) + \frac{1}{q}\partial^\nu\partial^\mu\alpha(x) \right) \\
&= F_{\mu\nu}F^{\mu\nu}.
\end{aligned}
\tag{11.2.68}
$$

So, putting it all together...

$$
\begin{aligned}
\mathcal{L}' &= \frac{1}{2}(D'_\mu\phi')^\dagger(D'^\mu\phi') - \frac{1}{4}F'_{\mu\nu}F'^{\mu\nu} - V(\phi'^\dagger\phi') \\
&= \frac{1}{2}(D_\mu\phi)^\dagger(D^\mu\phi) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - V(\phi^\dagger\phi) \\
&= \mathcal{L}.
\end{aligned}
\tag{11.2.69}
$$

So the Lagrangian is invariant under local U(1) transformations.

2. Introduce polar variables again:

$$
\phi = \rho e^{i\theta}.
\tag{11.2.70}
$$

The potential V then has the form:

$$
V(\phi^\dagger\phi) = V(\rho e^{i\theta} \cdot \rho e^{-\theta}) = V(\rho^2),
\tag{11.2.71}
$$

i.e.,

$$
V(\rho^2) = \frac{1}{2}m^2\rho^2 + \frac{1}{4}\lambda\rho^4.
\tag{11.2.72}
$$

Differentiating with respect to ρ^2, we obtained the minimum value of V (assuming $-m^2 > 0$) at

$$
\rho^2 = \frac{-m^2}{\lambda}.
\tag{11.2.73}
$$

As before, we have a degeneracy at this value of ρ^2, since θ can take any value 0 to 2π. This equation describing the space of possible vacua is just:

$$
\rho e^{i\theta} = \sqrt{\frac{-m^2}{\lambda}}e^{i\theta}.
\tag{11.2.74}
$$

3. We spontaneously pick a vacuum as given:

$$\langle \rho \rangle = \mathcal{V} \tag{11.2.75}$$

$$\langle \theta \rangle = 0 \tag{11.2.76}$$

$$\langle A_\mu \rangle = 0. \tag{11.2.77}$$

Now, consider small excitations about the vacuum:

$$\phi = (\mathcal{V} + h)e^{i\mathcal{E}/\mathcal{V}} \tag{11.2.78}$$

$$A_\mu = 0 + A_\mu. \tag{11.2.79}$$

Since we have said earlier that the equation describing the space of possible vacua is θ-independent, excitations in $\mathcal{E}$ (or $\mathcal{E}/\mathcal{V}$) does not leave the minimum of $\mathcal{V}$.

On the other hand, excitations in ρ leaves the minimum of V. If we let

$$\mathcal{V} = \rho = \sqrt{\frac{-m^2}{\lambda}}, \tag{11.2.80}$$

then obviously $(\mathcal{V} + h)^2 > \mathcal{V}^2$, or $\rho' > \rho$.

4. We re-parameterize ϕ as the inverse transformation

$$\phi = e^{-i\alpha(x)}\phi' \tag{11.2.81}$$

where $\phi = \rho = \mathcal{V} + h$ and $\alpha(x) = -\mathcal{E}/\mathcal{V}$, then

$$\phi' = e^{i\alpha(x)}\phi = \mathcal{V} + h \tag{11.2.82}$$

$$A'_\mu = A_\mu - \frac{1}{q}\partial_\mu\alpha(x) = A_\mu - \frac{1}{q}\partial_\mu\frac{\mathcal{E}}{\mathcal{V}} = A_\mu - \frac{1}{q\mathcal{V}}\partial_\mu\mathcal{E}. \tag{11.2.83}$$

The new Lagrangian is

$$\mathcal{L} = \frac{1}{2}(D'_\mu\phi')^\dagger(D'^\mu\phi') - \frac{1}{4}F'_{\mu\nu}F'^{\mu\nu} - V(\phi'^\dagger\phi'). \tag{11.2.84}$$

The potential changes as:

$$
\begin{aligned}
V(\phi'^\dagger\phi') &= \frac{1}{2}m^2(\mathcal{V} + h)^2 + \frac{1}{4}\lambda(\mathcal{V} + h)^4 \\
&= \frac{1}{2}m^2(\mathcal{V}^2 + 2\mathcal{V}h + h^2) + \frac{1}{4}\lambda(\mathcal{V}^4 + 4\mathcal{V}^3h + 6\mathcal{V}^2h^2 + 4\mathcal{V}h^3 + h^4) \\
&\approx h\mathcal{V}\left(m^2 + \lambda\mathcal{V}^2\right) + h^2\left(\frac{1}{2}m^2 + \frac{3}{2}\lambda\mathcal{V}^2\right) \\
&= \frac{1}{2}(-2m^2)h^2 \tag{11.2.85}
\end{aligned}
$$

where the second equality comes from the fact that

$$\mathcal{V} = \frac{-m^2}{\lambda}. \tag{11.2.86}$$

Finally, we look at the new covariant derivative in the Lagrangian:

$$\begin{aligned}
D'_\mu \phi' &= (\partial_\mu + iqA'_\mu)(\mathcal{V} + h) \\
&= (\partial_\mu + iqA'_\mu)(\mathcal{V} + h) \\
&= \partial_\mu h + iqA'_\mu(\mathcal{V} + h).
\end{aligned} \tag{11.2.87}$$

So,

$$\begin{aligned}
(D'_\mu \phi')^\dagger (D'^\mu \phi') &= \left(\partial_\mu h + iqA'_\mu(\mathcal{V} + h) \right) \left(\partial^\mu h - iqA'^\mu(\mathcal{V} + h) \right) \\
&= (\partial_\mu h)(\partial^\mu h) + q^2 A'_\mu A'^\mu (\mathcal{V} + h)^2 \\
&\quad - iq(\mathcal{V} + h)A'^\mu \partial_\mu h + iq(\mathcal{V} + h)A'_\mu \partial^\mu h \\
&= (\partial_\mu h)(\partial^\mu h) + q^2 A'_\mu A'^\mu (\mathcal{V} + h)^2 \\
&\quad - iq(\mathcal{V} + h)A'^\mu \partial_\mu h + iq(\mathcal{V} + h)A'^\mu \partial_\mu h \\
&= (\partial_\mu h)(\partial^\mu h) + q^2 A'_\mu A'^\mu (\mathcal{V} + h)^2.
\end{aligned} \tag{11.2.88}$$

Putting everything together...

$$\mathcal{L}' = \frac{1}{2}(\partial_\mu h)(\partial^\mu h) + \frac{1}{2}q^2 A'_\mu A'^\mu (\mathcal{V} + h)^2 - \frac{1}{4}F'_{\mu\nu}F'^{\mu\nu} - \frac{1}{2}(-2m^2)h^2. \tag{11.2.89}$$

Expanding this out gives the Lagrangian:

$$\begin{aligned}
\mathcal{L}' &= \frac{1}{2}(\partial_\mu h)(\partial^\mu h) - \frac{1}{2}(-2m^2)h^2 + \frac{q^2 \mathcal{V}^2}{2}A'_\mu A'^\mu \\
&\quad - \frac{1}{4}F'_{\mu\nu}F'^{\mu\nu} + \frac{1}{2}q^2(2\mathcal{V}h + h^2)A'_\mu A'^\mu.
\end{aligned} \tag{11.2.90}$$

5. The first two terms:

$$\boxed{\frac{1}{2}(\partial_\mu h)(\partial^\mu h) - \frac{1}{2}(-2m^2)h^2} \tag{11.2.91}$$

describe a massive $(-m^2 > 0)$, spin 0 scalar field. They describe the **Higgs boson**.

The next two terms:

$$\boxed{\frac{q^2 \mathcal{V}^2}{2}A'_\mu A'^\mu - \frac{1}{4}F'_{\mu\nu}F'^{\mu\nu}} \tag{11.2.92}$$

describe a **massive vector field** (which we have found in the first "assignment")

The last (significant) term:

$$\boxed{\frac{1}{2}q^2(2\mathcal{V}h + h^2)A'_\mu A^{'\mu}}$$ (11.2.93)

is the **interaction** between h and A'_μ.

We notice that the field associated with the $\mathcal{E}$ excitation has disappeared. This makes sense, since excitations in θ does not leave the space of possible vacua. We also observe that the massless Nambu-Goldstone field is no longer in the Lagrangian.

6. Recall the Lagrangian:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu h)(\partial^\mu h) - \frac{1}{2}(-2m^2)h^2 + \frac{q^2\mathcal{V}^2}{2}A'_\mu A^{'\mu}$$
$$- \frac{1}{4}F'_{\mu\nu}F^{'\mu\nu} + \frac{1}{2}q^2(2\mathcal{V}h + h^2)A'_\mu A^{'\mu}.$$ (11.2.94)

Assuming that $h \ll \mathcal{V}$ (negligible interaction), we can rewrite this as

$$\mathcal{L} = \frac{1}{2}(\partial_\mu h)(\partial^\mu h) - \frac{1}{2}(-2m^2)h^2 + \frac{q^2\mathcal{V}^2}{2}A'_\mu A^{'\mu} - \frac{1}{4}F'_{\mu\nu}F^{'\mu\nu}.$$ (11.2.95)

Call

$$\mathcal{L}_1 = \frac{1}{2}(\partial_\mu h)(\partial^\mu h) - \frac{1}{2}(-2m^2)h^2$$ (11.2.96)

and

$$\mathcal{L}_2 = \frac{q^2\mathcal{V}^2}{2}A'_\mu A^{'\mu} - \frac{1}{4}F'_{\mu\nu}F^{'\mu\nu}.$$ (11.2.97)

We have

$$\delta\mathcal{L} = \delta\mathcal{L}_1 + \delta\mathcal{L}_2.$$ (11.2.98)

Since we have eliminated the interaction term, $\mathcal{L}_1$ is only dependent on h, while $\mathcal{L}_2$ is only dependent on A_μ. We shall vary $\mathcal{L}_1$, $\mathcal{L}_2$ with respect to the corresponding fields they contain.

Varying $\mathcal{L}_1$ with respect to h is the same as varying the first box in (11.2.36), which simply gives the Klein-Gordon equation of motion for a massive spin 0 scalar field:

$$\boxed{(\Box + 2m^2)h = 0}$$ (11.2.99)

To vary the second box in (11.2.36) with respect to the vector potential A_μ, we have to make use of the results we got earlier in the "massive vector fields" assignment:

$$\delta \left(-\frac{1}{4} F^{\mu\nu} F_{\mu\nu} \right) = \partial_\mu(\partial^\mu A^\nu) \delta A_\nu - (\partial^\nu \partial_\mu A_\nu) \delta A^\mu$$

$$= (\Box A^\mu - \partial^\mu \partial^\nu A_\nu) \, \delta A_\mu$$

$$= (\Box A_\mu - \partial_\mu \partial^\nu A_\nu) \, \delta A^\mu \qquad (11.2.100)$$

and

$$\delta \left(\frac{1}{2} q^2 \mathcal{V}^2 A_\mu A^\mu \right) = \frac{q^2 \mathcal{V}^2}{2} \left(\delta A_\mu A^\mu + A_\mu \delta A^\mu \right)$$

$$= \frac{q^2 \mathcal{V}^2}{2} \left(\delta A_\mu A^\mu + A^\mu \delta A_\mu \right) \qquad (11.2.101)$$

$$= \left(q^2 \mathcal{V}^2 A^\mu \right) \delta A_\mu$$

$$= \left(q^2 \mathcal{V}^2 A_\mu \right) \delta A^\mu. \qquad (11.2.102)$$

Putting everything together,

$$\delta \mathcal{L} = \left(\Box A_\mu - \partial_\mu \partial^\nu A_\nu + q^2 \mathcal{V}^2 A_\mu \right) \delta A^\mu. \qquad (11.2.103)$$

Varying the action and requiring $\delta S = 0$ for all δA^μ forces $\delta \mathcal{L}_2 = 0$. We arrive at the equations of motion

$$\Box A_\mu - \partial_\mu \partial^\nu A_\nu + q^2 \mathcal{V}^2 A_\mu = 0. \qquad (11.2.104)$$

The given Lagrangian has no x^μ dependence, so we expect $\partial^\mu \mathcal{L}_2 = 0$. We will show that this adds a constraint to A^μ.

$$\partial^\mu \mathcal{L}_2 = \partial^\mu \left(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} q^2 \mathcal{V}^2 A_\mu A^\mu \right)$$

$$= -\frac{1}{4} \partial^\mu \left(F_{\mu\nu} F^{\mu\nu} \right) + \frac{1}{2} q^2 \mathcal{V}^2 \left(\partial^\mu A_\mu A^\mu + A_\mu \partial^\mu A^\mu \right)$$

$$= -\frac{1}{2} \partial^\mu \left(\partial^\mu A^\nu \, \partial_\mu A_\nu - \partial_\mu A_\nu \, \partial^\mu A^\nu \right) + q^2 \mathcal{V}^2 A^\mu \partial^\mu A_\mu$$

$$= -\frac{1}{2} \left[(\partial^\mu \partial^\mu A^\nu)(\partial_\mu A_\nu) - (\partial^\mu \partial^\mu A^\nu)(\partial_\mu A_\nu) \right.$$

$$\left. + (\Box A_\nu)(\partial^\mu A^\nu) - (\Box A_\nu)(\partial^\mu A^\nu) \right] + q^2 \mathcal{V}^2 A^\mu \partial^\mu A_\mu$$

$$= q^2 \mathcal{V}^2 A^\mu \partial^\mu A_\mu. \qquad (11.2.105)$$

Since $\partial^\mu \mathcal{L}_2 = 0$, a massive $m \neq 0$ field is assumed, and A^μ is arbitrary, $\partial^\mu A_\mu = 0$ must hold as desired. With this, the equations of motion become:

$$\boxed{\left(\Box + q^2 \mathcal{V}^2 \right) A_\mu = 0} \qquad (11.2.106)$$

7. When we start with

$$\mathcal{L} = \frac{1}{2}(D_\mu\phi)^\dagger(D^\mu\phi) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - V(\phi^\dagger\phi), \qquad (11.2.107)$$

there are 4 degrees of freedom: 2 for the complex scalar field ϕ and 2 for the massless A_μ gauge field.

Now, when we end with

$$\mathcal{L} = \frac{1}{2}(\partial_\mu h)(\partial^\mu h) - \frac{1}{2}(-2m^2)h^2 + \frac{q^2\mathcal{V}^2}{2}A'_\mu A^{'\mu}$$
$$- \frac{1}{4}F'_{\mu\nu}F^{'\mu\nu} + \frac{1}{2}q^2(2\mathcal{V}h + h^2)A'_\mu A^{'\mu}, \qquad (11.2.108)$$

there are still 4 degrees of freedom, but 1 belongs to the massive, spin 0 Higgs (scalar) field, while 3 belong to the massive vector gauge field A_μ.

11.3 Euler-Lagrange Equations (A. Lahiri + P. B. Pal)

Problem 11.3.1. The Lagrangian (density) of an electromagnetic field interacting with a current j^μ is given by

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - j_\mu A^\mu. \tag{11.3.1}$$

Treating the A^μ as the fields, find the Euler-Lagrange equations and show that they give the inhomogeneous Maxwell's equations.

Solution 11.3.1. Solution

Problem 11.3.2. For a real scalar field ϕ, the Lagrangian $\mathcal{L}$ is

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2}m^2\phi^2 - V(\phi). \qquad (11.3.2)$$

Find the Euler-Lagrange equations.

Solution 11.3.2. Solution

Problem 11.3.3. Do the same for a complex scalar field ϕ with Lagrangian

$$\mathcal{L} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - m^2 \phi^\dagger \phi - V(\phi^\dagger \phi). \qquad (11.3.3)$$

Solution 11.3.3.

11.4   Hamiltonian Formalism (A. Lahiri + P. B. Pal)

Problem 11.4.1. Treating the A^μ's as the fields, find the canonical momenta associated with them, using the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - j_\mu A^\mu. \tag{11.4.1}$$

Can these equations, expressing the momenta in terms of A^μ's and their derivatives, be inverted to solve the $\dot{A}^\mu$'s? If not, can you think of a reason why?

Solution 11.4.1. Solution

Problem 11.4.2. There is some freedom in defining the electromagnetic potentials A^μ. Using this freedom, we can make $A^0 = 0$. This is an example of *gauge fixing*. In this gauge, treating the components of $\mathbf{A}$ as the independent fields, find out the momenta. Can these equations be inverted to solve the $\dot{A}^i$'s, where the index i runs over the spatial components only?

Solution 11.4.2. Solution

Problem 11.4.3. Find the Hamiltonian for the complex scalar field whose Lagrangian is given by

$$\mathcal{L} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - m^2 \phi^\dagger \phi - V(\phi^\dagger \phi). \qquad (11.4.2)$$

Solution 11.4.3. Solution.

Problem 11.4.4. Consider a radiation field in a one-dimensional box or equivalently displacements on a string of length l with fixed ends. (This was the very first attempt at field theory.)

$$L = \int_0^l dx \left[\left(\frac{\partial u}{\partial t} \right)^2 - c^2 \left(\frac{\partial u}{\partial x} \right)^2 \right]. \tag{11.4.3}$$

Let us express the field u in a Fourier series:

$$u(x,t) = \sum_{k=1}^{\infty} q_k(t) \sin \frac{\omega_k x}{c}, \qquad \omega_k = \frac{k\pi c}{l}. \tag{11.4.4}$$

1. Calculate the 'momentum' p_k, canonically conjugate to q_k.

2. Write down the Hamiltonian as a function of q's and p's.

3. Take q_k and p_k to be operators with $[q_k, p_j] = i\hbar \delta_{kj}$. Define operators $a_k, a_k^\dagger$ by

$$q_k = \sqrt{\frac{\hbar}{2l\omega_k}} \left[a_k e^{-\omega_k t} + a_k^\dagger e^{i\omega_k t} \right]. \tag{11.4.5}$$

Find the commutators of a_k and $a_k^\dagger$ with themselves and one another.

4. Calculate the Hamiltonian in terms of a's and $a^\dagger$'s.
Solution 11.4.4. Solution.

11.5   Noether's Theorem (A. Lahiri + P. B. Pal)

11.5.1   Space-time translations

Problem 11.5.1. Find the stress-energy tensor for the following fields:

1. Real scalar.

2. Complex scalar.

3. Electromagnetic field with $j^\mu = 0$.

Solution 11.5.1. Solution.

11.6 Internal symmetries

Problem 11.6.1. Consider the complex scalar field. The Lagrangian remains invariant under $\phi \to e^{-iq\theta}\phi$, $\phi^\dagger \to e^{iq\theta}\phi^\dagger$.

1. Use Noether's theorem to find the corresponding conserved current j^μ. Verify that $p_\mu j^\mu = 0$.

2. Calculate the conserved Noether current for the same transformation when the electromagnetic field is also included, and the combined Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \left[(\partial^\mu - iqA^\mu)\phi^\dagger\right]\left[(\partial_\mu + iqA_\mu)\phi\right] - m^2\phi^\dagger\phi - V(\phi^\dagger\phi).$$

$$(11.6.1)$$

3. The Euler-Lagrange equations for A_μ should give the Maxwell equations

$$\partial_\mu F^{\mu\nu} = j^\nu. \tag{11.6.2}$$

Show that the 'current' appearing in this equation is the same as the current obtained in part 2. above. [Note: since the conserved current involves A^μ, the Lagrangian cannot be written like

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - j_\mu A^\mu. \tag{11.6.3}$$

for this case.]

Solution 11.6.1. Solution.

Problem 11.6.2. Consider the Lagrangian of a real scalar with $m^2 = 0$ and $V(\phi) = \lambda\phi^4$. Verify that the action has the symmetry

$$x \to bx, \qquad \phi \to \phi/b. \tag{11.6.4}$$

Find the conserved current corresponding to this symmetry.

Solution 11.6.2. Solution.

Problem 11.6.3. Suppose the action of a certain field theory is invariant under space-time translation as well as dilations

$$x \to bx, \qquad \Phi \to \Phi. \qquad (11.6.5)$$

Show that the stress-energy tensor in this case is traceless, i.e.,

$$T^{\mu}{}_{\mu} = 0. \qquad (11.6.6)$$

Solution 11.6.3. Solution.

Problem 11.6.4. Show that the Lie derivative written completely in covariant derivatives is equivalent to the version with only partial derivatives. In particular:

1. Show it holds for scalars

$$\mathcal{L}_\xi \phi = \xi^\alpha \partial_\alpha \phi = \xi^\alpha D_\alpha \phi. \tag{11.6.7}$$

2. Show it holds for the contravariant case

$$\begin{aligned}
\mathcal{L}_\xi A^\mu &= -(\partial_\alpha \xi^\mu) A^\alpha + \xi^\alpha (\partial_\mu A^\mu) \\
&= -(D_\alpha \xi^\mu) A^\alpha + \xi^\alpha (D_\mu A^\mu).
\end{aligned} \tag{11.6.8}$$

3. Show it holds for the covariant case

$$\begin{aligned}
\mathcal{L}_\xi A_\mu &= (\partial_\mu \xi^\alpha) A_\alpha + \xi^\alpha (\partial_\alpha A_\mu) \\
&= (D_\mu \xi^\alpha) A_\alpha + \xi^\alpha (D_\alpha A_\mu).
\end{aligned} \tag{11.6.9}$$

Solution 11.6.4.

1. For scalars:

2. For a contravariant vector:

3. For a covariant vector:

11.7 Diffeomorphism and Lie Derivatives

Problem 11.7.1. Show that

$$\mathcal{L} = \frac{1}{2}(D_\mu\phi)(D^\mu\phi) - \frac{1}{2}m^2\phi^2 \qquad (11.7.1)$$

is a scalar under diffeomorphisms: $\mathcal{L} \to \mathcal{L} + \xi^\alpha D_\alpha \mathcal{L}$.

Solution 11.7.1. We know that ϕ^2 is a scalar:

$$D_\mu\phi \to D_\mu\phi + (D_\mu\xi^\alpha)D_\alpha\phi + \xi^\alpha D_\alpha D_\mu\phi \qquad (11.7.2)$$

$$D^\mu\phi \to \dots \qquad (11.7.3)$$

Verify that

$$(D_\mu\phi)(D^\mu\phi) = (D_\mu\phi)(D^\mu\phi) + \xi^\alpha D_\alpha(D_\mu\phi)(D^\mu\phi). \qquad (11.7.4)$$