CLASSICAL FIELD THEORY

A Quick Guide

Huan Q. Bui

Colby College
Physics & Statistics
Class of 2021

February 14, 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. Some parts of the text will be derived from *Quantum Field Theory* by Ryder.

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.

Enjoy!

Contents

1 Introduction to Classical Field Theory

2 Group Theory: a quick study

3 Overview of Lagrangian Formulation of Classical Mechanics

4 Lagrangian Formulation in Field Theory

Proposition 4.1. All fundamental physics obeys least action principles.

The action S is defined as

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

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.

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

4.1 Field Theory: A 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:

$$\mathcal{L} = K - V, \tag{4.1}$$

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

$$\begin{aligned} K &= \frac{1}{2} m \dot{x}^2 \\ V &= V(x). \end{aligned}$$

Variations on the Lagrangian gives

$$\begin{aligned} \delta \mathcal{L} &= \delta \left(\frac{1}{2} m \dot{x}^2 - V(x) \right) \\ &= m \dot{x} \delta \dot{x} - \frac{dV}{dx} \delta x \\ &= m \dot{x} \dot{\delta x} - \frac{dV}{dx} \delta x \\ &= m \left(-\ddot{x} \delta x + \frac{d}{dt} \dot{x} \delta x \right) - \frac{dV}{dx} \delta x \\ &= -m \ddot{x} \delta x - m \frac{d}{dt} \dot{x} \delta x - \frac{dV}{dx} \delta x. \end{aligned}$$

It follows that the variations on the action gives

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

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

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

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

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

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

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

where d^4x implies integrating over all spacetime.

4.2 Introduction to Fields

In field theory, most physical objects are described as "fields." Let us dive into the first two fields that we are more or less familiar with: scalar fields and vector fields.

4.3 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

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

where

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

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):

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

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

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

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

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

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

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

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

$$\begin{aligned} \left(K^0\right)^2 - \vec{k}^2 &= m^2 \\ K^\mu K_\mu &= m^2. \end{aligned}$$

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

$$\begin{aligned} \text{Kinetic energy: } K &= \frac{1}{2}\dot{\phi}^2 \\ \text{Gradient energy: } G &= \frac{1}{2}\left(\nabla\phi\right)^2 \\ \text{Potential energy: } P &= V(\phi). \end{aligned}$$

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:

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

We shall verify that

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

This turns out to be quite straightforward:

$$\begin{aligned}\left(\partial_\mu \phi\right)\left(\partial^\mu \phi\right) &= \eta^{\mu\nu}\left(\partial_\mu \phi\right)\left(\partial_\nu \phi\right)\\ &= \left(\partial_0 \phi\right)^2 - \left(\partial_j \phi\right)^2\\ &= \dot{\phi}^2 - \left(\nabla\phi\right)^2.\end{aligned}$$

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

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

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

$$\begin{aligned}\delta\mathcal{L} &= \delta\left(-\frac{1}{2}\left(\partial_\mu \phi\right)\left(\partial^\mu \phi\right) - V(\phi)\right)\\ &= -\frac{1}{2}\left(\partial_\mu \phi\,\partial^\mu \phi + \partial_\mu \phi\,\partial^\mu \delta\phi\right) - \frac{dV(\phi)}{d\phi}\delta\phi\\ &= -\partial_\mu \delta\phi\,\partial^\mu \phi - \frac{dV(\phi)}{d\phi}\delta\phi.\end{aligned}$$

Now, integration by parts tells us that

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

So,

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

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

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

It follows that the action is

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

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

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

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

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

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

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

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

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

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:

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

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

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

is a solution to the KG equation. We simply unpack the d'Alembertian and attack the derivatives step-by-step. The first derivative is

$$\begin{aligned}
\partial_\mu \phi &= \partial_\mu \left(e^{-i\mathbf{k}\cdot\mathbf{x}}\right) \\
&= -i\partial_\mu \left(\mathbf{k}\cdot\mathbf{x}\right) e^{-i\mathbf{k}\cdot\mathbf{x}} \\
&= -i\partial_\mu \left(K_\nu X^\nu\right) e^{-iK_\alpha X^\alpha} \\
&= -iK_\nu\, \partial_\mu X^\nu\, \phi \\
&= -iK_\nu \delta^\nu_\mu \phi \\
&= -iK_\mu \phi
\end{aligned}$$

Next, we attack the second derivative:

$$\begin{aligned}
\partial^\mu \partial_\mu \phi &= \eta^{\mu\nu} \partial_\nu \partial_\mu \phi \\
&= \eta^{\mu\nu} \partial_\nu \left(-iK_\mu \phi\right) \\
&= -iK_\mu \eta^{\mu\nu} \left(-iK_\nu \phi\right) \\
&= (-i)^2 K^\mu K_\mu \phi.
\end{aligned}$$

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

$$\Box\phi + m^2\phi = (-m^2 + m^2)\phi = 0,$$

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

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

Recall the Lagrangian:

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

we get

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\phi} &= -\frac{dV}{d\phi}\\ \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} &= \partial^\mu\phi.\end{aligned}$$

So the Euler-Lagrange equation gives:

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

If we reasonably take

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

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

$$(\Box + m^2)\phi = 0.$$

4.4 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:

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

Next, recall our earlier definition of the variation:

$$\phi'(x) = \phi(x) + \delta\phi(x).$$

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

$$\begin{aligned}
\Delta\phi &= \phi'(x') - \phi(x) \\
&= [\phi'(x') - \phi(x')] + [\phi(x') - \phi(x)] \\
&\approx \delta\phi + (\partial_\mu \phi)\delta x^\mu.
\end{aligned}$$

So, the variation is now

$$\begin{aligned}
\delta S &= \int \mathcal{L}(\phi', \partial_\mu \phi', x'^\mu)\, d^4x' - \int \mathcal{L}(\phi, \partial_\mu \phi, x^\mu)\, d^4x \\
&= \int \mathcal{L}(\phi', \partial_\mu \phi', x'^\mu) J\left(\frac{x'}{x}\right)\, d^4x - \int \mathcal{L}(\phi, \partial_\mu \phi, x^\mu)\, d^4x,
\end{aligned}$$

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

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

So, the variation becomes

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

where

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

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

$$\begin{aligned}
\delta S &= \int \left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\mu(\delta\phi) + \frac{\partial\mathcal{L}}{\partial x^\mu}\delta x^\mu + \mathcal{L}\partial_\mu(\delta x^\mu)\right] d^4x \\
&= \int \left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\mu(\delta\phi) + \left(\frac{\partial\mathcal{L}}{\partial x^\mu}\delta x^\mu + \mathcal{L}\partial_\mu(\delta x^\mu)\right)\right] d^4x \\
&= \int \left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\mu(\delta\phi) + \partial_\mu(\mathcal{L}\delta x^\mu)\right] d^4x.
\end{aligned}$$

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

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

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

$$\begin{aligned}\delta S &= \int_R \left[\frac{\partial \mathcal{L}}{\partial \phi}\delta\phi + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\delta\phi\right) - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right)\delta\phi + \partial_\mu(\mathcal{L}\delta x^\mu)\right] d^4x \\ &= \int_R \left\{\delta\phi \left[\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right)\right] + \partial_\mu \left[\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\delta\phi\right) + (\mathcal{L}\delta x^\mu)\right]\right\} d^4x.\end{aligned}$$

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

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

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:

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

(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:

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

Recall that the total variation is defined as

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

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

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

So, once again, the action variation becomes:

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

Let the infinitesimal transformations be

$$\Delta\phi = \Phi_\nu \delta\omega^\nu$$
$$\Delta x^\mu = X^\mu_\nu \delta\omega^\nu \approx \delta x^\mu,$$

where Φ^μ_ν 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

$$\int_{\partial R} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \Delta\phi - \theta^\mu_\nu \delta x^\nu \right] d\sigma_\mu = 0$$
$$\int_{\partial R} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \Phi_\nu \delta\omega^\nu - \theta^\mu_\nu X^\nu_\mu \delta\omega^\nu \right] d\sigma_\mu = 0$$
$$\int_{\partial R} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \Phi_\nu - \theta^\mu_\nu X^\nu_\mu \right] \delta\omega^\nu \, d\sigma_\mu = 0.$$

Let us define

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

Now, because

$$\int_{\partial R} J^\mu_\nu \delta\omega^\nu \, d\sigma_\mu = 0$$

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

$$\int_{\partial R} J^\mu_\nu \, d\sigma_\mu = 0.$$

Now, recall Gauss' theorem one more time:

$$\int_{\partial R} J^\mu_\nu \, d\sigma_\mu = \int_R \partial_\mu J^\mu_\nu \, d^4x.$$

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

$$\partial_\mu J^\mu_\nu = 0.$$

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

$$Q_\nu = \int_\sigma J^\mu_\nu \, d\sigma_\mu.$$

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

$$Q_\nu = \int_V J^0_\nu \, d^3x.$$

Now, revisit Gauss' theorem:

$$\int_V \partial_0 J^0_\nu \, d^3x + \int_V \partial_i J^i_\nu \, d^3x = 0$$
$$\int_V \partial_0 J^0_\nu \, d^3x + 0 = 0$$
$$\frac{d}{dt}\int_V J^0_\nu \, d^3x = \frac{dQ_\nu}{dt} = 0.$$

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

$$X^\mu_\nu = \delta^\mu_\nu$$
$$\Phi_\mu = 0.$$

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

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

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

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

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:

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

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

$$\int \theta^0_i \, d^3x$$

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 4.2. 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

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

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

$$\begin{aligned}
\theta^{\mu\nu} &= g^{\lambda\nu}\theta^{\mu}_{\lambda} \\
&= g^{\lambda\nu}\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\partial_\lambda \phi - \delta^{\mu}_{\lambda}\mathcal{L}\right) \\
&= g^{\lambda\nu}(g^{\mu\lambda}(\partial_\lambda \phi) - \delta^{\mu}_{\lambda}\mathcal{L}) \\
&= g^{\lambda\nu}(\partial^\mu \phi\, \partial_\lambda \phi - \delta^{\mu}_{\lambda}\mathcal{L}) \\
&= (\partial^\nu \phi)(\partial^\mu \phi) - g^{\mu\nu}\mathcal{L}.
\end{aligned}$$

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

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

such that $\partial_\mu T^{\mu\nu} = 0$ and $T^{\mu\nu}$ is symmetric. We will not go into detail about this, but 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.

4.5 Complex Scalar Fields

4.6 Vector Fields: An Electromagnetic Example

Vector fields describe particles of spin 1 such as photons. Unlike 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.

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

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:

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

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

Let us consider the (cleverly chosen) Lagrangian density for electromagnetism:

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

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

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

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

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

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

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

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

$$\begin{aligned} \vec{\nabla} \times \vec{E} &= -\frac{\partial \vec{B}}{\partial t} \\ \vec{\nabla} \cdot \vec{B} &= 0. \end{aligned}$$

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:

$$\delta\left(j^\mu A_\mu\right) = j^\mu\, \delta A_\mu$$

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:

$$\begin{aligned} \delta\left(\frac{-1}{4}F^{\mu\nu}F_{\mu\nu}\right) &= \frac{-1}{4}\,\delta\left[\left(\partial^\mu A^\nu - \partial^\nu A^\nu\right)\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right)\right] \\ &= \frac{-1}{2}\,\delta\left(\partial^\mu A^\nu\, \partial_\mu A_\nu - \partial_\mu A_\nu\, \partial^\mu A^\nu\right) \\ &= \frac{-1}{2}\left(\partial^\mu\, \delta A^\nu\, \partial_\mu A_\nu + \partial^\mu A^\nu\, \partial_\mu\, \delta A_\nu - \partial_\mu\, \delta A_\nu\, \partial^\mu A^\nu - \partial_\mu A_\nu\, \partial^\mu\, \delta A^\nu\right). \end{aligned}$$

Raising and lowering indices gives

$$\begin{aligned} \delta\left(\frac{-1}{4}F^{\mu\nu}F_{\mu\nu}\right) &= \frac{-1}{2}\left(\partial^\mu\, \delta A^\nu\, \partial_\mu A_\nu + \partial_\nu A_\mu\, \partial^\nu\, \delta A^\mu - \partial_\mu\, \delta A_\nu\, \partial^\mu A^\nu - \partial^\nu A^\mu\, \partial_\nu\, \delta A_\mu\right) \\ &= \partial^\nu\, \delta A^\mu\, \partial_\nu A_\mu - \partial_\mu\, \delta A_\nu\, \partial^\mu A^\nu. \end{aligned}$$

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

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

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

$$\partial_\mu(\partial^\mu A^\nu)\,\delta A_\nu - (\partial^\nu\partial_\mu A_\nu)\delta A^\mu \equiv (\Box A^\mu - \partial^\mu\partial^\nu A_\nu)\;\delta A_\mu = 0$$

for all δA_μ, which forces the following identity:

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

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

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

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

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

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 $\mu = 0$:

$$\begin{aligned}\Box A_0 - \partial_0\,\partial^\nu A_\nu &= \partial^0\partial_0 A_0 + \partial^j\partial_j A_0 - \partial_0\partial^0 A_0 - \partial_0\partial^j A_j \\ &= \partial^j\partial_j A_0 - \partial_0\partial^j A_j \\ &= 0.\end{aligned}$$

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:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu,$$

and attempt to transform (**gauge transform**)

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

then we observe that

$$\begin{aligned} F'_{\mu\nu} &= \partial_\mu (A_\nu + \partial_\nu \Lambda(x)) - \partial_\nu (A_\mu + \partial_\mu \Lambda(x)) \\ &= \partial_\mu A_\nu - \partial_n u A_\mu \\ &= F^{\mu\nu}, \end{aligned}$$

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.

Note: This is a very hand-wavy-type of explanation. For a better, more detailed, and more satisfactory explanation, please refer to Sean Carroll's book, p. 40. I will come back to this section later when my understanding of classical field theory develops, and hopefully provide a better explanation here.

5 Gravitation and Lagrangian Formulation of General Relativity

5.1 Review of General Relativity & Curved Spacetime

5.1.1 General Relativity

Taking the speed of light, c, to be 1, the Einstein equation for general relativity is

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G T_{\mu\nu}.$$

Sometimes, it is more useful to write the Einstein's equation as

$$R_{\mu\nu} = 8\pi G \left(T_{\mu\nu} - \frac{1}{2}T g_{\mu\nu} \right).$$

In vacuum, where all components of the energy-momentum stress tensor are zero, the Einstein's equation becomes

$$R_{\mu\nu} = 0.$$

5.1.2 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:

$$\begin{aligned} D_\mu w_\nu &= \partial_\mu w_\nu - \Gamma^\lambda_{\mu\nu} w_\lambda \\ D_\mu v^\nu &= \partial_\mu v^\nu + \Gamma^\nu_{\lambda\nu} v^\lambda. \end{aligned}$$

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 5.1. 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**:

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2} g^{\lambda\rho} \left(\partial_\mu g_{\rho\nu} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu} \right).$$

Now, consider a contravariant vector V^μ, we can define the **divergence** in curved spacetime based on the flat 3-space definition

$$\text{div}\vec{V} = \partial_i V^i$$

by "contracting" an index and adding the Christoffel symbols:

$$D_\mu V^\mu = \partial_\mu V^\mu + \Gamma^{\mu}_{\mu\lambda} V^\lambda.$$

Next, invoking the definition of the Christoffel symbols, we can compute

$$\begin{aligned} \Gamma^{\mu}_{\mu\lambda} &= \frac{1}{2} g^{\mu\rho} (\partial_\mu g_{\rho\lambda} + \partial_\lambda g_{\rho\mu} - \partial_\rho g_{\mu\lambda}) \\ &= a \end{aligned}$$

5.2 Lagrangian Formulation

The action is

$$S = \int \mathcal{L}(\Phi^i, \nabla_\mu \Phi^i)\, d^n x.$$

$\mathcal{L}$ is a density.

$$\mathcal{L} = \sqrt{-g}\hat{\mathcal{L}}$$

where $\hat{\mathcal{L}}$ is a scalar. The associated Euler-Lagrange equation is

$$\frac{\partial \hat{\mathcal{L}}}{\partial \Phi} - \nabla_\mu \left(\frac{\partial \hat{\mathcal{L}}}{\partial (\nabla_\mu \Phi)} \right) = 0.$$

Recall Stokes' theorem:

$$\int_\Sigma \nabla_\mu V^\mu \sqrt{|g|}\, d^n x = \int_{\partial\Sigma} n_\mu V^\mu \sqrt{|\gamma|}, d^{n-1} x$$