# Notes on The Theoretical Minimum
# — Special Relativity and Classical Field Theory

Gui-Rong Liang

August 22, 2019

## Contents

# 1 Special Relativity

## 1.1 Derivation of the Lorentz transformation

To establish the Lorentz transformation, we need five small steps, or properties.

1.*Synchronization by light*

Suppose a primed Inertial Reference Frame is moving right relative to an unprimed IRF by a velocity $v$. Everyone will agree that $x' = 0$ corresponds to $x = vt$, but what's the expression of $t' = 0$ in terms of $x$ and $t$? $t' = 0$ means every event on $x'$-axis happens at the same time. We use light to synchronize two points and then we can determine the $x'$-aixs.

Imagine in the primed IRF there're three observers $A$, $B$, and $C$, who are expressed in the unprimed IRF as $x = vt$, $x - vt = 1$, $x - vt = 2$, if the observer $B$ in the middle simultaneously receives two signals from both sides, we can say that $A$ and $C$ are synchronized. The two light signals are represented by $x = t$ and $x + t = a$, where we have set the speed of light as $c = 1$ and $a$ is a parameter to be determined.

We first combine $x = t$ and $x - vt = 1$ to give the coordinate of the event $B$ receives the signal, it's $(x_B, t_B) = (\frac{1}{1-v}, \frac{1}{1-v})$. And then we plug in this coordinate to give the parameter in the second signal $a = \frac{2}{1-v}$. Next we combine the second signal $x + t = \frac{2}{1-v}$ and observer $C$'s world line $x - vt = 2$ to give the event $C$ sent the signal, it's $(x_C, t_C) = (\frac{2v}{1-v^2}, \frac{2}{1-v^2})$. Thus by connecting $(x_C, t_C)$ and the origin, we have the synchronization line as $t = vx$. Note that it is symmetric to $x = vt$ with respect to the light ray $x = t$.

We conclude the corresponding relationship as follow,

$$
\begin{aligned}
x' = 0 &\iff x - vt = 0 \\
t' = 0 &\iff t - vx = 0,
\end{aligned}
\tag{1}
$$

then we may write the general form of the transformation between primed and unprimed coordinates as

$$
\begin{aligned}
x' &= f(x - vt), \quad \text{with} \quad f(0) = 0, \\
t' &= g(t - vx), \quad \text{with} \quad g(0) = 0.
\end{aligned}
\tag{2}
$$

2.*Linearity property*

To proceed, we impose some physical considerations.

Firstly, if two meter-sticks have the same length in unprimed RF, they will still have the same length in the primed RF. If we join them head to tail, the coordinate relationship will be

$$
\begin{aligned}
x_1 - x_0 = x_2 - x_1 &\implies x_1' - x_0' = x_2' - x_1' \\
\text{or} \quad \Delta x_1 = \Delta x_2 &\implies \Delta x_1' = \Delta x_2'.
\end{aligned}
\tag{3}
$$

If the coordinate difference is infinitesimal, the relation can be re-expressed as

$$
\begin{aligned}
\mathrm{d}x_1 = \mathrm{d}x_2 &\implies \mathrm{d}x_1' = \mathrm{d}x_2' \\
\text{or} \quad \left.\frac{\partial x'}{\partial x}\right|_{x_1} &= \left.\frac{\partial x'}{\partial x}\right|_{x_2},
\end{aligned}
\tag{4}
$$

where in the second equation the $x_1$ and $x_2$ are arbitrary adjacent points in $x$-axis, so we conclude that the derivative must be a constant,

$$
\frac{\partial x'}{\partial x} = \text{const}(v),
\tag{5}
$$

where we should note that the constant may depend on the relative velocity $v$. This is the linearity property of $x'$ with respect to $x$.

Secondly, imagine you are standing in a train station, and three trains pass you by the same time interval, then you will agree in the train RF, they are evenly spaced. By formulas, we have

$$t_1 - t_0 = t_2 - t_1 \implies x'_1 - x'_0 = x'_2 - x'_1$$

$$\text{or} \qquad \frac{\partial x'}{\partial t} = \text{const}(v). \tag{6}$$

This is the linearity property of $x'$ with respect to $t$. To conclude the above two considerations, we my write the first transformation relation as,

$$x' = f(v)(x - vt), \tag{7}$$

where we have taken $f(0) = 0$ into account, thus the two constant are the same, and $x - vt$ always appears as a whole.

Thirdly, if two time intervals are the same in one RF, they will be the same in another, so

$$\frac{\partial t'}{\partial t} = \text{const}(v). \tag{8}$$

Fourthly, if two distance are the same in the unprimed RF, observers in the primed RF will feel the same time interval when passing by them, so

$$\frac{\partial t'}{\partial x} = \text{const}(v). \tag{9}$$

To conclude the above two linearities, considering $g(0) = 0$, we have

$$t' = g(v)(t - vx). \tag{10}$$

3.*Invariance of light speed*

Next we use the fact that the speed of light is the same in every RFs,

$$x = t \implies x' = t', \tag{11}$$

we can easily find that

$$f(v) = g(v). \tag{12}$$

4.*Reflection symmetry*

Note that we have chosen the right-oriented coordinate system so that the equation describes the primed RF moving right relative to the unprimed RF, manifestly, we write the equation as

$$x'_R = f(v)(x_R - vt). \tag{13}$$

But the same equation can refer to a left-moving primed RF relative to the unprimed RF in a left-oriented coordinate system,

$$x'_L = f(v)(x_L - vt). \tag{14}$$

Then how do we describe a right-moving RF in a left-oriented coordinate system? We change the sign of the speed,

$$x'_L = f(-v)(x_L + vt). \tag{15}$$

Note that coordinates in left-oriented coordinate system differs from those in right-oriented coordinate system by a minus sign, $x_L = -x_R$ and $x'_L = -x'_R$. So we change the above equation back to the right-oriented coordinate system,

$$-x'_R = f(-v)(-x_R + vt)$$

$$\text{or} \qquad x'_R = f(-v)(x_R - vt). \tag{16}$$

So we have two equations describing the same phenomenon,

$$\begin{cases} x'_R = f(v)(x_R - vt) \\ x'_R = f(-v)(x_R - vt). \end{cases} \tag{17}$$

Thus there must be

$$f(v) = f(-v). \tag{18}$$

Actually the $v$ in $f(v)$ can point to any directions in 2 or 3 dimensions, so $f(v)$ can only depend on the magnitude of $v$, not direction. At the end we will find it as $f(v^2)$.

5.*Equivalence of reference frames*

When looking at the primed RF in unprimed RF, we have

$$\begin{cases} x' = f(v)(x - vt) \\ t' = f(v)(t - vx). \end{cases} \tag{19}$$

Relatively, when looking at the unprimed RF in primed RF, it's moving to the opposite direction, so we have

$$\begin{cases} x = f(-v)(x' + vt') = f(v)(x' + vt') \\ t = f(-v)(t' + vx') = f(v)(t' + vx'). \end{cases} \tag{20}$$

Joining the above two sets of equations, we have

$$\begin{aligned} x' &= f(v)(x - vt) \\ &= f(v)^2[(x' + vt') - v(t' + vx')] \\ &= f(v)^2(1 - v^2)x', \end{aligned} \tag{21}$$

thus we find out the form of $f(v)$,

$$f(v) = \frac{1}{\sqrt{1 - v^2}}. \tag{22}$$

So the Lorentz transformation is given by

$$\begin{cases} x' = \dfrac{x - vt}{\sqrt{1 - v^2}} \\ t' = \dfrac{t - vx}{\sqrt{1 - v^2}}. \end{cases} \tag{23}$$

Recovering the normal units, we have

$$\begin{cases} x' = \dfrac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}} \\ t' = \dfrac{t - \frac{v}{c^2}x}{\sqrt{1 - \frac{v^2}{c^2}}}. \end{cases} \tag{24}$$

We see that when $v/c \ll 1$, we recover the Galilieo's transformation,

$$\begin{cases} x' = x - vt \\ t' = t. \end{cases} \tag{25}$$

Moreover, the positivity of $1 - \frac{v^2}{c^2}$ gives the famous conclusion: nothing's moving faster than light.

## 1.2  Inferences from LT and relativistic effects

*1.Addition of the velocity*

Lorentz transformations can be done twice, through which we could find the rule of addition velocity. Now let's say the double-primed RF is moving relative to the primed RF with speed $u$, and primed to the rest $v$, so we have

$$x'' = \frac{x' - ut'}{\sqrt{1 - u^2}} = \frac{\frac{x - vt}{\sqrt{1 - v^2}} - u\frac{t - vx}{\sqrt{1 - v^2}}}{\sqrt{1 - u^2}} = \frac{(1 + uv)x - (u + v)t}{\sqrt{1 - u^2}\sqrt{1 - v^2}}. \tag{26}$$

By analysis the numerator, we know that $x'' = 0$ is the world line the double-primed observer, its equation in the rest RF is $x = \frac{u+v}{1+uv}t$, so we conclude the total velocity $w$ is $\frac{u+v}{1+uv}$. But more specifically, we want to verify it in explicit transformation between $x''$ and $x$, $t$, by dividing both the numerator and the denominator the factor $1 + uv$, we have the denominator as,

$$\sqrt{\frac{1 + u^2v^2(+2uv - 2uv) - u^2 - v^2}{(1 + uv)^2}} = \sqrt{1 - \left(\frac{u + v}{1 + uv}\right)^2}, \tag{27}$$

so, to conclude,

$$x'' = \frac{x - wt}{\sqrt{1 - w^2}}, \quad \text{where} \quad w = \frac{u + v}{1 + uv}. \tag{28}$$

The extreme example of the velocity addition rule is that both $u$ and $v$ are equal to the speed of light 1, we see that the resulting $w$ is also 1. Nothing's faster than light.

*2.Length contraction and time dilation*

Next we're going to explore some of the relativistic effects simply from Lorentz transformations.

If a stick with length $L$ is at rest in the unprimed RF, what's its length measured in the primed RF? The point here is, in the primed RF, we must measure it's both ends at same time. To convert it mathematically, we want to find the $x'$ coordinate of the cross point of $x = L$ and $t' = 0\,(t = vx)$. Using the Lorentz transformation, we soon get

$$L' = x' = \frac{x - vt}{\sqrt{1 - v^2}} = \frac{L - v^2L}{\sqrt{1 - v^2}} = \sqrt{1 - v^2}L < L, \tag{29}$$

thus the length is contracted in the moving RF. This is the *length contraction.*

If a clock at the origin runs time $T$ at unprimed RF, how long does the observer feel in the primed RF? Mathematically, we want to find the $t'$ coordinate of the point $(0, t)$. Very soon by Lorentz transformation we get

$$t' = \frac{t - vx}{\sqrt{1 - v^2}} = \frac{T}{\sqrt{1 - v^2}} > T, \tag{30}$$

which says time dilates in the moving RF. This is the *time dilation.*

*3.Spacetime intervals and 4-velocity*

By a simple calculation, we could find an invariant quantity under Lorentz transformations.

$$t'^2 - x'^2 = \frac{t^2 + v^2x^2 - 2vtx}{1 - v^2} - \frac{x^2 + v^2t^2 - 2vtx}{1 - v^2} = t^2 - x^2 \equiv \tau^2, \tag{31}$$

where $\tau$ is called the proper time. Its physical meaning is shown as follows, on the $t'$-axis of the primed RF, $x' = 0$, the $t'$ coordinate is nothing but $\tau$, so $\tau$ represents time of a moving body. In 4 dimensions, the spacetime interval is represented as,

$$s^2 = -\tau^2 = -t^2 + x^2 + y^2 + z^2, \tag{32}$$

which also includes rotation symmetries. And more generally it's represented in terms of infinitesimal intervals,

$$ds^2 = -d\tau^2 = -dt^2 + dx^2 + dy^2 + dz^2. \tag{33}$$

When $ds^2 < 0$, it's called a timelike separation; when $ds^2 > 0$, spacelike; and $ds^2 = 0$, lightlike.

We denote $(t, x, y, z)$ by $(x^0, x^1, x^2, x^3)$, or $x^\mu$, where $\mu$ runs from 0 to 4. The 4-velocity is defined as,

$$U^\mu := \frac{dx^\mu}{d\tau}. \tag{34}$$

The zeroth component is

$$U^0 = \frac{dt}{d\tau} = \frac{dt}{\sqrt{dt^2 - dx^2}} = \frac{1}{\sqrt{1 - v^2}}, \tag{35}$$

and the spatial components are

$$U^i = \frac{dx^i}{d\tau} = \frac{dx^i}{dt}\frac{dt}{d\tau} = \frac{V^i}{\sqrt{1 - v^2}}, \tag{36}$$

where $V^i$ is the 3-velocity. The 4 components of the 4-velocity are not independent, they are constraint by

$$d\tau^2 = dt^2 - \sum_i (dx^i)^2 \implies 1 = (U^0)^2 - (\vec{U})^2 \tag{37}$$

## 1.3 Particle mechanics

The **Least Action Principle** serves as one of the fundamental rules of physics, and will be our starting point.

To make sure the laws of motion are the same in every RF, the action should be invariant. There's really only one thing that's invariant when a particle moves from some position to a neighboring position — the proper time separating the two. If the particle moves from spacetime $A$ to spacetime $B$, the invariant should be the integral of the proper time, or multiplying an arbitrary constant.

$$\mathcal{A} = -m \int_a^b d\tau, \tag{38}$$

where the minus sign is just due to convention, and $m$ just a parameter but eventually we will find it's the mass of the particle.

To go back to our familiar form of action, we rewrite it as

$$\mathcal{A} = -m \int_a^b \sqrt{1 - v^2} \, dt, \tag{39}$$

thus we can extract the Lagrangian as

$$L = -m\sqrt{1 - v^2}, \tag{40}$$

or to recover the normal units as

$$L = -mc^2 \sqrt{1 - \frac{v^2}{c^2}}. \tag{41}$$

When taking the low speed limit we have

$$L = -mc^2 \left(1 - \frac{v^2}{2c^2}\right) = \frac{1}{2}mv^2 - mc^2. \tag{42}$$

We see the good old Lagrangian of a free particle $\frac{1}{2}mv^2$, and adding or subtracting a constant doesn't affect the motion of the particle.

Let's come back to our relativistic version and derive its canonical momentum,

$$P^i = \frac{\partial L}{\partial \dot{x}^i} = \frac{m\dot{x}^i}{\sqrt{1-v^2}} = \frac{m\,\mathrm{d}x^i}{\mathrm{d}t\sqrt{1-v^2}} = m\frac{\mathrm{d}x^i}{\mathrm{d}\tau} = mU^i, \tag{43}$$

where we can see that it's just mass times the spatial component of the 4-velocity.

The most important quantity the Hamiltonian is given by

$$H = \sum_i P^i \dot{x}^i - L = \frac{mv^2}{\sqrt{1-v^2}} + m\sqrt{1-v^2} = \frac{m}{\sqrt{1-v^2}} = mU^0 \equiv P^0, \tag{44}$$

which shows that the energy is just the zeroth component of the 4-momentum $P^\mu$. Thus the conservation of energy and momentum can be expressed in a single equation,

$$\frac{\mathrm{d}P^\mu}{\mathrm{d}t} = 0. \tag{45}$$

Now we recover the energy expression in normal units and take slow speed limit,

$$E = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} \sim mc^2 + \frac{1}{2}mv^2, \tag{46}$$

where the second term is the kinetic energy, and the first term is the rest energy contained in the mass. This is the origin of the famous Einstein's mass-energy relation,

$$E_0 = mc^2. \tag{47}$$

Then what's the energy of a massless particle? Since the speed of a massless particle is the light speed 1, the numerator and the denominator in the Hamiltonian tends to be the 0/0 type, which is undetermined. The knob here is to represent the energy not through the velocity (every massless particle has the same velocity — the speed of light), but through the momentum (and energies of massive particles can also be expressed in terms of momentum).

We multiply the normalization relationship of the 4-velocity by mass squared,

$$m^2 = m^2(U^0)^2 - m^2(\vec{U})^2 = E^2 - P^2, \tag{48}$$

thus we have the energy in terms of momentum

$$E = \sqrt{P^2 + m^2} \longrightarrow \sqrt{P^2c^2 + m^2c^4}. \tag{49}$$

Now we can see when the mass is zero, the energy would be

$$E = c\,P, \tag{50}$$

where $P$ is the magnitude of the momentum.

# 2    Scalar field theory

## 2.1    The field Euler-Lagrangian equation and generalized harmonic oscillator

From the basics of classical mechanics, an action for a non-relativistic particle moving along one dimensional space is given by,

$$\mathcal{A} = \int_a^b L(q, \dot{q}) \, \mathrm{d}t, \tag{51}$$

in which $q(t)$ denotes the spatial coordinate varying with time.
The Lagrangian $L$ is has the form of kinetic energy minus potential energy,

$$L = \frac{1}{2} \left( \frac{\mathrm{d}q}{\mathrm{d}t} \right)^2 - V(q), \tag{52}$$

where the mass $m$ is set to 1.
The Euler-Lagrange equation tell us how to minimize the action and provide the equation of motion,

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q}, \tag{53}$$

by solving which we have the Newton's equation,

$$\frac{\mathrm{d}^2 q}{\mathrm{d}t^2} = -\frac{\partial V(q)}{\partial q}. \tag{54}$$

Note that here $q$ is the spatial coordinate, but it can be any scalar quantity which depends only on $t$, for example, the temperature T(t) at some place. So we may give this kind of quantity a general name $\phi$, and treat it a field theory in $(0 + 1)$ dimension.

*A particle theory in one dimension of space has the same mathematical structure as a scalar field theory in zero dimension of space.*

Now we want to generalize it to $(d + 1)$ dimensions, but firstly focus on the case in $d = 1$. Parallel to the above case, the action is constructed by summing Lagrangian over tiny spacetime cells,

$$\mathcal{A} = \int \int \mathcal{L} \, \mathrm{d}t \, \mathrm{d}x, \tag{55}$$

with boundary value fixed.
Then we see from the above that the particle Lagrangian depends only on the position and the velocity. We may expect the field Lagrangian also depends on the field values and its first-order derivatives (including its spatial derivatives), which is the minimum requirement of **Locality**. *Locality means that things happening at one place only directly affect conditions nearby in space and time.* Higher derivatives are ruled out because they're "less local' than first derivatives.
Then we may write our field Lagrangian as

$$\mathcal{L} = \mathcal{L}(\phi, \frac{\partial \phi}{\partial t}, \frac{\partial \phi}{\partial x}). \tag{56}$$

In multi-dimensional case, the Euler-Lagrangian equation is generalized by incorporating contributions from all spacetime direction, thus it would be

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial \left( \frac{\partial \phi}{\partial t} \right)} + \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial \mathcal{L}}{\partial \left( \frac{\partial \phi}{\partial x} \right)} = \frac{\partial \mathcal{L}}{\partial \phi}. \tag{57}$$

We proceed to modify the kinetic term of the original Lagrangian by accounting the effect of the $x$ component,

$$\mathcal{L} = \frac{1}{2}\left[\left(\frac{\partial \phi}{\partial t}\right)^2 - \left(\frac{\partial \phi}{\partial x}\right)^2\right] - V(\phi), \tag{58}$$

where the minus sign before the quadratic term arises as a must of **Lorentz Invariance**. This will be the Lagrangian for our field theory, and we view it as a prototype.

The resulting EOM is,

$$\frac{\mathrm{d}^2\phi}{\mathrm{d}t^2} - \frac{\mathrm{d}^2\phi}{\mathrm{d}x^2} + \frac{\partial V}{\partial \phi} = 0. \tag{59}$$

We list several choices of $V(\phi)$ to give explicitly different EOMs.

In the first case, we consider the equation without the source term by setting $V(\phi) = 0$, and meanwhile we go back to the conventional units by restoring the factor of $c$, we get the familiar wave equation,

$$\frac{1}{c^2}\frac{\mathrm{d}^2\phi}{\mathrm{d}t^2} - \frac{\mathrm{d}^2\phi}{\mathrm{d}x^2} = 0, \tag{60}$$

and any function of the form $F(x + ct)$ or $F(x - ct)$ will be its solution.

Another basic example is to choose $V(\phi) = \frac{\mu^2}{2}\phi^2$, where $\mu^2$ resembles the spring constant $k$,

$$\frac{\mathrm{d}^2\phi}{\mathrm{d}t^2} - \frac{\mathrm{d}^2\phi}{\mathrm{d}x^2} + \mu^2\phi = 0, \tag{61}$$

and this is the analog version of the Simple Harmonic Oscillator, called Klein-Gordon equation.

Since the above cases are all concerned with linear terms (with only first powers of $\phi$ and its derivatives), their general solutions are linear combination of the basics forms like $\phi = e^{-i\omega t}e^{ikx}$, taking which into the equation gives the constraint relation $\omega = \pm\sqrt{k^2 + \mu^2}$, a solution is given when $k$ is determined.

A third example is to choose $V(\phi) = \frac{\mu^2}{2}\phi^2 + g\phi^3$, containing a $\phi$ to the third power, then the EOM is given by

$$\frac{\mathrm{d}^2\phi}{\mathrm{d}t^2} - \frac{\mathrm{d}^2\phi}{\mathrm{d}x^2} = -\mu^2\phi - 3g\phi^2, \tag{62}$$

which is non-linear in $\phi$, and this form reflects the interaction phenomenom.

If the Lagrangian contains two fields $\phi$ and $\psi$, then there may or may not be a mix term representing the interaction (like scattering) between them,

$$\mathcal{L} = \mathcal{L}(\phi, \partial_\mu\phi) + \mathcal{L}(\psi, \partial_\mu\psi) + \mathcal{L}_{int}(\phi, \psi, \partial_\mu\phi, \partial_\mu\psi), \tag{63}$$

where the last term is the interaction term, we set an example to choose it to be $\mathcal{L}_{int} = \phi\psi^2$, then the extra terms in separate EOM of $\phi$ and $\psi$ are given by

$$\begin{cases} \dfrac{\partial \mathcal{L}_{int}}{\partial \phi} = \psi^2 \\[2mm] \dfrac{\partial \mathcal{L}_{int}}{\partial \psi} = 2\phi\psi, \end{cases} \tag{64}$$

which says the EOM of $\phi$ has the contribution from $\psi$ and the EOM of $\psi$ has the contribution from $\phi$, the two fields talk to each other.

Then it's easy to generalize all the above set to $(3 + 1)$ dimensions, where we will use the metric $\eta_{\mu\nu} = \text{diag}\{-1, 1, 1, 1\}$ to raise and lower indices,

$$\mathcal{A} = \int \mathcal{L}(\phi, \partial_\mu\phi)\,\mathrm{d}^4x = \int\left[\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi)\right]\mathrm{d}^4x,$$

$$\partial_\mu\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)} - \frac{\partial \mathcal{L}}{\partial \phi} = 0, \tag{65}$$

$$\partial_\mu\partial^\mu\phi + \frac{\partial V}{\partial \phi} = 0.$$

Above formalisms are just simple extensions of non-relativistic particle motion. Generally, due to the fact that an action should be invariant in any reference frame, the Lagrangian must be a scalar. Constructed out of the field $\phi$ and its first-time derivative $\partial_\mu \phi$, there are many applicants for the Lagrangian, such as: any functions of $\phi$, $\partial_\mu \phi \partial^\mu \phi$, and any combinations of them (eg. $\mathcal{L} = \phi^2(\partial_\mu \phi \partial^\mu \phi)$).

## 2.2 The canonical momentum, the Hamiltonian density, and the Noether theorem

The canonical quantities are deduced from classical mechanics. Firstly we treat the field $\phi$ as a coordinate as compared with $q$, and the variable $x$ a degree of freedom as compared with the label $i$, which means at a given point in $x$, the field value $\phi_x$ is free to change, without any influence from values at other points $\phi_y$s. Then the action of a field can be rewrite in the classical form,

$$
\begin{aligned}
\mathcal{A} &= \int \int \mathcal{L}(\phi(x,t), \frac{\partial \phi(x,t)}{\partial t}, \frac{\partial \phi(x,t)}{\partial x}) \, \mathrm{d}x \, \mathrm{d}t \\
&= \int \int \mathcal{L}(\phi_x, \dot{\phi}_x) \, \mathrm{d}x \, \mathrm{d}t \\
&= \int L(\phi_{x_0}, \phi_{...}, \dot{\phi}_{x_0}, \dot{\phi}_{...}) \, \mathrm{d}t, \quad (x_0, ... \in \{x\})
\end{aligned}
\tag{66}
$$

where $L = \int \mathcal{L} \, \mathrm{d}x$, so we call the $\mathcal{L}$ the *Lagrangian density*. Note here the derivative of the field with respect to space is nothing but a function of the field $\phi(x)$s,

$$
\frac{\partial \phi}{\partial x} = \frac{\phi(x+\epsilon) - \phi(x)}{\epsilon}
\tag{67}
$$

where $\epsilon$ is a small quantity. So it's absorbed in the $\phi_x$ in $\mathcal{L}(\phi_x, \dot{\phi}_x)$. From the classical definition of canonical momentum, we construct the field canonical momentum,

$$
p^i = \frac{\partial L}{\partial \dot{q}_i} \quad \rightarrow \quad p^y = \frac{\partial L}{\partial \dot{\phi}_y} = \frac{\partial \int \mathcal{L}(\phi_x, \dot{\phi}_x) \, \mathrm{d}x}{\partial \dot{\phi}_y} = \frac{\partial \mathcal{L}(\phi_y, \dot{\phi}_y)}{\partial \dot{\phi}_y} \equiv \pi^y,
\tag{68}
$$

where $y$ is a specific number in $x$, and $\pi^y = \frac{\partial \mathcal{L}}{\partial \dot{\phi}_y}$ is the $y$th component of the field canonical momentum. One can treat the integral in the numerator as a sum over the dummy index $x$, thus the derivative of $L$ with respect to the specific label $y$ is just the derivative of the term corresponding to that specific label. Since this $y$ is free to change, we recover the normal variable $x$ to give its definition,

$$
\pi(x) := \frac{\partial \mathcal{L}}{\partial \dot{\phi}}(x).
\tag{69}
$$

We proceed to construct the field Hamiltonian still from the classical definition,

$$
H = \sum_i \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L \quad \rightarrow \quad H = \int \frac{\partial \mathcal{L}}{\partial \dot{\phi}_x} \dot{\phi}_x \, \mathrm{d}x - L = \int \left( \frac{\partial \mathcal{L}}{\partial \dot{\phi}_x} \dot{\phi}_x - \mathcal{L} \right) \mathrm{d}x \equiv \int \mathcal{H} \, \mathrm{d}x,
\tag{70}
$$

where we have use the fact that $\frac{\partial L}{\partial \dot{\phi}} = \frac{\partial \mathcal{L}}{\partial \dot{\phi}}$. Thus we have the definition of a *Hamiltonian density*,

$$
\mathcal{H} := \pi(x) \dot{\phi}(x) - \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \dot{\phi}}(x) \dot{\phi}(x) - \mathcal{L}.
\tag{71}
$$

Next we try the simplest Lagrangian $\mathcal{L} = \frac{1}{2} \dot{\phi}^2 - \frac{1}{2} \phi_{,x}^2 - V(\phi)$, and explore its canonical quantities. The canonical momentum is just

$$
\pi = \dot{\phi},
\tag{72}
$$

and the Hamiltonian density, (or the energy density,) is

$$
\mathcal{H} = \frac{1}{2} \dot{\phi}^2 + \frac{1}{2} \left( \frac{\partial \phi}{\partial x} \right)^2 + V(\phi) \equiv \frac{1}{2} \dot{\phi}^2 + V_{eff}(\phi),
\tag{73}
$$

where the $V_{eff}$ plays the role of potential energy rather than the field potential $V(\phi)$.

We can see from its form that apart from $V(\phi)$, the energy is always positive. Further if the field potential $V(\phi)$ is bounded form below (,which is a desired assumption), we can easily change the zero point by adding a constant. This is our first artificial requirement — *the positivity of energy.*

We previously treat $\phi(t, x)$ as an independent degree of freedom for each value of $x$. If it is really so, the difference of fields at two adjacent point changing rapidly will cause the spatial derivative to be huge, or even infinite, let alone the square of it — the energy. Definitely we want a finite energy, so this must be a constraint to the "independence of degree", telling $\phi(t, x)$ to be smooth, and putting the brakes on how wildly $\phi$ can vary from point to point. This will be the second artificial requirement — *the finiteness of energy.*

Now let's recall the Noether Theorem in classical mechanics, which says,

*If the Lagrangian is invariant under the change of $\delta q_i = f_i(q)\epsilon$, the corresponding charge $Q = \sum_i p_i f_i(q)$ will be a conserved quantiy.*

As a review, we prove it in the regime of classical field theory, where we treat the variable $x$ as a degree of freedom. If the Lagrangian is invariant under the transformation $\delta\phi_x = f_x(\phi)\epsilon$, we have

$$
\begin{aligned}
0 = \delta L &= \delta \int \mathcal{L}(\phi_x, \dot{\phi}_x)\,\mathrm{d}x \\
&= \int \left( \frac{\partial \mathcal{L}}{\partial \dot{\phi}_x} \delta\dot{\phi}_x + \frac{\partial \mathcal{L}}{\partial \phi_x} \delta\phi_x \right) \mathrm{d}x \\
&= \int \left( \pi^x \delta\dot{\phi}_x + \dot{\pi}^x \delta\phi_x \right) \mathrm{d}x \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \int \pi^x \delta\phi_x \,\mathrm{d}x \\
&= \epsilon \frac{\mathrm{d}}{\mathrm{d}t} \int \pi^x f_x(\phi)\,\mathrm{d}x,
\end{aligned}
\tag{74}
$$

thus we have the conserved quantity

$$
Q = \int \pi(x) f(\phi)\,\mathrm{d}x.
\tag{75}
$$

But in field theory, the change of the field is generally due to the change in the coordinate, which is

$$
\delta\phi = \phi(x - \epsilon) - \phi(x) = -\epsilon \frac{\partial \phi}{\partial x},
\tag{76}
$$

where we take an *active view* of change — translating the real field from coordinate $x$ to coordinate $x + \epsilon$. Here $f(\phi) = -\frac{\partial\phi}{\partial x}$, thus the conserved charge is

$$
Q = -\int \pi(x) \frac{\partial\phi(x)}{\partial x}\,\mathrm{d}x.
\tag{77}
$$

To conclude Noether theorem in classical field theory, we state that

*If a physical system has the symmetry under transformation $\delta x = -\epsilon$, the corresponding quantity $Q = -\int \pi(x)\partial_x\phi(x)\,\mathrm{d}x$ is conserved.*

## 2.3 Interactions between a particle and the scalar field

The action of a free particle, which is Lorentz invariant, is given by,

$$
\mathcal{A}_{par} = -m \int \mathrm{d}\tau = -m \int \sqrt{\mathrm{d}t^2 - \mathrm{d}x^2} = -m \int \sqrt{1 - \dot{x}^2}\,\mathrm{d}t,
\tag{78}
$$

11

where the minus sign is just a convention.

We can immediately recognize the Lagrangian as $-m\sqrt{1 - \dot{x}^2}$ and found that it matches the familiar classical Lagrangian at low velocities when expanded.

Now we impose a scalar field into this Lagrangian, the simplest way is by adding a term which indicates integrating the field along the trajectory of the moving particle, thus making the field variables $t$ and $x$ to be the same with the particle's variables,

$$
\begin{aligned}
\mathcal{A}_{par-field} &= -m \int d\tau - \int \phi(t,x)\, d\tau \\
&= -\int [m + \phi(t,x)]\, d\tau \\
&= -\int [m + \phi(t,x)]\sqrt{1 - \dot{x}^2}\, dt
\end{aligned}
\tag{79}
$$

*1. Higgs Mechanism*

From the action we will discover an interesting property of the effect acted on the particle by the field — if the field $\phi(t,x)$ somehow in some region approaches a non-zero constant, the motion of the particle would be like a particle moves with a mass $m + \phi$; and what's more, if a massless particle is coupled to this field, this coupling can effectively shift the particle's mass to a non-zero value! This is much like the case when Higgs field gives a particle its mass!

To make this simplified Higgs Mechanism clear again, we extract the Lagrangian as

$$
\mathcal{L} = -[m + \phi(t,x)]\sqrt{1 - \dot{x}^2},
\tag{80}
$$

and take its derivative with respect to $x$ to get the canonical momentum,

$$
\frac{\partial \mathcal{L}}{\partial \dot{x}} = \frac{[m + \phi(t,x)]\dot{x}}{\sqrt{1 - \dot{x}^2}},
\tag{81}
$$

which supports the notion that the expression $[m + \phi(t,x)]$ behaves like a position-dependent mass.

*2. Low-speed expansion and potential term*

Now we recover the normal units and expand the Lagrangian in the low speed limit,

$$
\begin{aligned}
\mathcal{L} &= -[mc^2 + g\phi(t,x)]\sqrt{1 - \frac{\dot{x}^2}{c}} \\
&\approx -[mc^2 + g\phi(t,x)]\left(1 - \frac{\dot{x}^2}{2c^2}\right) \\
&\sim \frac{1}{2}m\dot{x}^2 - g\phi(t,x),
\end{aligned}
\tag{82}
$$

where $g$ is coupling constant, telling us how strong the effect the field acts on the particle (the most familiar coupling constant being the electric charge $e$). In the last step we dropped the constant term $mc^2$ and the infinitesimal term $g\phi(t,x)\frac{\dot{x}^2}{2c^2}$.

We readily recognize this form as $\mathcal{L} = T - V$ in our classical mechanism class, with $g\phi(t,x)$ serves as the potential energy term. At this point, we clearly see the fact that *interaction gives rise to potential*.

*3. The interaction term when the particle is at rest*

Suppose the particle is at rest with $x = 0$ (and $\dot{x} = 0$), so its world line is just a vertical line. In this case, the Lagrangian is only left with an interaction term,

$$
\begin{aligned}
\mathcal{L}_{int} &= -g\phi(t, 0) \\
&= \int -g\phi(t, x)\delta(x)\,\mathrm{d}x.
\end{aligned} \tag{83}
$$

So that the interaction action will be

$$
\mathcal{A}_{int} = \int \mathcal{L}_{int}\,\mathrm{d}t = \int \int -g\phi(t, x)\delta(x)\,\mathrm{d}x\,\mathrm{d}t. \tag{84}
$$

We claim that *this interaction term is shared both by the particle action and the field action*.

*4.The back-reaction to the field by the particle*

The action of a field, when taking into account the interaction by the rest particle, is given by

$$
\begin{aligned}
\mathcal{A}_{field-par} &= \mathcal{A}_{field} + \mathcal{A}_{int} \\
&= \int \int \left\{ \frac{1}{2}\left[ \left(\frac{\partial\phi}{\partial t}\right)^2 - \left(\frac{\partial\phi}{\partial x}\right)^2 \right] - g\phi(t, x)\delta(x) \right\}\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned} \tag{85}
$$

Note that this is also the total action of the "field and particle", where the the action of the rest particle $\mathcal{A}_{par}(rest) = \int \int \frac{1}{2}m\dot{x}^2\,\mathrm{d}x\,\mathrm{d}t = 0$. We extract its Lagrangian,

$$
\mathcal{L}_{field-par} = \frac{1}{2}\left[ \left(\frac{\partial\phi}{\partial t}\right)^2 - \left(\frac{\partial\phi}{\partial x}\right)^2 \right] - g\phi(t, x)\delta(x), \tag{86}
$$

and get its EOM through the Euler-Lagrangian equation,

$$
\frac{\mathrm{d}^2\phi}{\mathrm{d}t^2} - \frac{\mathrm{d}^2\phi}{\mathrm{d}x^2} = -g\delta(x). \tag{87}
$$

When generalizing to $d = 3$ case, and consider a dynamic field resulting from a moving particle with its positon $a(t)$, we have

$$
\partial_\mu\partial^\mu\phi(t, x) = g\,\delta^3(x - a(t)), \tag{88}
$$

or consider a static field $\phi(x)$ with a rest particle in it, we have

$$
\nabla^2\phi(x) = g\,\delta^3(x), \tag{89}
$$

which is nothing but the *Poisson's equation*.

13

# 3 Vector field theory — the electro-magnetic field

## 3.1 A particle moving in the electro-magnetic field and the Lorentz force law

The electric and magnetic field can be derived from a more intrinsic field – *the vector potential* $A_\mu(x,t)$ (note here $x$ represents for all spatial coordinates); the simplest scalar quantity constructed from the vector potential and the particle's local coordinates is $A_\mu(x,t)\,\mathrm{d}x^\mu$ (another less simple scalar is $A_\mu A^\mu\,\mathrm{d}\tau$, which however violates the gauge invariance we will mention later), thus we integrate it along the trajectory of the particle to get the action, which is much similar to the case of the scalar field,

$$
\begin{aligned}
\mathcal{A} &= -m \int \mathrm{d}\tau + e \int A_\mu(x,t)\,\mathrm{d}x^\mu \\
&= \int \left( -m\sqrt{1-\dot{x}^2} + eA_\mu(x,t)\dot{x}^\mu \right)\,\mathrm{d}t.
\end{aligned}
\tag{90}
$$

where $e$ is the coupling constant, which we will later recognize it as the electric charge.

1.*Equation of Motion, the electro-magnetic tensor*

We extract its Lagrangian as,
$$
\mathcal{L} = -m\sqrt{1-\dot{x}^2} + eA_\mu(x,t)\dot{x}^\mu,
\tag{91}
$$
and take its derivative with respect to the $i$-component of the velocity to get the canonical momentum,

$$
\begin{aligned}
\mu P^i = \frac{\partial \mathcal{L}}{\partial \dot{x}^i} &= \frac{m\dot{x}_i}{\sqrt{1-\dot{x}^2}} + eA_i(x,t) \\
&= \frac{m\,\mathrm{d}x_i}{\sqrt{1-\dot{x}^2}\,\mathrm{d}t} + eA_i(x,t) \\
&= m\frac{\mathrm{d}x_i}{\mathrm{d}\tau} + eA_i(x,t) \\
&\equiv mU_i + eA_i(x,t),
\end{aligned}
\tag{92}
$$

where $U_i$ is the 4-velocity of the particle. We proceed to take its time derivative to get prepared for the EOM,

$$
\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^i}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m\dot{x}_i}{\sqrt{1-\dot{x}^2}}\right) + e\frac{\partial A_i(x,t)}{\partial x^j}\dot{x}^j + e\frac{\partial A_i(x,t)}{\partial t},
\tag{93}
$$

where $A_i(x,t)$ not only depends directly on $t$, but also through $x$. We expand the right-hand side of the EOM as,

$$
\frac{\partial \mathcal{L}}{\partial x^i} = e\frac{\partial A_\mu(x,t)}{\partial x^i}\dot{x}^\mu = e\frac{\partial A_0(x,t)}{\partial x^i} + e\frac{\partial A_j(x,t)}{\partial x^i}\dot{x}^j.
\tag{94}
$$

We equate the LHS and the RHS and collect terms with and without the velocity $\dot{x}^j$,

$$
m\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{x}_i}{\sqrt{1-\dot{x}^2}}\right) = e\left(\frac{\partial A_0}{\partial x^i} - \frac{\partial A_i}{\partial x^0}\right) + e\left(\frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j}\right)\dot{x}^j.
\tag{95}
$$

In the low-speed limit the LHS of the above is nothing but "mass times acceleration". We will assign the 1st term of the RHS as the electric field,

$$
E_i \equiv \frac{\partial A_0}{\partial x^i} - \frac{\partial A_i}{\partial x^0}
\tag{96}
$$

and manipulate the 2nd term of the RHS to recognize it as $e\,\vec{v}\times(\vec{\nabla}\times\vec{A})$. Thus equation (95) is nothing but the Lonrentz force law,

$$
m\vec{a} = e\,(\vec{E} + \vec{v}\times\vec{B}).
\tag{97}
$$

To make it a nice neat form and manifestly Lorentz invariant, we proceed to simplify it, firstly multiplying the whole equation (95) by a factor of $\frac{dt}{d\tau}$ to get

$$m\frac{d^2x_i}{d\tau^2} = m\frac{dt}{d\tau}\frac{d}{dt}\left(\frac{dx_i}{d\tau}\right) = e\frac{dx^0}{d\tau}\left(\frac{\partial A_0}{\partial x^i} - \frac{\partial A_i}{\partial x^0}\right) + e\frac{dx^j}{d\tau}\left(\frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j}\right)$$
$$= e\frac{dx^\nu}{d\tau}\left(\frac{\partial A_\nu}{\partial x^i} - \frac{\partial A_i}{\partial x^\nu}\right). \tag{98}$$

Observing it we find there's a free downstairs spatial index $i$ on each side. If we can replace it by the spacetime index $\mu$, we will have a covariant equation,

$$m\frac{d^2x_\mu}{d\tau^2} = e\frac{dx^\nu}{d\tau}\left(\frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}\right). \tag{99}$$

Actually doing this is allowable. By making sure at the beginning that the action was a scalar, we guaranteed that our EOM would be Lorentz invariant.

*If the equations are Lorentz invariant, and the three space components of a certain 4-vector are equal to the three space components of some other 4-vector, then we know automatically that their zeroth components (the time components) must also match.*

Yet, we still want to explore the meaning of the 0th component of the equation, we put it here,

$$m\frac{d^2x_0}{d\tau^2} = e\frac{dx^\nu}{d\tau}\left(\frac{\partial A_\nu}{\partial x^0} - \frac{\partial A_0}{\partial x^\nu}\right)$$
$$= e\frac{dx^0}{d\tau}\left(\frac{\partial A_0}{\partial x^0} - \frac{\partial A_0}{\partial x^0}\right) + e\frac{dx^j}{d\tau}\left(\frac{\partial A_j}{\partial x^0} - \frac{\partial A_0}{\partial x^j}\right) \tag{100}$$
$$= e\frac{dx^j}{d\tau}\left(\frac{\partial A_j}{\partial x^0} - \frac{\partial A_0}{\partial x^j}\right),$$

restoring the $\frac{d\tau}{dt}$ factor we rewrite it as,

$$\frac{d}{dt}(mU_0) = e\frac{dx^j}{dt}\left(\frac{\partial A_j}{\partial x^0} - \frac{\partial A_0}{\partial x^j}\right), \tag{101}$$

where the LHS is the time derivative of such a quantity

$$mU_0 = -mU^0 = -m\frac{dt}{d\tau} = -\frac{m}{\sqrt{1-v^2}} \sim -\frac{mc^2}{\sqrt{1-v^2/c^2}}, \tag{102}$$

which represents minus kinetic energy. The RHS of the 0-component EOM is nothing but $-e\vec{E}\cdot\vec{v}$, or $-\vec{F}_{elec}\cdot\vec{v}$, the minus work power. To conclude, the 0th-component EOM can be reduced as,

$$\frac{dE_{kin}}{dt} = \vec{F}_{elec}\cdot\vec{v}, \tag{103}$$

which is just the energy balance equation, expressing the fact that the change of kinetic energy is the work done on a system. Note here magnetic force doesn't contribute, because magnetic force always does no work.

To further simplify the EOM, we introduce an anti-symmetric tensor,

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu \equiv \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}. \tag{104}$$

It is easy to check that

$$\begin{cases} E_i = F_{i0} \\ B_i = \epsilon_i^{\ jk}F_{jk} \end{cases} \tag{105}$$

15

then the manifestly Lorentz invariant equation has the form,

$$m\mathbf{A}_\mu = e\,F_{\mu\nu}\,U^\nu, \quad \text{or} \quad m\mathbf{A}^\mu = e\,F^\mu{}_\nu\,U^\nu, \tag{106}$$

where $\mathbf{A}_\mu$ is the covariant 4-acceleration, not to confuse with the vector potential $A_\mu$.

Here we see the fact that $\vec{E}$ and $\vec{B}$ are frame-dependent quantities, which are just components of the tensor $F_{\mu\nu}$. The transformation rule for $F_{\mu\nu}$ is,

$$(F^{\mu\nu})' = L^\mu{}_\rho L^\nu{}_\sigma F^{\rho\sigma}, \tag{107}$$

where $L^\mu{}_\rho$ is the Lorentz transformation matrix (supposing two frames moving relatively along $x$-axis),

$$(L^\mu{}_\rho) = \begin{pmatrix} \frac{1}{\sqrt{1-v^2}} & \frac{-v}{\sqrt{1-v^2}} & 0 & 0 \\ \frac{-v}{\sqrt{1-v^2}} & \frac{1}{\sqrt{1-v^2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{108}$$

Then we will find that,

$$(E')^y = (F')^{0y} = L^0{}_x L^y{}_y F^{xy} = \frac{-vB_z}{\sqrt{1-v^2}}, \tag{109}$$

which indicates the Lorentz force of a moving particle along $x$-axis in the $B_z$ field is just the electric force exerted on the particle in its own rest frame with the $B_z$ field moving along the negative $x$-axis. (magnitude is to be matched.) (Einstein's paper)

## 2. Gauge invariance

Suppose we make a change to the vector potential by adding a total derivative of an arbitrary scalar field,

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu S(x), \tag{110}$$

where $x$ represents all 4 spacetime components. The new interaction action to the particle would become

$$\mathcal{A}'_{int} = \int_a^b A'\,\mathrm{d}x^\mu = \int_a^b A_\mu\,\mathrm{d}x^\mu + \int_a^b \partial_\mu S(x)\,\mathrm{d}x^\mu = \mathcal{A}_{int} + S(b) - S(a). \tag{111}$$

When we wiggle the vector field and the trajectory, the end points $a$ and $b$ are kept fixed, so $S(b) - S(a)$ remains a constant, meaning that if $\mathcal{A}$ is minimized, $\mathcal{A}'$ is also minimized, thus the change does not have any effect on the EOM of the particle. This change is called a *gauge transformation*.

Further if we see the new electro-magnetic tensor with this change, we'll find it's also invariant,

$$F'_{\mu\nu} = \partial_\mu(A_\nu + \partial_\nu S) - \partial_\nu(A_\mu + \partial_\mu S) = \partial_\mu A_\nu - \partial_\nu A_\mu + \partial_\mu\partial_\nu S - \partial_\nu\partial_\mu S = F_{\mu\nu}. \tag{112}$$

As an interesting special case, if $A_\mu$ itself is a total derivative of some scalar field, i.e, $A_\mu(x) = \partial_\mu S(x)$, there would be not any effect on the particle, and we see that the $F_{\mu\nu} = 0$, no electro-magnetic field at all.

Since such a change to $A_\mu$ doesn't change any physics, $A_\mu$ is not a measurable physical quantity, it's just an auxiliary field, which has certain freedom called gauge choice. With this freedom we could simplify equations in some cases. Choosing different gauges can illustrate different properties of a theory, and by choosing all possible choices we know all properties of a theory.

Here we recall that when we construct the scalar for the Lagrangian, we didn't use $\int A^\mu A_\mu\,\mathrm{d}\tau$, because it violates gauge invariant,

$$A'^\mu A'_\mu = (A^\mu + \partial^\mu S)(A_\mu + \partial_\mu S) = A^\mu A_\mu + 2A_\mu\partial^\mu S + \partial^\mu S \partial_\mu S \neq A^\mu A_\mu. \tag{113}$$

*3.Bianchi identity — the first half of Maxwell's equations*

An identity is a mathematical fact that follows from a definition. Due to the fact that $\vec{E}$ and $\vec{B}$ are defined as,

$$\begin{cases} \vec{E} = \vec{\nabla} A_0 - \dfrac{\partial \vec{A}}{\partial t} \\ \vec{B} = \vec{\nabla} \times \vec{A} \end{cases}, \tag{114}$$

it's easy to verify with vector analysis that

$$\begin{cases} \vec{\nabla} \times \vec{E} + \dfrac{\partial \vec{B}}{\partial t} = 0 \\ \vec{\nabla} \cdot \vec{B} = 0 \end{cases}, \tag{115}$$

which are called homogenous Maxwell equations. In terms of electro-magnetic tensor, it's rephrased as,

$$\partial_\sigma F_{\mu\nu} + \partial_\mu F_{\nu\sigma} + \partial_\nu F_{\sigma\mu} = 0, \tag{116}$$

or

$$\partial_{[\sigma} F_{\mu\nu]} = 0, \tag{117}$$

with its components containing 1-time coordinate representing the 1st equation, and its all-spatial components representing the 2nd. This is the so-called *Bianchi identity.*

To conclude this subsection, when a vector potential field $A_\mu$ is given, we could define $\vec{E}$ and $\vec{B}$, we can derive the EOM, and we naturally have the Bianchi identity which are homogenous Maxwell equations. Meanwhile, all of the above can parallel be expressed in terms of the electro-magnetic tensor $F_{\mu\nu}$. Furthermore, with the principle of gauge invariance, we exclude some possibilities of the scalar Lagrangian .

## 3.2   The Lagrangian of the Maxwell's equations and the effect by a charged particle

Now we are ready to construct the theory of the free EM field, on the basis of the 4 principles:
**Least action**, **Locality**, **Lorentz invariance**, and **Gauge invariance**.

*1.EOM of the e-m field — the second half of free Maxwell equations*

We can use the components $F_{\mu\nu}$ in any way we like without worrying about gauge invariance. And locality is also guaranteed since $F_{\mu\nu}$ contains first derivatives of the vector potential, $\partial_\mu A_\nu$. The only thing we have to do is to make out a Lagrangian which is a scalar.

The simplest scalar out of $F_{\mu\nu}$ we can think of is to contract with itself, as $F_\mu{}^\mu$; but we know that this is a trivial quantity since $F_\mu{}^\mu = 0$. Then, we could try something that's non-linear, which would be the quadratic term $F_{\mu\nu} F^{\mu\nu}$. First, we are going to figure out what it is. We expand it as,

$$\begin{aligned} F_{\mu\nu} F^{\mu\nu} &= F_{i0} F^{i0} + F_{0i} F^{0i} + F_{jk} F^{jk} + F_{kj} F^{kj} \\ &= 2(F_{i0} F^{i0} + F_{jk} F^{jk}) \\ &= -2(\vec{E}^2 - \vec{B}^2). \end{aligned} \tag{118}$$

We find the familiar result in classical electrodynamics that $\vec{E}^2 - \vec{B}^2$ is a Lorentz invariant quantity.
By convention, we choose the Lagrangian to be

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \ \left( = \frac{1}{2}(\vec{E}^2 - \vec{B}^2) \right). \tag{119}$$

The Euler-Lagrangian equation is given by

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} - \frac{\partial \mathcal{L}}{\partial A_\nu} = 0. \tag{120}$$

To proceed, we rephrase the Lagrangian as,

$$\mathcal{L} = -\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu). \tag{121}$$

Take derivatives of each four terms with respect to $\partial_\mu A_\nu$, we get

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} = -\frac{1}{4}[(\partial^\mu A^\nu - \partial^\nu A^\mu) - (\partial^\nu A^\mu - \partial^\mu A^\nu) + (\partial^\mu A^\nu - \partial^\nu A^\mu) - (\partial^\nu A^\mu - \partial^\mu A^\nu)] = -F^{\mu\nu}, \tag{122}$$

and the 2nd term of the E-L equation is zero, so we have a neat form of the EOM of the e-m field as,

$$\partial_\mu F^{\mu\nu} = 0. \tag{123}$$

To recover the second half of Maxwell equation, we extract its time-contained and all-spatial components,

$$\begin{cases} \partial_0 F^{00} + \partial_i F^{i0} = 0 \implies \vec{\nabla} \cdot \vec{E} = 0 \\ \partial_0 F^{0j} + \partial_i F^{ij} = 0 \implies \vec{\nabla} \times \vec{B} - \dfrac{\partial \vec{E}}{\partial t} = 0 \end{cases} \tag{124}$$

To conclude the free Maxwell equations, we list the tensor form,

$$\begin{cases} \partial_{[\sigma} F_{\mu\nu]} = 0 \\ \partial_\mu F^{\mu\nu} = 0, \end{cases} \tag{125}$$

and the vector form,

$$\begin{cases} \vec{\nabla} \times \vec{E} = -\dfrac{\partial \vec{B}}{\partial t} \\ \vec{\nabla} \cdot \vec{B} = 0 \\ \vec{\nabla} \cdot \vec{E} = 0 \\ \vec{\nabla} \times \vec{B} = \dfrac{\partial \vec{E}}{\partial t}. \end{cases} \tag{126}$$

At this point, it's time to have an interlude to present the plane wave solution to the free Maxwell equations.(to be updated)

2.*The charged particle's effect on the e-m field — Lagrangian with a source term*

The charge density and current can compose a 4-component current as, $J^\mu = (\rho, \vec{j})$ (transformation property to be updated). The local conservation law gives the continuity equation,

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0, \quad \text{or} \quad \partial_\mu J^\mu = 0. \tag{127}$$

The simplest scalar we can constrict out of the charge source and the vector field is $J^\mu A_\mu$. Now we check its gauge invariance, by changing $A_\mu \to A'_\mu = A_\mu + \frac{\partial S}{\partial x^\mu}$, and integrate it over the spacetime volume to get the interaction action,

$$\mathcal{A}'_{int} = \int J^\mu A'_\mu \, \mathrm{d}^4 x = \int J^\mu A_\mu \, \mathrm{d}^4 x + \int J^\mu \frac{\partial S}{\partial x^\mu} \, \mathrm{d}^4 x = \mathcal{A}_{int} + J^\mu S|_{boundary} - \int S \, \partial_\mu J^\mu \, \mathrm{d}^4 x = \mathcal{A}_{int}. \tag{128}$$

18

The boundary term vanishes because the current is localized and it's zero at infinity; the second new term vanishes due to the continuity equation. Thus the scalar we made obeys gauge invariance.

We may write the Lagrangian as,

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - J^\mu A_\mu, \tag{129}$$

where the minus sign is due to convention. When taking it into the E-L equation, the only change is the non-zero term,

$$\frac{\partial\mathcal{L}}{\partial A_\nu} = -J^\nu. \tag{130}$$

By combining it with the other part of the E-L equation, we get the EOM of the e-m field with a source term,

$$\partial_\mu F^{\mu\nu} = J^\nu. \tag{131}$$

Note that in other papers there may be a minus before the 4-current term, which may be due three possible conventions: the sign of $\eta_{\mu\nu}$, the definition of $E_i$, and the sign in the Lagrangian. But one can always make it consistent. Then we could recover the second half of Maxwell equation with the source term, which is called the *inhomogeneous Maxwell equations*,

$$\begin{cases} \vec{\nabla}\cdot\vec{E} = \rho \\ \vec{\nabla}\times\vec{B} - \dfrac{\partial\vec{E}}{\partial t} = \vec{j}. \end{cases} \tag{132}$$

There's minor thing to mention, if someone happened to, (or probably truly in history,) discover the above equations without knowing the continuity equation, he could also derive it naturally from the inhomogeneous equations,

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\vec{j} = \frac{\partial}{\partial t}(\vec{\nabla}\cdot\vec{E}) + \vec{\nabla}\times(\vec{\nabla}\times\vec{B} - \frac{\partial\vec{E}}{\partial t}) = 0, \tag{133}$$

or simply,

$$\partial_\nu J^\nu = \partial_\nu\partial_\mu F^{\mu\nu} = \partial_{(\nu}\partial_{\mu)}F^{[\mu\nu]} = 0, \tag{134}$$

which says the Maxwell equations and the continuity equation are consistent. Also, gauge invariance is consistent with the above two.

## 3.3 The energy and momentum of the electro-magnetic field

To apply the idea of canonical momentum and energy, we treat the four components of the vector potential $A_\mu$ as four independent fields (,which are not scalar field but coordinate-dependent).

Note that we can take advantage of the gauge choice to make the vector potential simpler, without affecting any physics. Our choice is to make the zeroth component of $A_\mu$ vanish, in other words, we'd like to choose an $S$ such that

$$A_0 + \frac{\partial S}{\partial t} = 0 \quad\text{or}\quad \frac{\partial S}{\partial t} = -A_0, \tag{135}$$

which means our new vector potential will be

$$A'_\mu = A_\mu + \frac{\partial S}{\partial t} = (0, A'_i). \tag{136}$$

This is called *gauge fixing*. With this gauge, our electro and magnetic field are expressed as,

$$\begin{cases} \vec{E} = \vec{\nabla}A_0 - \dfrac{\partial\vec{A}}{\partial t} = -\dfrac{\partial\vec{A}}{\partial t} \\ \vec{B} = \vec{\nabla}\times\vec{A}. \end{cases} \tag{137}$$

Thus the Lagrangian density will be

$$\mathcal{L} = \frac{1}{2}(\vec{E}^2 - \vec{B}^2) = \frac{1}{2}\left(\frac{\partial\vec{A}}{\partial t}\right)^2 - \left(\vec{\nabla}\times\vec{A}\right)^2, \tag{138}$$

which has the structure of "squares of time derivatives minus squares of space derivatives", much similar with the case in free scalar field,

$$\mathcal{L} = \frac{1}{2}\left(\frac{\partial \phi}{\partial t}\right)^2 - \frac{1}{2}\left(\frac{\partial \phi}{\partial x}\right)^2. \tag{139}$$

The canonical momentum conjugate to $A_i$ is given by

$$\vec{\pi}_A = \frac{\partial \mathcal{L}}{\partial(\partial_t \vec{A})} = \frac{\partial \vec{A}}{\partial t} = -\vec{E}. \tag{140}$$

By observing the Lagrangian density, we see that it has the form of a "kinetic energy" term that depends on squares of time derivatives, minus a "potential energy" term that has no time derivatives at all. We may immediately write the Hamiltonian density as

$$\mathcal{H} = \frac{1}{2}(\vec{E}^2 + \vec{B}^2), \tag{141}$$

which is the electromagnetic energy. We can see that it's positive and both $\vec{E}$ and $\vec{B}$ contribute same amounts to the total energy.

Using the Noether Theorem we constructed in scalar field theory, the $x$ component of the momentum carried by the e-m field is given by

$$P_x = \int E_i \frac{\partial A_i}{\partial x} \, \mathrm{d}^3 x. \tag{142}$$

It is straightforward to generalize in any other directions,

$$P_i = \int E_j \frac{\partial A_j}{\partial x^i} \, \mathrm{d}^3 x. \tag{143}$$

Now we have the momentum density in the form of $E_j \partial_i A_j$, which is however gauge-dependent. We make it a gauge-independent form by adding a term, which actually turns to be zero,

$$-\int E_j \frac{\partial A_i}{\partial x^j} \, \mathrm{d}^3 x = \int \frac{\partial E_j}{\partial x^j} A_i \, \mathrm{d}^3 x = \int (\vec{\nabla} \cdot \vec{E}) A_i \, \mathrm{d}^3 x = 0, \tag{144}$$

where we assume the fields go to zero beyond a certain distance, therefore there are no boundary terms. Thus we rewrite the momentum as

$$P_i = \int E_j \left(\frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j}\right) \, \mathrm{d}^3 x, \tag{145}$$

which is nothing but

$$\vec{P} = \int \vec{E} \times \vec{B} \, \mathrm{d}^3 x. \tag{146}$$

The momentum density is called the *Poynting vector*, pointing to the direction of the momentum,

$$\vec{S} := \vec{E} \times \vec{B}. \tag{147}$$

## 3.4   Energy and momentum in four dimensions

Recall that in the particle mechanics in the 1st chapter, the energy and the momentum are not Lorentz invariant, but they form a Lorentz covariant 4-vector $P^\mu$. Global conservation law tells us that the each of the four components of the vector is time-independent,

$$\frac{\mathrm{d}P^\mu}{\mathrm{d}t} = 0. \tag{148}$$

This suggests that each component has a local conservation law, there's a density and a flux correspondingly, which satisfy the continuity equation.

For the 0th component, a change of energy within some region of space is always accompanied by a flow of energy through the boundaries of the region. What we're going to do is to find the energy flux corresponding to the energy density. First we write the continuity equation below,

$$\frac{\partial \mathcal{H}}{\partial t} + \vec{\nabla} \cdot \vec{J}_{\mathcal{H}} = 0. \tag{149}$$

We rearrange the equation and plug in the expression of Hamiltonian density to give,

$$\begin{aligned} \vec{\nabla} \cdot \vec{J}_{\mathcal{H}} = -\frac{\partial \mathcal{H}}{\partial t} &= -\frac{1}{2}\frac{\partial}{\partial t}(\vec{E}^2 + \vec{B}^2) \\ &= -(\vec{E} \cdot \dot{\vec{E}} + \vec{B} \cdot \dot{\vec{B}}) \\ &= \vec{B} \cdot (\vec{\nabla} \times \vec{E}) - \vec{E} \cdot (\vec{\nabla} \times \vec{B}) \\ &= \vec{\nabla} \cdot (\vec{E} \times \vec{B}), \end{aligned} \tag{150}$$

where we have make use of the Maxwell equation. So it is natural to define the energy flux as,

$$\vec{J}_{\mathcal{H}} := \vec{E} \times \vec{B}, \tag{151}$$

which is nothing but the Poynting vector.

*The Poynting vector has two meanings: We can think of it as either an <u>energy flow</u> or a <u>momentum density</u>.*

Now we develop a new set of symbol, $T^{\mu\nu}$, called the energy-momentum temsor. *The first index tells us which of the four quantities the element refers; the second index tells us whether we're talking about a density or flow.*

For example, the following equation

$$P^\mu = \int T^{\mu 0} \, \mathrm{d}^3 x \tag{152}$$

combines the fact that the total energy $P^0$ is the integral of the energy density $T^{00}$ over space, and the total momentum $P^i$ is the integral of the momentum density $T^{i0}$ over space. And the continuity equation of energy density we constructed above can be re-expressed as

$$\frac{\partial \mathcal{H}}{\partial t} + \vec{\nabla} \cdot \vec{J}_{\mathcal{H}} = 0 \quad \Longrightarrow \quad \partial_0 T^{00} + \partial_i T^{0i} = 0, \quad \text{or} \quad \partial_\nu T^{0\nu} = 0. \tag{153}$$

Still, we have the continuity equation of the momentum density and its flux,

$$\partial_0 T^{i0} + \partial_j T^{ij} = 0, \quad \text{or} \quad \partial_\nu T^{i\nu} = 0. \tag{154}$$

Combining them we have

$$\partial_\nu T^{\mu\nu} = 0. \tag{155}$$

This is the unified relativistic form of the local conservation law of all the four components.

To proceed, we construct $T^{\mu\nu}$ of the e-m fields by taking considerations on symmetries. The energy-momentum tensor is constituted of densities and fluxes of energy and momentum, which are physical quantities, so gauge invariance tells us they should depend only on the observable fields $\vec{E}$ and $\vec{B}$, not on the vector potential $A_\mu$. The Lorentz invariance tells us it's a rank two tensor.

Now that we have already know what $T^{00}$ is, we can make use of this significant information,

$$T^{00} = \frac{1}{2}(E^2 + B^2). \tag{156}$$

It tells us two pieces of information.

First, from its form we are suggested that $T^{\mu\nu}$ is quadratic in the components of the field tensor; in other words, it is formed from products of two components of $F^{\mu\nu}$. There are only two ways to construct a tensor from the product of two copies of $F^{\mu\nu}$: contract itself once, and twice. So it must be a sum of two terms of the form

$$T^{\mu\nu} = a\, F^{\mu\sigma} F^{\nu}{}_{\sigma} + b\, \eta^{\mu\nu} F^{\sigma\tau} F_{\sigma\tau}, \tag{157}$$

where $a$ and $b$ are numerical constants to be determined.

Second, the form of $T^{00}$ just tells us what $a$ and $b$ is. Simply by plugging it in, we have

$$\begin{aligned} \frac{1}{2}(E^2 + B^2) = T^{00} &= a\, F^{0\sigma} F^{0}{}_{\sigma} + b\, \eta^{00} F^{\sigma\tau} F_{\sigma\tau} \\ &= a\, E^2 + 2b\,(E^2 - B^2), \end{aligned} \tag{158}$$

and find that $a = 1$ and $b = -1/4$. Thus our energy-momentum tensor is constructed as

$$T^{\mu\nu} = F^{\mu\sigma} F^{\nu}{}_{\sigma} - \frac{1}{4}\eta^{\mu\nu} F^{\sigma\tau} F_{\sigma\tau}. \tag{159}$$

We can see that it's a symmetric tensor, and we can compute all its various components through this equation. We display the the form of the tensor matrix below,

$$(T^{\mu\nu}) = \begin{pmatrix} \frac{1}{2}(E^2 + B^2) & S_x & S_y & S_z \\ S_x & -\sigma_{xx} & -\sigma_{xy} & -\sigma_{xz} \\ S_y & -\sigma_{yx} & -\sigma_{yy} & -\sigma_{zz} \\ S_z & -\sigma_{zx} & -\sigma_{zy} & -\sigma_{zz} \end{pmatrix}, \tag{160}$$

where $S_i$ are components of the Poynting vector, and $\sigma_{ij}$ are components of the tensor $T^{ij}$ called the *electromagnetic stress tensor*.

However, $T^{\mu\nu}$ does not directly appear in the Maxwell equations for $\vec{E}$ and $\vec{B}$, it's only in the theory of General Relatvity that energy-momentum tensor take its rightful role as the source of the gravitational field.

Finally, what we left is an interesting subtlety. If we restore the factor of the speed of light $c$, we would find that the energy flux and momentum density differ by a factor of $c^2$.

*For an electromagnetic wave, the density of momentum is equal to the flux of energy divided by the speed of light squared. Both are proportional to the Poynting vector.*

Now we can see why on the one had sunlight warms us when it's absorbed, but on the other hand it exerts such a feeble force. The reason is that the energy density is $c^2$ times the momentum density, and $c^2$ is a very big number. And this is a very good way to memorize the relationship between the energy flux and momentum density.

Figures to be added: synchronization by light, scalar field and particle trajectory, vector field and particle trajectory.