

PHYS 323 (Advanced Electrodynamics II) Notes

Rio Weil

This document was typeset on April 8, 2025

Introduction:

This is a set of lecture notes taken from UChicago's PHYS 323 (Advanced Electrodynamics II), taught by David Kutasov. Topics covered include...

Contents

1	Relativistic Kinematics	2
1.1	Symmetries of Classical Mechanics	2
1.2	Symmetry Group, Invariants of Classical Mechanics	3
1.3	Generalization to Fields	3
1.4	Symmetry actions in Maxwell theory	4
2	Lorentz Transformations	6
2.1	Review of last class	6
2.2	Spacetime Intervals and Poincare Symmetry	7
2.3	Solutions to the Wave Equation	7
2.4	Lorentz Transformations - Rotations and Boosts	8
2.5	Doppler Effect	10
2.6	Teaser - Relativistic Formulation of Maxwell Theory	11
3	Relativistic Formulation of a Free Particle	11
3.1	Parametrization	12
3.2	Parametrizing Worldlines and Time Dilation	13
3.3	Lagrangian for Relativistic Particle	14
3.4	Relativistic energy and momentum	15
4	Relativistic Formulation of Maxwell Theory	16
4.1	Maxwell under Rotations	16
4.2	Gauge Fields	17
4.3	Field Strength	17
4.4	Lorentz Covariant Maxwell Equations	18
4.5	Lagrangian Formulation	20
5	Lagrangian for EM field	21
5.1	From particles to fields	21
5.2	EM field Lagrangian - no sources	22
5.3	EM field Lagrangian - with sources	23
5.4	The stress-energy tensor	24

1 Relativistic Kinematics

Course logistics - see syllabus.

1.1 Symmetries of Classical Mechanics

We will start with Ch. 8/9 of Wald's book, and discuss relativistic kinematics.

The basic fact discovered after the formulation of Maxwell's theory of electromagnetism is that there is a clash between Maxwell's theory (Lorentz invariance) and the symmetries of classical mechanics (Galilean invariance). We will start by reviewing these symmetries, and then discuss how to promote these symmetries to field theory.

Consider the following (relatively general) class of problems. We have N particles, labelled by coordinates $\mathbf{x}_1(t), \mathbf{x}_2(t), \dots, \mathbf{x}_N(t)$, which have pairwise interactions that only depend on the particle separation $V_{ij}(|\mathbf{x}_i - \mathbf{x}_j|)$. In classical mechanics (as you learned two quarters ago), we use a Lagrangian to describe the system:

$$\mathcal{L} = \frac{1}{2} \sum_{i=1}^N m_i \dot{\mathbf{x}}_i^2 - \sum_{i < j} V_{ij}(|\mathbf{x}_i - \mathbf{x}_j|) \quad (1.1)$$

and we have the Euler-Lagrange equation(s) which are the equations of motion of the system:

$$m_i \ddot{\mathbf{x}}_i = -\nabla_i \sum_{j \neq i} V_{ij}(|\mathbf{x}_i - \mathbf{x}_j|) \quad (1.2)$$

where we work in 3-dimensions, so $\mathbf{x}_i = (x_i, y_i, z_i)$ and $\nabla_i = (\frac{\partial}{\partial x_i}, \frac{\partial}{\partial y_i}, \frac{\partial}{\partial z_i})$. This theory has a couple symmetries (i.e. if we have a solution to the equations of motion, under some transformation the equations are still solutions. This manifests as a transformation that leaves \mathcal{L} invariant):

1. Translational invariance:

$$\mathbf{x}_i \rightarrow \mathbf{x}_i + \mathbf{a} \quad (1.3)$$

The kinetic term does not change because the \mathbf{a} is not time-dependent (and hence is killed by the time derivative) and the potential term does not change because the potential only depends on the difference of the positions (and hence the \mathbf{a} s cancel). Physically, we can take our entire system and shift it over by \mathbf{a} and nothing changes. Why are we wasting our time with this nonsense? Thanks to Noether, we know that this symmetry that looks nontrivial implies *momentum conservation*.

2. Time translation invariance:

$$t \rightarrow t + t_0 \quad (1.4)$$

Again, we see that this is a symmetry of \mathcal{L} because there is no explicit time-dependence in the Lagrangian. Noether tells us that this implies *energy conservation*.

3. Rotational invariance:

$$\mathbf{x}_i \rightarrow R\mathbf{x}_i \quad (1.5)$$

for a rotation matrix R , i.e. that satisfying $RR^t = R^tR = \mathbb{I}$. The origin of this symmetry is that there is no preferred direction in space. Noether tells us that this implies *angular momentum conservation*.

4. Boost invariance:

$$\mathbf{x}_i \rightarrow \mathbf{x}_i - \mathbf{v}t \quad (1.6)$$

The previous 3 symmetries seem so robust that it seems like things cannot go wrong. Indeed, it is this fourth symmetry that must be modified when we consider the Maxwell theory. What this symmetry means is if we take the entire system and put it on a train going on a constant velocity relative to us, the physics of the system look the same (e.g. if we take the solar system and put it on a galactic train, the motion of the planets are left invariant).

What about other transformations that depend on time? For example a transformation that depends on t^2 ? This would not be a symmetry - indeed we can feel if we are in an airplane and we are accelerating. This leads to the notion of an inertial frame. If we have some frame in which Newton's equations are valid, then any other frame that differs from our original frame via a boost will also be inertial. But accelerating frames are not inertial.

1.2 Symmetry Group, Invariants of Classical Mechanics

The symmetry group describing the symmetries of classical mechanics is the Galilean group. It is 10-dimensional; there are:

- 3 spatial translation symmetries
- 1 time translation symmetry
- 3 rotation symmetries
- 3 boost symmetries

Side note - people like to divide the Galilean group into the homogenous Galilean group (rotations and boosts - which map coordinates under linear combinations of coordinates) and the inhomogenous Galilean group (space and time translations).

Consider two points in space + time - $(t_1, \mathbf{x}_1), (t_2, \mathbf{x}_2)$ (say, what time your alarm clock went off and what time you got out of bed). What combination of these coordinates are invariants?:

- $\Delta t = t_2 - t_1$ is invariant under all of the above transformations.
- $\Delta x = |\mathbf{x}_2 - \mathbf{x}_1|$ - is not invariant under boosts, as under $\mathbf{x}' = \mathbf{x} - \mathbf{v}t$ we have:

$$|\mathbf{x}'_2 - \mathbf{x}'_1| = |(\mathbf{x}_2 - \mathbf{v}t_2) - (\mathbf{x}_1 - \mathbf{v}t_1)| = |\mathbf{x}_2 - \mathbf{x}_1 - \mathbf{v}(t_2 - t_1)| \neq |\mathbf{x}_2 - \mathbf{x}_1| \quad (1.7)$$

However, if $t_1 = t_2$ then it is indeed an invariant.

1.3 Generalization to Fields

We now think about how to generalize these symmetries to fields. Consider the wave equation, which you saw in PHYS 322:

$$\square \phi(\mathbf{x}, t) := \left(-\frac{1}{c^2} \partial_t^2 + \nabla^2\right) \phi(\mathbf{x}, t) = 0 \quad (1.8)$$

Actually, we don't need to go to Maxwell/electromagnetic theory to see this; the wave equation already appears in classical mechanics, e.g. to describe waves on guitar strings, water etc.

The solutions to the wave equation are:

$$\phi(\mathbf{x}, t) = C e^{i\mathbf{k} \cdot \mathbf{x} - \omega t}, \quad \omega = c|\mathbf{k}| \quad (1.9)$$

where c is the speed of propagation in the medium and C some arbitrary multiplicative constant. Let's consider again the symmetries of our discrete classical theory and see what the effects of the transformations are.

- Translations: if we take $\mathbf{x} \rightarrow \mathbf{x} + \mathbf{a}, t \rightarrow t + t_0$, then we have:

$$\phi(\mathbf{x}, t) \rightarrow \phi(\mathbf{x} + \mathbf{a}, t + t_0) = C e^{i\mathbf{k} \cdot (\mathbf{x} + \mathbf{a}) - i\omega(t + t_0)} = C e^{i\mathbf{k} \cdot \mathbf{a} - i\omega t_0} e^{i\mathbf{k} \cdot \mathbf{x} - i\omega t} \quad (1.10)$$

The translation simply yields a multiplicative constant/phase which is yet another solution to the wave equation.

- With boosts the story is different. Taking $\mathbf{x} \rightarrow \mathbf{x} - \mathbf{v}t$, we have:

$$\phi(\mathbf{x}, t) \rightarrow \phi(\mathbf{x} - \mathbf{v}t, t) = e^{i\mathbf{k} \cdot \mathbf{x} - i(\omega + \mathbf{k} \cdot \mathbf{v})t} \quad (1.11)$$

so we have a new ω' :

$$\omega' = \omega + \mathbf{k} \cdot \mathbf{v} \quad (1.12)$$

but we require that $\omega' = c|\mathbf{k}|$, and indeed this does not hold generally:

$$\omega + \mathbf{k} \cdot \mathbf{v} \neq c|\mathbf{k}| \quad (1.13)$$

unless the $\mathbf{k} \cdot \mathbf{v}$ term vanishes, i.e. the boost is in a perpendicular direction to the direction of propagation of the wave. In other words, while monochromatic waves do map to monochromatic waves, they do not preserve the dispersion relation.

At its face this seems to be a paradox. How is this consistent with the fact that boosts are a symmetry of our classical theory? There was indeed something we hid here - for the wave equation to be valid, the medium needs to be at rest - Eq. (1.8) assumes that the medium is at rest, for example (though we could generalize to the case of a moving medium)

Let us restrict ourselves to 1-d for moment and think about this further. Then, the solutions to the wave equation become:

$$\phi(x, t) = e^{ikx - i\omega t} \quad (1.14)$$

with boosted solutions:

$$\phi'(x, t) = \phi(x + vt, t) = e^{ikx - i(\omega + kv)t} \quad (1.15)$$

so in the boosted frame, it looks like the waves are travelling at speed:

$$V = \frac{\omega + kv}{k} = \frac{\omega}{k} + v = c + v \quad (1.16)$$

so we can see the law of composition of velocities emerge.

1.4 Symmetry actions in Maxwell theory

When Maxwell theory was proposed, people that believed that the same discussion holds for electromagnetic waves. As a result they believed two things:

1. There was an “Ether”, a frame where the light waves were at rest, analogous to the frame in the previous example where the water waves were at rest.
2. The speed of light depends on the frame of reference.

Michelson wasted 2 decades of his life trying to measure this effect (at least he got a UChicago building named after him!) and after this sequence of failures, people realized that our understanding needed to be amended. Indeed, the discussion of classical waves we just went through is *not* a feature of electromagnetism. The actual situation is that Maxwell's equations hold in all inertial frames, and the symmetry group of nature is still generated by translations, rotations, and boosts, but the action of boosts is different from the classical case.

As a first step towards generalizing the action, let us remind ourselves how rotations act on monochromatic plane waves. As previously mentioned, we obtain the rotated coordinates via acting on our coordinates with a 3x3 rotation matrix:

$$\mathbf{x}' = R\mathbf{x} \quad (1.17)$$

Looking at $\mathbf{k} \cdot \mathbf{x}$:

$$\mathbf{k} \cdot \mathbf{x} = \begin{pmatrix} k_x & k_y & k_z \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (1.18)$$

If we replace \mathbf{x} with $\mathbf{x}' = R^t \mathbf{x}$ we have:

$$\begin{pmatrix} k_x & k_y & k_z \end{pmatrix} R^t \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \quad (1.19)$$

We then have the rotated wavenumber:

$$\begin{pmatrix} k'_x & k'_y & k'_z \end{pmatrix} = \begin{pmatrix} k_x & k_y & k_z \end{pmatrix} R^t \implies \begin{pmatrix} k'_x \\ k'_y \\ k'_z \end{pmatrix} = R^t \begin{pmatrix} k_x \\ k_y \\ k_z \end{pmatrix} \quad (1.20)$$

So - If $\phi(\mathbf{x}, t)$ is a solution to the wave equation, then $\phi'(\mathbf{x}, t)$ (i.e. the solution where each \mathbf{k} is mapped to $\mathbf{k}' = R^t \mathbf{k}$) is also a (different) solution of that equation, which satisfies $\phi'(\mathbf{x}', t) = \phi(\mathbf{x}, t)$. This is an exercise which you can go through at home by taking a general solution and fourier decomposing it and applying the above wavenumber argument to each fourier component.

Now, let's return back to boosts - if we want the wave equation to be invariant under boosts (which it was not using the original definition of boosts), the best way to proceed is the following. We go from a 3-d story to a 3+1-d story:

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z) \quad (1.21)$$

Now, we want to generalize the rotations in 3-d space to the full 3+1-d spacetime.

In space, we have the line element:

$$ds^2 = dx^i dx^i = dx^2 + dy^2 + dz^2 \quad (1.22)$$

where the length of a trajectory is given by (in the 2-d case where $y = y(x)$ so $dy = y' dx$):

$$L = \int ds = \int dx \sqrt{1 + y'^2} \quad (1.23)$$

ds is rotationally invariant:

$$ds^2 = dx^i dx^i = dx'^i dx'^i \quad (1.24)$$

i.e. if you rotate your frame the distance between two points does not change. This is the thing that we generalize to 3+1d. We define the line element:

$$ds^2 = dx^\mu dx^\nu \eta_{\mu\nu} = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2 \quad (1.25)$$

with $\eta_{\mu\nu}$ the Minkowski metric:

$$\eta = \text{diag}(-1, 1, 1, 1). \quad (1.26)$$

How does this help? Let's look at the wave equation. ϕ originally was a function of \mathbf{x}, t , which we've repackaged into x^μ :

$$\square \phi(x^\mu) = 0, \quad \square = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} + \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^i} = -\partial_0^2 + \partial_i \partial_i = \frac{\partial}{\partial x^\mu} \frac{\partial}{\partial x^\mu} \eta^{\mu\nu} \quad (1.27)$$

An important note; previously, indices on coordinates were just an aesthetic choice. But here now one of the signs of the metric is different, so we need to take a little more care with where we put our indices. In particular, given x^μ we can define:

$$x_\mu = \eta_{\mu\nu} x^\nu \quad (1.28)$$

so:

$$x_0 = -x^0, x_i = x^i \quad (1.29)$$

Note that this implies that $\frac{\partial}{\partial x^\mu}$ actually carries a lower index. $x^\mu x_\mu$ is Lorentz invariant (note the Einstein summation convention/repeated indices are summed over), and taking its derivative:

$$\frac{\partial}{\partial x^\mu}(x^\nu x_\nu) = 2x_\mu \quad (1.30)$$

so since taking the derivative results in an object with a lower index, $\frac{\partial}{\partial x^\mu}$ indeed carries a lower index. This motivates the $\frac{\partial}{\partial x^\mu} \frac{\partial}{\partial x^\mu} \eta^{\mu\nu}$ we wrote in the D'Alembertian.

Note that fortunately, the metric with upper indices is identical to the metric with lower indices:

$$\eta_{00} = -1, \quad \eta_0^0 = +1 \implies \eta^{00} = -1 \quad (1.31)$$

This discussion was a bit fast, but we will continue it next class. It's recommended that you read the relevant sections in Wald to refresh yourself on all of this index manipulation.

2 Lorentz Transformations

2.1 Review of last class

Last time, we were discussing the question of how pre-relativistic spacetime symmetries can be promoted to be consistent with the Maxwell theory. We generalized from rotations; we have three space coordinates $(x^1, x^2, x^3) = (x, y, z)$ and the line element:

$$ds^2 = dx^2 + dy^2 + dz^2 = dx^i dx^i \quad (2.1)$$

where the summation over the repeated index i is implied. Rotations have the action on the coordinates:

$$x'^i = R^i_j x^j \quad (2.2)$$

with $R^t R = R R^t = \mathbb{I}$. The line element ds^2 is invariant under rotations:

$$ds'^2 = dx'^i dx'^i = dx^i dx^i = ds^2 \quad (2.3)$$

The idea is that given two points $\mathbf{x}_1, \mathbf{x}_2$, the squared distance:

$$D^2 = |\mathbf{x}_2 - \mathbf{x}_1|^2 \quad (2.4)$$

is invariant under rotations. For our purpose, it is also useful to recall that $\nabla^2 = \partial_i \partial_i$ is also invariant:

$$\nabla'^2 = \frac{\partial}{\partial x'^i} \frac{\partial}{\partial x'^i} = \nabla^2 \quad (2.5)$$

The idea was to now include time:

$$ds^2 = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2 = dx^\mu dx^\nu \eta_{\mu\nu} \quad (2.6)$$

with $x^0 = ct$, and $\mu, \nu = 0, 1, 2, 3$ and:

$$\eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1). \quad (2.7)$$

Note that we can write the line element equivalently as:

$$ds^2 = dx^\mu dx_\mu \quad (2.8)$$

where:

$$dx_\mu = \eta_{\mu\nu} dx^\nu \quad (2.9)$$

generally, we are able to lower any index using the spacetime metric (here a metric on $R^{1,3}$). With this definition, note:

$$dx_0 = -dx^0, \quad dx_i = dx^i \quad (2.10)$$

Note that greek indices run over all 4 spacetime components, while alphabetic indices (i, j, k) only run over spatial.

2.2 Spacetime Intervals and Poincare Symmetry

For two points in spacetime x^μ, y^μ , we can define the spacetime interval:

$$I = -(x^\mu - y^\mu)(x^\nu - y^\nu)\eta_{\mu\nu} = -(x^0 - y^0)^2 + |\mathbf{x} - \mathbf{y}|^2 \quad (2.11)$$

There are three cases:

- For $I > 0$, x, y are spacelike separated.
- For $I < 0$, x, y are timelike separated.
- For $I = 0$, x, y are lightlike separated.

I behaves like an inner product of sorts on spacetime, but it is not positive definite.

Note that ds^2 is invariant under transformations of the form:

$$x'^\mu = \Lambda^\mu_\nu x^\nu \quad (2.12)$$

It is very important to get the location of indices correct!¹

Λ satisfies:

$$\Lambda^\mu_\lambda \Lambda^\nu_\rho \eta_{\mu\nu} = \eta_{\lambda\rho} \quad (2.13)$$

2.3 Solutions to the Wave Equation

Recall the box operator, which we can now write compactly:

$$\square = -\frac{1}{c^2} \partial_t^2 + \nabla^2 = \partial_\mu \partial^\mu = \partial_\mu \partial_\nu \eta^{\mu\nu} \quad (2.14)$$

Like the Laplacian was invariant under rotations, the box operator is invariant under Lorentz transformations/ Λ s:

$$\square_x = \square_{x'} \quad (2.15)$$

What do we learn from this? If $\phi(x^\mu)$ is a solution of the wave equation:

$$\square_x \phi(x) = 0 \quad (2.16)$$

then $\phi'(x)$ is also a solution of the wave equation. ϕ' is defined by:

$$\phi'(x') = \phi(x) \quad (2.17)$$

An important example is plane waves:

$$\phi(x^\mu) = e^{ik_\mu x^\mu} = e^{ik^\mu x^\nu \eta_{\mu\nu}} \quad (2.18)$$

Note that if ϕ is a solution:

$$(x^\mu) = ik^\mu ik_\mu e^{ik^\mu x_\mu} = 0 \quad (2.19)$$

thus:

$$k^\mu k_\mu = 0 \implies (k^0)^2 = |\mathbf{k}|^2 \implies k^0 = |\mathbf{k}| \quad (2.20)$$

where we take the positive root. In the discussion we had on Tuesday, we quoted the familiar result from last quarter:

$$\phi(\mathbf{x}, t) = e^{-i\omega t + i\mathbf{k} \cdot \mathbf{x}} \quad (2.21)$$

¹If you wrote the LHS as x'_μ on the midterm, the grader is going to be in pain.

with $\omega = c|\mathbf{k}|$. This condition is exactly the same as that we write above:

$$\phi(\mathbf{x}, t) = e^{-i\omega t + i\mathbf{k} \cdot \mathbf{x}} = e^{-i\frac{\omega}{c}x^0 + i\mathbf{k} \cdot \mathbf{x}} = e^{-ik^0x^0 + i\mathbf{k} \cdot \mathbf{x}} \quad (2.22)$$

So if we substitute k^0 into the dispersion relation:

$$k^0 = |\mathbf{k}|. \quad (2.23)$$

For this $\phi(x)$, $\phi'(x)$ looks like:

$$e^{ik'_\mu x^\mu} \quad (2.24)$$

where:

$$k'_\mu = \Lambda_\mu^\nu k_\nu \quad (2.25)$$

Note that in order for this be a solution, we must have that the k s are lightlike, as they are for the original solution:

$$k'^\mu k'_\mu = 0 \quad (2.26)$$

But this is guaranteed by the fact that the Lorentz transformations preserve the metric (check it if you aren't convinced).

Last class we had the problem that applying a boost to a solution of the wave equation was not a solution - we have now fixed this.

2.4 Lorentz Transformations - Rotations and Boosts

There is a special class of Lorentz transformations where $\Lambda_0^0 = 1$ and $\Lambda_0^i = 0$, which look like:

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & R & \\ 0 & & & \end{pmatrix} \quad (2.27)$$

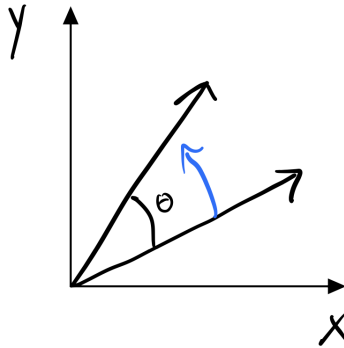
Note that for these:

$$x'^0 = \Lambda^0_\nu x^\nu = x^0 \quad (2.28)$$

$$x'^i = \Lambda^i_\nu x^\nu = R^i_j x^j \quad (2.29)$$

for $R^t R = R R^t = \mathbb{I}$, so rotations are just a subclass of Lorentz transformations. Given that rotations of solutions to the wave equation are also solution, this is not too surprising.

A couple examples; for a rotation about the z-axis by the angle θ :



we have:

$$R = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.30)$$

For a boost in the x' direction with velocity v , we have:

$$\Lambda_{\mu}^{\nu} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.31)$$

where the parameters obey (in order for Λ to preserve the metric):

$$\gamma^2(1 - \beta^2) = 1 \quad (2.32)$$

$\beta = \frac{v}{c} \in (-1, 1)$ as nothing travels faster than the speed of light. Rearranging for γ , we have the familiar expression:

$$\gamma = \frac{1}{\sqrt{1 - \beta^2}} \quad (2.33)$$

We can write:

$$\beta = \tanh \zeta \quad (2.34)$$

where $\zeta \in (-\infty, \infty)$ is the rapidity. In this parameterization:

$$\gamma = \frac{1}{\sqrt{1 - \beta^2}} = \frac{1}{\sqrt{1 - \tanh^2 \zeta}} = \cosh \zeta \quad (2.35)$$

where we have used $\cosh^2 \zeta = \sinh^2 \zeta + 1$. Note then that $\beta\gamma = \sinh \zeta$. So we can write the boost transformation as:

$$\Lambda = \begin{pmatrix} \cosh \zeta & -\sinh \zeta & 0 & 0 \\ -\sinh \zeta & \cosh \zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.36)$$

Which looks quite analogous to what we had with the spatial rotations, except we have replaced the trigonometric functions with their hyperbolic counterparts. Indeed, we can view boosts as hyperbolic rotations.

Under a boost, we have:

$$\begin{pmatrix} x'^0 \\ x'^1 \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \end{pmatrix} = \begin{pmatrix} \gamma x^0 - \beta\gamma x^1 \\ \gamma x^1 - \beta\gamma x^0 \end{pmatrix} \quad (2.37)$$

In the nonrelativistic limit of $v/c \ll 1$:

$$\gamma = (1 - \beta^2)^{-\frac{1}{2}} \approx 1 + \frac{1}{2}\beta^2 \quad (2.38)$$

so:

$$x'^0 \approx x^0, \quad x'^1 = x^1 - \frac{v}{c} \cdot ct \quad (2.39)$$

so we recover the expressions for old (Galilean) boosts. So classical physics was not wrong per se, it just emerges as the nonrelativistic limit of the correction theory. But when we look at relativistic physics (e.g. electromagnetic waves, high-energy particle scattering...) we need to use the relativistic theory.

On Tuesday, we discussed the 10 dimensional Galilean group which quantified the symmetry of classical mechanics. Now, we have modified boosts (i.e. the homogenous part of the Galilean group). The group consisting of 3 spatial rotations + 3 relativistic boosts yields the 6-dimensional Lorentz group, and the 4 spacetime translations yields the 10-dimensional Poincare group.

2.5 Doppler Effect

Recall the monochromatic wave solution:

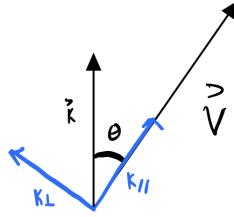
$$\phi = e^{ik_\mu x^\mu} \quad (2.40)$$

with:

$$k^2 = k_\mu k^\mu = 0 \quad (2.41)$$

$$\omega = ck^0 = ck, \quad \mathbf{k} = \hat{\mathbf{k}}k = \hat{\mathbf{k}} \frac{2\pi}{\lambda} \quad (2.42)$$

The two heroes of the story are the frequency and $\hat{\mathbf{k}}$ the direction of propagation.



The question is now - what does a boosted observer see? We consider the wavevector \mathbf{k} and decompose it into the components perpendicular and parallel to the boost velocity. θ is the angle between the two vectors. Let us calculate:

$$k'^\mu = \Lambda^\mu_\nu k^\nu \quad (2.43)$$

We have:

$$k'^0 = \gamma(k^0 - \boldsymbol{\beta} \cdot \mathbf{k}) = \gamma(k^0 - \beta k_\parallel) \quad (2.44)$$

$$k'_\parallel = \gamma(k_\parallel - \beta k^0) \quad (2.45)$$

$$\mathbf{k}'_\perp = \mathbf{k}_\perp \quad (2.46)$$

the last equality follows from the fact that the Λ matrix is trivial in this sector.

A good consistency check is that:

$$k'^2 = k'_\mu k'^\mu = 0 \quad (2.47)$$

which you can verify. This tells us that the boosted observer still sees the EM wave propagating at speed:

$$k'^0 = |\mathbf{k}'| \quad (2.48)$$

This is not something to be surprised about - the formalism was built for this to be true.

If the speed of propagation is not different, what is different? Well, let's look at the k'^0 relation. We can write this as:

$$\omega' = \gamma\omega(1 - \beta \cos \theta) \quad (2.49)$$

So, we see a different frequency, which is the Doppler effect! The direction of propagation is also modified:

$$\tan \theta = \frac{|\mathbf{k}_\perp|}{k_\parallel}, \quad \tan \theta' = \frac{|\mathbf{k}'_\perp|}{k'_\parallel} = \frac{|\mathbf{k}_\perp|}{\gamma(k_\parallel - \beta k^0)} \quad (2.50)$$

Since the denominator changes, so too must the angle.

Note the special case of the frequency relation when $\theta = \pi/2$, then:

$$\omega' = \gamma\omega \quad (2.51)$$

so the shift is manifestly relativistic.

2.6 Teaser - Relativistic Formulation of Maxwell Theory

You might say - hey, weren't we studying Maxwell theory the whole time? Indeed we were studying the wave equation, but there is more to Maxwell theory than this! We have to talk about two things:

- The dynamics of (charged) particles in electromagnetic fields
- The dynamics of the electromagnetic field itself

Consider the Lorentz force; suppose we have particle with spatial coordinate $\mathbf{x}(t)$ and charge q . In the presence of EM fields, the equation of motion is given by:

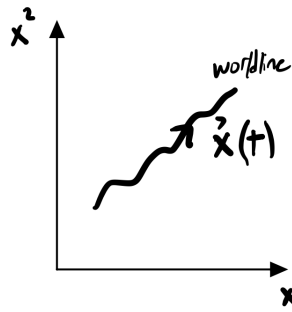
$$\frac{d(m\dot{\mathbf{x}})}{dt} = \frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \quad (2.52)$$

These equations are written in a highly inconvenient way for our purposes, as they treat time and space very differently. Time is a label and \mathbf{x} is the position of interest. But for us, we want to package things in terms of 4-vectors/spacetime coordinates $x^\mu = (x^0 = ct, \mathbf{x})$. Thus the above formulation is very inconvenient, e.g., to see how Lorentz transformations work in this theory. We will see how we can formulate the Maxwell theory in a relativistically covariant fashion next time.

3 Relativistic Formulation of a Free Particle

At the end of last week, we discovered that the wave equation (in Maxwell theory) must hold in all inertial frames, and as such the notion of boosts in Maxwell theory is different from the Galilean version in classical mechanics. Today what we will do is follow Einstein, and take the attitude that all physical laws must be Lorentz invariance.

To this end, we will need to generalize the idea of the coordinate $\mathbf{x}(t)$. In classical mechanics, the momentum $\mathbf{p} = m\dot{\mathbf{x}}$ is conserved, and can take on any value. But in the relativistic theory $|\dot{\mathbf{x}}|$ is bounded above by the speed of light. Some kind of reformulation is necessary. We also have the EM fields $\mathbf{E}(\mathbf{x}, t)$, $\mathbf{B}(\mathbf{x}, t)$ - these are already coming out of a Lorentz invariant theory, but we can rewrite the Maxwell equations such that they are manifestly Lorentz covariant.



In classical mechanics, we write down the Lagrangian:

$$\mathcal{L} = \frac{1}{2}m\dot{\mathbf{x}}^2 - U(\mathbf{x}) \quad (3.1)$$

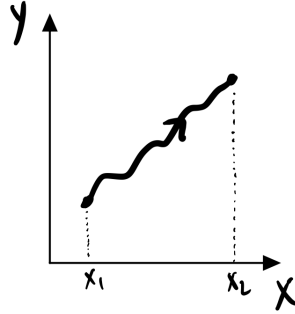
and then solve the EL equations to find the trajectories.

3.1 Parametrization

However, for a relativistically covariant formulation, we would like to put \mathbf{x}, t on the same footing. To this end, we can introduce a parameter λ , and describe the worldline of the particle as:

$$x^\mu = x^\mu(\lambda) \quad (3.2)$$

with $\mu = 0, 1, 2, 3$. This might look outlandish, but we do such parametrizations in simpler contexts. For example, consider a line in 2D space:



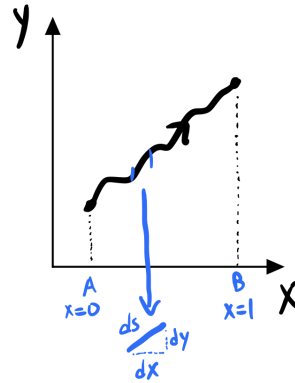
Then we could either write $y = y(x)$, or we could write $y = y(\lambda), x = x(\lambda)$ and make them depend on a common parameter:

$$x = \lambda, y = \lambda^2, \lambda \in [0, 1] \quad (3.3)$$

In doing so, we actually introduced an additional symmetry in the form of reparametrization invariance; indeed we can have x, y in terms of an arbitrary (monotonic, positive) function of λ :

$$x = f(\lambda), y = f^2(\lambda), \lambda \in [0, 1] \quad (3.4)$$

This can be useful for calculations in various contexts. For example, suppose we were interested in calculating the length of the line:



Then:

$$L = \int_A^B ds = \int_A^B \sqrt{dx^2 + dy^2} = \int_0^1 dx \sqrt{1 + \left(\frac{dy}{dx}\right)^2} = \int_0^1 dx \sqrt{1 + y'^2} \quad (3.5)$$

But if $x = x(\lambda), y = y(\lambda)$ for some parameter λ , we could then write:

$$ds^2 = (dx(\lambda))^2 + (dy(\lambda))^2 = \left(\frac{dx}{d\lambda} d\lambda\right)^2 + \left(\frac{dy}{d\lambda} d\lambda\right)^2 = d\lambda^2 (\dot{x}^2 + \dot{y}^2) \quad (3.6)$$

So:

$$L = \int_A^B ds = \int_A^B \sqrt{d\lambda^2 (\dot{x}^2 + \dot{y}^2)} = \int_A^B d\lambda \sqrt{\dot{x}^2 + \dot{y}^2} \quad (3.7)$$

the length calculation is now symmetric in x, y . If we replace $\lambda \rightarrow f(\lambda)$, the length should not change (for it shouldn't matter on the parameterization of the curve!) This must be a symmetry of the integral, and indeed we can easily check that it is. Under $\lambda \rightarrow f(\lambda)$:

$$x^\mu = \frac{dx^\mu}{d\lambda} \rightarrow \frac{dx^\mu}{df(\lambda)} = \frac{dx^\mu}{d\lambda} \frac{d\lambda}{df(\lambda)} \quad (3.8)$$

and the same for y^μ . Thus the length becomes:

$$L = \int_A^B df(\lambda) \sqrt{\left(\frac{dx^\mu}{df(\lambda)}\right)^2 + \left(\frac{dy^\mu}{df(\lambda)}\right)^2} = \int_A^B d\lambda \dot{f} \sqrt{\left(\frac{\dot{x}}{\dot{f}}\right)^2 + \left(\frac{\dot{y}}{\dot{f}}\right)^2} = \int_A^B d\lambda \sqrt{\dot{x}^2 + \dot{y}^2} \quad (3.9)$$

which we can see is invariant (as it should be).

3.2 Parametrizing Worldlines and Time Dilation

Back to our problem of parametrizing worldlines. We can repeat the idea from the simple context here, by describing the particle worldline as $x^\mu(\lambda)$. Other than the fact that the time component has a different sign:

$$ds^2 = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2 \quad (3.10)$$

we can play the same game as before; with $x^\mu = x^\mu(\lambda)$:

$$ds^2 = d\lambda^2 \left(-(\dot{x}^0)^2 + (\dot{x}^1)^2 + (\dot{x}^2)^2 + (\dot{x}^3)^2 \right) = d\lambda^2 \dot{x}^\mu \dot{x}_\mu = d\lambda^2 \dot{x}^\mu \dot{x}^\nu \eta_{\mu\nu} \quad (3.11)$$

This line element is reparametrization invariant, but also is Lorentz invariant (as it is written as a contraction with the Minkowski metric).

To do physics, we want to fix reparametrization symmetry. It is natural to fix $x^0(\lambda) = \lambda c$, so then $t(\lambda) = \lambda$. The logic is we have four functions, and we can freely reparametrize - using this symmetry, we may pick the time function to be simple. Let's see what happens to ds^2 when we do this:

$$ds^2 = dt^2 \left(-c^2 + \left(\frac{dx}{dt} \right)^2 \right) = -c^2 dt^2 (1 - \beta^2) \quad (3.12)$$

with $\beta = \frac{v}{c} = \frac{\dot{x}}{c}$. A natural definition is:

$$d\tau^2 = dt^2 (1 - \beta^2) \quad (3.13)$$

so:

$$ds^2 = -c^2 d\tau^2 \quad (3.14)$$

τ is called the proper time. In the frame where $\beta = 0$, we can see from the above $\tau = t$; it thus has the interpretation as the time in the inertial frame in which the particle is at rest momentarily. The relation between t, τ can be written as:

$$dt = \gamma d\tau \quad (3.15)$$

with:

$$\gamma^2 = \frac{1}{1 - \beta^2} \quad (3.16)$$

This is the celebrated time dilation formula. This says that things that take time $d\tau$ in the frame of the particle are dilated/increased by $\frac{1}{1-\beta^2}$ in the frame of the external observer. As a concrete example, if a particle has a lifetime T , someone who sees the particle moving will see its lifetime as γT . You may be familiar with the fact that cosmic rays create muons with, $T = 2.2 \cdot 10^{-6}\text{s}$. In the muon frame, it can only travel $2.2 \cdot 10^{-6}\text{s} \cdot 3 \times 10^8\text{ms}^{-1} = 0.66\text{km}$. But we observe that cosmic ray muons propagate for hundreds of kilometers, and the explanation is that their lifetime is extended via time dilation.

3.3 Lagrangian for Relativistic Particle

The usual Lagrangian:

$$\mathcal{L} = \frac{1}{2}m\dot{\mathbf{x}}^2 - U(\mathbf{x}) \quad (3.17)$$

does not work in our relativistic formulation as the particle could have any velocity, and velocities are bounded by the speed of light. Further, the action $S = \int dt \mathcal{L}$ is not invariant under Lorentz transformations. Can we modify this Lagrangian in some way? A very natural candidate is:

$$S = A \int d\lambda \sqrt{-\dot{x}_\mu \dot{x}^\mu}. \quad (3.18)$$

This is automatically Lorentz invariant. It is also reparametrization invariant, so let us fix the reparametrization symmetry by taking $\lambda = t$:

$$S = \tilde{A} \int dt \sqrt{1 - \frac{\dot{\mathbf{x}}^2}{c^2}}. \quad (3.19)$$

This is now invariant under the boosts that we defined last week - even though it looks like it mixes x/t because they are written separately, we know from the form of Eq. (3.18) that it is indeed invariant. It is also now clear why we took the minus sign there; we want the argument of the square root to be always positive.

We are not quite done yet. We have to check that the action passes basic scrutiny. One sanity check is that in the non-relativistic limit of $\dot{\mathbf{x}} \ll c$. Using the binomial approximation:

$$S = \tilde{A} \int dt \left[1 - \frac{1}{2} \frac{\dot{\mathbf{x}}^2}{c^2} + O\left(\frac{|\dot{\mathbf{x}}|^4}{c^4}\right) \right] \quad (3.20)$$

This motivates the choice that $\tilde{A} = -mc^2$:

$$S = \int dt \left(-mc^2 + \frac{1}{2}m\dot{\mathbf{x}}^2 + \dots \right) \quad (3.21)$$

This looks very good! We have the kinetic term with the correct normalization, and the first term can be thought of the rest energy of the particle $U = mc^2$. Note that the zero of the energy has been fixed - in regular classical mechanics, we usually have a freedom to choose what our “zero” is, but it does matter here.

The bottom line is our action:

$$S = -mc^2 \int dt \sqrt{1 - \frac{\dot{\mathbf{x}}^2}{c^2}} \quad (3.22)$$

has the properties that it is Lorentz invariant and it produces the correct formula in the low-velocity limit. This is the action we will be working with!

3.4 Relativistic energy and momentum

We want to turn the crank of the classical mechanics machinery and write down the Euler-Lagrange equations, and define energy and momentum. The E-L equations are:

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}^i} = \frac{\partial \mathcal{L}}{\partial x^i} \quad (3.23)$$

for $i = 0$. Since the Lagrangian is independent of x^i :

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}^i} = 0 \quad (3.24)$$

Thus we can define the (conserved) momentum:

$$p^i = \frac{m\dot{x}^i}{\sqrt{1 - \frac{\mathbf{v}^2}{c^2}}} = \gamma m \dot{x}^i \quad (3.25)$$

So the difference between the non-relativistic case ($m\dot{x}^i$) and relativistic case ($\gamma m \dot{x}^i$) is that γm is an effective “relativistic mass”; the faster the particle travels, the heavier it becomes/harder it becomes to change its momentum (and hence it can never surpass the speed of light).

We can also write down the Hamiltonian (equal to the energy because \mathcal{L} does not depend on time):

$$H = E = p_i \dot{x}^i - \mathcal{L} = \gamma mc^2 \quad (3.26)$$

where again the energy becomes larger and larger the more the particle approaches the speed of light.

You can easily convince yourself that:

$$E^2 = c^2 |\mathbf{p}|^2 + (mc^2)^2 \quad (3.27)$$

This is an interesting formula for the following reason. (E, \mathbf{p}) are a scalar and vector under rotations. Because rotations are a subgroup of Lorentz transformations, they presumably can be combined to a four-vector $p^\mu = (\frac{E}{c}, \mathbf{p})$. $p^\mu \rightarrow p'^\mu = \Lambda^\mu_\nu p^\nu$ under Lorentz transformations. In this language, Eq. (3.27) has the nice interpretation:

$$p_\mu p^\mu = -(mc)^2 \quad (3.28)$$

So the combination $p_\mu p^\mu$ is a Lorentz invariant quantity.

Next time, we will look at the relativistic formulation of Maxwell’s theory; we have fields $\mathbf{E}(\mathbf{x}, t)$, $\mathbf{B}(\mathbf{x}, t)$, which satisfy:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \quad (3.29)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (3.30)$$

$$\nabla \times \mathbf{B} - \frac{1}{c} \mathbf{E} = \mu_0 \mathbf{J} \quad (3.31)$$

$$\nabla \times \mathbf{E} + \dot{\mathbf{B}} = 0 \quad (3.32)$$

we will repackage these in such a way that it is clearer to see that they are Lorentz invariant. We will then combine our work on a relativistically invariant formulation of the mechanics of a particle with this.

4 Relativistic Formulation of Maxwell Theory

A small comment from last class - we discussed the fact that we could package the energy and momentum into a four vector:

$$(E, \mathbf{p}) = p^\mu \quad (4.1)$$

and that:

$$p_\mu p^\mu = p^\mu p^\nu \eta_{\mu\nu} = -(mc)^2 \quad (4.2)$$

is a Casimir of the Poincare group/is a Poincare invariant. That is to say - the above is an invariant under boosts/rotations/translations, and has the interpretation that it is the mass of the particle, which is unchanged under all such transformations.

Today, we move onto the relativistic formulation of the Maxwell theory - in your EM courses you have surely seen the 4 Maxwell equations:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \quad (4.3)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (4.4)$$

$$\nabla \times \mathbf{B} - \frac{1}{c} \mathbf{E} = \mu_0 \mathbf{J} \quad (4.5)$$

$$\nabla \times \mathbf{E} + \dot{\mathbf{B}} = 0 \quad (4.6)$$

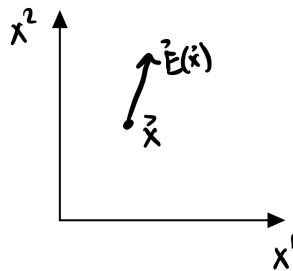
We suspect ahead of time that these are Lorentz invariant (as the wave equation is a consequence of these equations, and it is Lorentz invariant) but there is a way we can write them down such that they are manifestly so.

4.1 Maxwell under Rotations

Let's start with rotations, which are a subset of Lorentz transformations. It is not hard to show that the Maxwell equations as written above are already invariant under rotations - let's work this out. The charge density is a scalar, and so:

$$\rho'(\mathbf{x}') = \rho(\mathbf{x}) \quad (4.7)$$

Where $\mathbf{x}' = R\mathbf{x}$, or in index notation, $x'^i = R^{ij}x^j$. What of the vector quantities? $\mathbf{E}, \mathbf{B}, \mathbf{J}$ are vector fields, which have a direction at each point in space:



and you can convince yourself that vector fields have the following transformation:

$$E'_i(\mathbf{x}') dx'^i = E_i(\mathbf{x}) dx^i \quad (4.8)$$

4.2 Gauge Fields

What we would like to do is generalize this to all Lorentz transformations, where four-vectors $x^\mu = (t, \mathbf{x})$ transform as:

$$x'^\mu = \Lambda^\mu_\nu x^\nu. \quad (4.9)$$

The issue now becomes that it is unclear how \mathbf{E}, \mathbf{B} (which are *not* 4-vectors) transform. The main hint is that we may write the electric and magnetic fields in terms of a scalar and vector potential:

$$\mathbf{E} = -\nabla\phi - \dot{\mathbf{A}} \quad (4.10)$$

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (4.11)$$

there is a sense in which ϕ, \mathbf{A} are more fundamental than \mathbf{E}, \mathbf{B} , which we will see later in the lecture. Now, we are back to a problem that we have previously encountered. We have a scalar ϕ and a vector \mathbf{A} , so it makes sense to package them into a single four-vector, known as the gauge field:

$$A^\mu = \left(\frac{\phi}{c}, \mathbf{A}\right) \quad (4.12)$$

The question then becomes whether or not this is a useful definition, and the answer is yes!

Under Lorentz transformations, we expect:

$$A_\mu(x') dx'^\mu = A_\nu(x) dx^\nu \quad (4.13)$$

before we move on, it's worth picking apart exactly what this means; if we divide both sides by dx^ν :

$$A'_\mu(x') \frac{dx'^\mu}{dx^\nu} = A_\nu(x) \quad (4.14)$$

Now, the derivative we can read off from Eq. (4.9):

$$\frac{dx'^\mu}{dx^\nu} = \Lambda^\mu_\nu \quad (4.15)$$

so then:

$$A'_\mu(x') \Lambda^\mu_\nu = A_\nu(x) \quad (4.16)$$

4.3 Field Strength

Now, we have to check what the implication this is for on the \mathbf{E}, \mathbf{B} - the answer is that we can define the field strength of the gauge field:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (4.17)$$

where $\partial_\mu = \frac{\partial}{\partial x^\mu}$. I have a suspicion that $F_{\mu\nu}$ has something to do with \mathbf{E}, \mathbf{B} . The first reason is that $F_{\mu\nu}$ is a 4×4 antisymmetric matrix:

$$F_{\mu\nu} = -F_{\nu\mu} \quad (4.18)$$

which contains 6 independent components, which matches the total number of independent components stored in the \mathbf{E}, \mathbf{B} . The other reason is that $F_{\mu\nu}$ is gauge invariant. Last quarter, you noticed that, for an arbitrary scalar function χ , we could transform the scalar/vector potential as:

$$\phi' = \phi - \partial_t \chi, \quad \mathbf{A}' = \mathbf{A} + \nabla \chi \quad (4.19)$$

and under such a transformation \mathbf{E}, \mathbf{B} are invariant. Such transformations are known as gauge transformations, and are labelled by an arbitrary scalar function of spacetime. In order to discuss what happens

to $F_{\mu\nu}$ under such transformations, we first act how the gauge field $A_\mu(x)$ transforms under such a transformation:

$$A'_\mu(x) = A_\mu(x) + \partial_\mu \chi \quad (4.20)$$

Let's check this for the 0-th/time component:

$$A'_0 = A_0 + \frac{\partial}{\partial x^0} \chi \implies -\frac{\phi'}{c} = -\frac{\phi}{c} + \frac{\partial}{\partial(ct)} \chi \implies \phi' = \phi - \partial_t \chi \quad (4.21)$$

and the story for the spatial components is analogous. With this, let's check how the field strength transforms:

$$F'_{\mu\nu} = \partial_\mu A'_\nu - \partial_\nu A'_\mu = \partial_\mu (A_\nu + \partial_\nu \chi) - \partial_\nu (A_\mu + \partial_\mu \chi) = F_{\mu\nu} + \partial_\mu \partial_\nu \chi - \partial_\nu \partial_\mu \chi = F_{\mu\nu} \quad (4.22)$$

it is invariant! So now we're pretty confident that the components of F will give the components of the \mathbf{E}/\mathbf{B} fields. Let's work them out:

$$F_{i0} = -F_{0i} = \partial_i A_0 - \partial_0 A_i = -\frac{1}{c} \partial_i \phi - \partial_0 A_i = \frac{1}{c} E_i \quad (4.23)$$

$$F_{12} = -F_{21} = \partial_1 A_2 - \partial_2 A_1 = \epsilon_{3jk} \partial_j A_k = B_3 \quad (4.24)$$

(where in the last equality we recall that $B_i = \epsilon_{ijk} \partial_j A_k$). Thus for the general component of the B-field:

$$B_i = \frac{1}{2} \epsilon_{ijk} F_{jk} \quad (4.25)$$

the $\frac{1}{2}$ coming from the fact that when we do the j, k summations we get two nonzero terms. The bottom line of this discussion:

$$F_{\mu\nu} = \begin{pmatrix} 0 & -\frac{1}{c} E_1 & -\frac{1}{c} E_2 & -\frac{1}{c} E_3 \\ \frac{1}{c} E_1 & 0 & B_3 & -B_2 \\ \frac{1}{c} E_2 & -B_3 & 0 & B_1 \\ \frac{1}{c} E_3 & B_2 & -B_1 & 0 \end{pmatrix} \quad (4.26)$$

why did we do all of this? When we first wrote down the Maxwell equations, we had confusion about how \mathbf{E}, \mathbf{B} transform under Lorentz transformations. Now we have answered it - we can combine \mathbf{E}, \mathbf{B} into $F_{\mu\nu}$ and it is easy to see how this transforms, because we know how A_μ transforms under Lorentz:

$$F'^{\mu\nu}(x') = \Lambda^\mu_\lambda \Lambda^\nu_\rho F^{\lambda\rho}(x) \quad (4.27)$$

This looks very abstract, but it answers very physical questions. Suppose we have a charged particle, emanating an \mathbf{E} -field. I could then get on a bike² and ask what \mathbf{E}, \mathbf{B} -fields I observe - this allows us to answer it.

So far, we have just written down a bunch of definitions we hope would be useful. We had Maxwell's equations, and we wanted to make sure they were invariant under Lorentz transformations. Now, we are in a position to go back to this question, and ask if we can write Maxwell's equations in a way where the Lorentz invariance is manifest.

4.4 Lorentz Covariant Maxwell Equations

We require one more definition, that of the dual field strength. We have the field strength:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (4.28)$$

²Assuming I was very good at riding a bike and I could go close to the speed of light...

and then its dual is defined by:

$$F^{*\alpha\beta} = \tilde{F}^{\alpha\beta} := \frac{1}{2}\epsilon^{\alpha\beta\gamma\delta}F_{\gamma\delta} \quad (4.29)$$

where:

$$\epsilon^{0123} = 1 \quad (4.30)$$

and every time two indices are flipped we get a -1 sign. The $\frac{1}{2}$ in the definition is there to account for the fact that we get two factors in the summation over γ, δ .

We can explicitly work out the components:

$$\tilde{F}_{\mu\nu} = \begin{pmatrix} 0 & B_1 & B_2 & B_3 \\ -B_1 & 0 & \frac{1}{c}E_3 & -\frac{1}{c}E_2 \\ -B_2 & -\frac{1}{c}E_3 & 0 & \frac{1}{c}E_1 \\ -B_3 & \frac{1}{c}E_2 & -\frac{1}{c}E_1 & 0 \end{pmatrix} \quad (4.31)$$

from which we can see that F, \tilde{F} are related via $\mathbf{E} \rightarrow c\mathbf{B}, c\mathbf{B} \rightarrow -\mathbf{E}$. This is the beginning of a rabbit hole known as electromagnetic duality, and it has lead theorists to believe that there exist magnetic monopoles.

We claim that the dual tensor satisfies:

$$\partial_\mu \tilde{F}^{\mu\nu} = 0 \quad (4.32)$$

Let's prove this; fully writing out the dual tensor:

$$\tilde{F}^{\mu\nu} = \epsilon^{\mu\nu\gamma\delta}(\partial_\gamma A_\delta - \partial_\delta A_\gamma) \quad (4.33)$$

Now taking the derivative;

$$\partial_\mu \tilde{F}^{\mu\nu} = \epsilon^{\mu\nu\gamma\delta}(\partial_\mu \partial_\gamma A_\delta - \partial_\mu \partial_\delta A_\gamma) \quad (4.34)$$

Note that the derivative terms are symmetric under $\mu \leftrightarrow \gamma$ and $\mu \leftrightarrow \delta$, but $\epsilon^{\mu\nu\gamma\delta}$ is antisymmetric under such interchanges; hence the derivative vanishes:

$$\boxed{\partial_\mu \tilde{F}^{\mu\nu} = 0} \quad (4.35)$$

and we call this the *Bianchi identity*. Actually, it is four equations, one for each $\nu = 0, 1, 2, 3$. We can write this identity in terms of electric and magnetic fields:

$$\nabla \cdot \mathbf{B} = 0 \quad (4.36)$$

$$\nabla \times \mathbf{E} + \dot{\mathbf{B}} = 0 \quad (4.37)$$

which we can see are just two of the Maxwell equations - the equations are really Lorentz covariant, because we can write them in terms of $\tilde{F}^{\mu\nu}$, so if they are satisfied in one frame we can easily deduce that they are in a boosted frame.

We then have the other two equations which don't straightforwardly follow from the Bianchi identity:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \quad (4.38)$$

$$\nabla \times \mathbf{B} - \frac{1}{c}\dot{\mathbf{E}} = \mu_0 \mathbf{J} \quad (4.39)$$

The charge and current densities satisfy the continuity equation:

$$\dot{\rho} + \nabla \cdot \mathbf{J} = 0 \quad (4.40)$$

at this point, we're Pavlovian dogs who start drooling whenever we see a scalar and a vector together, so we're motivated to define:

$$J^\mu = (c\rho, \mathbf{J}) \quad (4.41)$$

wherein the continuity equation just reads:

$$\partial_\mu J^\mu = 0 \quad (4.42)$$

Under Lorentz transformations:

$$J'^\mu(x') dx'_\mu = J^\nu(x) dx_\nu \quad (4.43)$$

So let's get back to the remaining two Maxwell equations; you can convince yourself that:

$$\boxed{\partial_\mu F^{\mu\nu} = -\mu_0 J^\nu} \quad (4.44)$$

which is the relativistically covariant form of the Maxwell equations relating the fields to charges/currents.

One thing to point out - when we write this, one thing that is obvious is that if we can write this at all, the current must be conserved. The reason is that if I take the derivative ∂_ν :

$$\partial_\nu \partial_\mu F^{\mu\nu} = -\mu_0 \partial_\nu J^\nu \quad (4.45)$$

on the LHS we have a trivial zero (symmetric derivative of an antisymmetric tensor) so:

$$\partial_\nu J^\nu = 0 \quad (4.46)$$

the current must be conserved. Said another way, the electromagnetic field can only couple consistently to conserved currents.

A comment - on Tuesday, we discussed the dynamics of free particles. We might ask whether we've solved the problem of charged particles here - the answer is not quite, since throughout we've assumed ρ, \mathbf{J} are smooth and a charged pointlike particle would give rise to a delta-function concentrated charge/current. But we'll see how to deal with this later.

4.5 Lagrangian Formulation

We have found the Maxwell equations:

$$\partial_\mu \tilde{F}^{\mu\nu} = 0 \quad (4.47)$$

$$\partial_\mu F^{\mu\nu} = -\mu_0 J^\nu \quad (4.48)$$

and we want to know whether we can get these equations from Euler-Lagrange equations that follow from some Lagrangian.

First, we review the situation in point-particle mechanics. Then, we generalize to field theory. Finally, we will specialize to E& M.

In classical mechanics, we have generalized coordinates $q_i(t)$ for $i = 1, \dots, n$. The action is then:

$$S = \int dt \mathcal{L}(q_i, \dot{q}_i) \quad (4.49)$$

for $\mathcal{L} = T - U$. The Euler-Lagrange equations are then:

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} = \frac{\partial \mathcal{L}}{\partial q_i} \quad (4.50)$$

and you have solved this, e.g., when $T = \frac{1}{2} \sum_i m_i \dot{q}_i^2$ and $U = U(q_1, \dots, q_n)$.

We now want to promote these coordinates to fields, $q_i(t) \rightarrow \phi_i(t, \mathbf{x})$, and we shall continue with this next lecture.

5 Lagrangian for EM field

In relativistic notation, Maxwell's equations are:

$$\partial_\mu \tilde{F}^{\mu\nu} = 0 \quad (5.1)$$

$$\partial_\mu F^{\mu\nu} = -\mu_0 J^\nu \quad (5.2)$$

where:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (5.3)$$

$$\tilde{F}^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma} \quad (5.4)$$

we then raised the question: could we derive these equations from a Lagrangian?

5.1 From particles to fields

Brief review of classical mechanics of particles - we consider an action:

$$S = \int dt \mathcal{L}(q_i(t), \dot{q}_i(t)) \quad (5.5)$$

Which, extremizing, we obtain the Euler-Lagrange equations:

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} = \frac{\partial \mathcal{L}}{\partial q_i} \quad (5.6)$$

which are equivalent to Newton's second law.

We now generalize, in two steps. First, we generalize the generalized coordinates $q_i(t)$ to a field ϕ :

$$q_i(t) \rightarrow \phi(x) = \phi(\mathbf{x}, t) \quad (5.7)$$

Comparing to what we had in the case of particle mechanics, instead of the i index we have the position \mathbf{x} - we have a continuous number of degrees of freedom. Field theory is nothing more than classical mechanics of an infinite degrees of freedom.

We can further generalize:

$$\phi(x) \rightarrow \phi_i(x) \quad (5.8)$$

where $i = 1, \dots, n$. This is not to be confused with the index i for the particle label. This is just saying we have $i = 1, \dots, n$ different fields.

In our application, we take $\phi_i \rightarrow A_\mu$. The spacetime index μ plays the role of i . A_μ is a vector field, and in full gory detail:

$$A_\mu = A_\mu(x^0, x^1, x^2, x^3) \quad (5.9)$$

with $\mu = 0, 1, 2, 3$. Lorentz transformations act on both indices.

Comment: this is similar to the role of spin in quantum mechanics; we had wavefunctions $\phi(\mathbf{x}, t)$ which we obtained by solving the Schrodinger equation. When we solved for the energy levels of the hydrogen atom, we likely ignored the fact that the electron had spin. But in the presence of the magnetic field, there is a term in the Hamiltonian that acts on the spin, so we add an extra index α to $\psi_\alpha(\mathbf{x}, t)$ where $\alpha = \uparrow, \downarrow$. μ plays an analogous role - we have spin-1.

Going back to the action, we can write (restricting to just one field for now):

$$S = \int d^4x \mathcal{L}(\phi(x), \partial_\mu \phi(x)). \quad (5.10)$$

We want to understand what is the analog of the Euler-Lagrange equations? We will follow the same logic in the particle case - we vary the field ϕ :

$$\phi(x) \rightarrow \phi(x) + \delta\phi \quad (5.11)$$

where the variation $\delta\phi$ is arbitrary and small³. We then ask by how much does $S(\phi)$ change under the variation, to first order in $\delta\phi$:

$$\begin{aligned} \delta S &= S(\phi + \delta\phi) - S(\phi) \\ &= \int d^4x \left[\mathcal{L}(\phi + \delta\phi, \partial_\mu\phi + \partial_\mu\delta\phi) - \mathcal{L}(\phi, \partial_\mu\phi) \right] \\ &= \int d^4x \left[\delta\phi \frac{\partial \mathcal{L}}{\partial \phi} + \partial_\mu\delta\phi \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)} \right] \\ &= \int d^4x \left[\delta\phi \frac{\partial \mathcal{L}}{\partial \phi} + \partial_\mu \left(\delta\phi \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)} \right) - \delta\phi \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)} \right] \end{aligned} \quad (5.12)$$

Now, the middle term is an integral of a total derivative. We can thus write it as (for $\mu = 0$):

$$\int d^4x \partial_0 \left[\left(\frac{\partial \mathcal{L}}{\partial(\partial_0\phi)} \right) \delta\phi \right] = \int d^3x \frac{\partial \mathcal{L}}{\partial(\partial_0\phi)} \delta\phi \Big|_{x^0=-\infty}^{x^0=\infty} \quad (5.13)$$

and we can choose the variation $\delta\phi$ to go to zero at the boundary of time. This is analogously true for the spatial μ s. Thus the total derivative term drops out, and the variation in the action becomes:

$$\delta S = \int d^4x \delta\phi(x) \left[\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)} \right] = 0 \quad (5.14)$$

and if we want this to hold for all variations $\delta\phi(x)$, we thus obtain:

$$\frac{\partial \mathcal{L}}{\partial \phi} = \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)} \quad (5.15)$$

or in the case of multiple fields:

$$\boxed{\frac{\partial \mathcal{L}}{\partial \phi_i} = \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi_i)}} \quad (5.16)$$

5.2 EM field Lagrangian - no sources

We want:

$$S = \int d^4x \mathcal{L}(A_\mu(x), \partial_\nu A_\mu(x)) \quad (5.17)$$

note that the Lagrangian depends on the gauge potential A_μ as opposed to the electric/magnetic field. We want to get the Maxwell equations out of the Euler-Lagrange equations:

$$\frac{\partial \mathcal{L}}{\partial A_\nu} = \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \quad (5.18)$$

What \mathcal{L} should we choose? First, we have the equation/Bianchi identity:

$$\partial_\mu \tilde{F}^{\mu\nu} = 0 \quad (5.19)$$

³The proof of the fact that Kusatov has free will is that he can choose any variation he likes, here.

but this equation is a triviality. This is just true by the definition of the (dual) field strength. So we don't want this from the Lagrangian. What we actually want to derive from the Lagrangian is:

$$\partial_\mu F^{\mu\nu} = -\mu_0 J^\nu \quad (5.20)$$

For simplicity, let's start in the case with $J^\nu = 0$. We then want to reproduce:

$$\partial_\mu F^{\mu\nu} = \partial_\mu (\partial^\nu A^\mu - \partial^\mu A^\nu) = 0. \quad (5.21)$$

\mathcal{L} should satisfy some properties:

- \mathcal{L} should be Lorentz invariant - this is because S should be Lorentz invariant, and d^4x is Lorentz invariant.
- \mathcal{L} should be gauge invariant - the Maxwell equation and physics we get out of it should be gauge invariant.
- The Euler-Lagrange equation is linear in A_μ - thus \mathcal{L} should be quadratic in A_μ .

Given these conditions, we have only three objects we could conceivably write⁴:

- $-F_{\mu\nu}F^{\mu\nu}$
- $-F_{\mu\nu}\tilde{F}^{\mu\nu}$
- $-\tilde{F}_{\mu\nu}\tilde{F}^{\mu\nu}$

Note that $-\tilde{F}_{\mu\nu}\tilde{F}^{\mu\nu}$ is proportional to $-F_{\mu\nu}F^{\mu\nu}$ and so is not a unique possibility, and $-F_{\mu\nu}\tilde{F}^{\mu\nu}$ is a total derivative and so does not contribute to the physics. We thus have one possibility remaining:

$$\mathcal{L}_{\text{EM}} = -\frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu} \quad (5.22)$$

The proportionality constant we put there for later convenience, but this is not important for the purposes of the Euler-Lagrange equation. In the next problem set, you can verify that the E-L equation yields precisely $\partial_\mu F^{\mu\nu} = 0$. This must work - if it did not, there would be no Lagrangian that could possibly work, and we just wasted 2 weeks of your time. Writing \mathcal{L} in terms of the electromagnetic fields, we have:

$$\mathcal{L}_{\text{EM}} = \frac{1}{2\mu_0} \left(\frac{1}{2}|\mathbf{E}|^2 - |\mathbf{B}|^2 \right) \quad (5.23)$$

wherein we can immediately see that the Euler-Lagrange procedure would be doomed to fail if we had used \mathbf{E}, \mathbf{B} as the fundamental degrees of freedom.

5.3 EM field Lagrangian - with sources

If we want to reproduce the case with sources:

$$\partial_\mu F^{\mu\nu} = -\mu_0 J^\nu \quad (5.24)$$

we can simply add the term:

$$\mathcal{L} = -\frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu} + A_\mu J^\mu. \quad (5.25)$$

But actually there is a problem here. This Lagrangian is *not* gauge invariant. If we perform the gauge transformation:

$$A_\mu \rightarrow A_\mu - \partial_\mu \chi \quad (5.26)$$

⁴If you can think of another, please tell Kusatov, and he will immediately run to write a paper on it without you.

Then the Lagrangian picks up a new term:

$$\delta\mathcal{L} = -J^\mu\partial_\mu\chi \quad (5.27)$$

Is this a problem? Actually, no. What we *really* want is not that \mathcal{L} is gauge invariant, but that the action S is gauge invariant. In other words, does δS from making the gauge transformation vanish?

$$\delta S = \int d^4x \delta\mathcal{L} = - \int d^4x J^\mu \partial_\mu\chi = - \int d^4x \partial_\mu [J^\mu\chi] \quad (5.28)$$

In the last line, we wrote things as a total derivative; this is valid as:

$$\partial_\mu [J^\mu\chi] = (\partial_\mu J^\mu)\chi + J^\mu(\partial_\mu\chi) = J^\mu(\partial_\mu\chi) \quad (5.29)$$

as $\partial_\mu J^\mu = 0$ (continuity equation/current conservation). Going back to δS , we can have some discussion about the integral of the total derivative. As we discussed earlier in lecture, this only receives contributions from the boundaries of spacetime. So, the action is invariant under gauge transformations that go to zero rapidly to zero at the boundaries of spacetime. But if we include χ s that do *not* go to zero, we get physically important boundary terms - they are relevant to charge conservation. But in this course we do not consider such cases.

Another point to make - the discussion heavily relies on the fact that the current J^μ is conserved. If we want the action to be invariant under gauge transformations, we can only couple the electromagnetic field to conserved currents.

5.4 The stress-energy tensor

Now, we can ask what this Lagrangian is good for - of course, we already have the Euler-Lagrange equation; another answer is we can get the stress-energy tensor for the EM field. Why is this physically important? You are already viscerally aware of the fact that particles have energy and momenta. But you also expect that the electromagnetic field also has this property. What is the analogue of the energy/momentum of a bowling ball for an EM wave?

To review, let's just remind ourselves of how things work for particles. There, we have momenta:

$$p_i = \frac{\partial\mathcal{L}}{\partial\dot{q}_i} \quad (5.30)$$

and the Hamiltonian:

$$\mathcal{H} = p_i\dot{q}^i - \mathcal{L} \quad (5.31)$$

where if the Lagrangian does not depend on time, $\mathcal{H} = E$ is the conserved energy. We want to generalize this discussion to fields. We have the action:

$$S = \int d^4x \mathcal{L}(\phi, \partial_\mu\phi) \quad (5.32)$$

and if there is no explicit dependence on x^μ , then we expect that the action is invariant under $x^\mu \rightarrow x^\mu + a^\mu$, and so we expect p_μ to be conserved.

Let us recall the example of electric charge. We have a conserved current:

$$\partial_\mu J^\mu = 0. \quad (5.33)$$

Why do we care? If we define a charge:

$$Q = \int d^3x J^0(\mathbf{x}, t) \quad (5.34)$$

where $\dot{Q} = 0$ as:

$$\dot{Q} = (\text{Const.}) \int d^3x \partial_0 J^0 = (\text{Const.}) \int d^3x \nabla \cdot \mathbf{J} = 0 \quad (5.35)$$

where the last equality follows from current conservation if we integrate over all space (if we integrate over a finite region, we keep track of the charge that flows in/out of a region). Note that Q here is a scalar quantity. This is true because J^μ is a Lorentz vector.

What about p^μ ? What is the conserved current in this case? Here we expect a tensor $T^{\mu\nu}$ which satisfies $p_\mu T^{\mu\nu} = 0$. Why is this the case?

$$p^\nu = \int d^3x T^{0\nu} \quad (5.36)$$

so:

$$\dot{p}^\nu = (\text{Const.}) \int d^3x \partial_0 T^{0\nu} = (\text{Const.}) \int d^3x \partial_i T^{i\nu} = 0 \quad (5.37)$$

if we integrate over all space. We want a tensor so that the conserved quantity p^ν we get out of it is a Lorentz vector.

We define the stress-energy tensor as:

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^i)} \partial^\nu \phi^i - \eta^{\mu\nu} \mathcal{L}. \quad (5.38)$$

Let us verify that the conservation equation:

$$\partial_\mu T^{\mu\nu} = 0 \quad (5.39)$$

is satisfied.

$$\begin{aligned} \partial_\mu T^{\mu\nu} &= \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^i)} \partial^\nu \phi^i \right) - \eta^{\mu\nu} \partial_\mu \mathcal{L} \\ &= \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^i)} \right) \partial^\nu \phi^i + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^i)} \partial_\mu \partial^\nu \phi^i - \partial^\nu \mathcal{L}(\phi^i, \partial_\mu \phi^i) \\ &= \frac{\partial \mathcal{L}}{\partial \phi^i} \partial^\nu \phi^i + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^i)} \partial_\mu \partial^\nu \phi^i - \partial^\nu \mathcal{L}(\phi^i, \partial_\mu \phi^i) \\ &= 0 \end{aligned} \quad (5.40)$$

where in the third equality we use the Euler-Lagrange equation, and in the last equality we can recognize the terms to cancel out using the chain rule to evaluate the derivative of the Lagrangian.

Next time, we study $T^{\mu\nu}$ for an EM field.