PHYS 323 (Advanced Electrodynamics II) Notes

Rio Weil

This document was typeset on April 1, 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

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 $\mathbf{x}_1(t), \mathbf{x}_2(t), \ldots, \mathbf{x}_N(t)$, 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:

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

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

$$m_i \ddot{\mathbf{x}}_i = -\boldsymbol{\nabla}_i \sum_{y \neq i} V_{ij}(|\mathbf{x}_i - \mathbf{x}_j|) \tag{1.2}$$

where we work in 3-dimensions, so $\mathbf{x}_i = (x_i, y_i, z_i)$ and $\boldsymbol{\nabla}_i = (\frac{\partial}{\partial x_i}, \frac{\partial}{\partial y_i}, \frac{\partial}{\partial z_i})$. 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:
$$\mathbf{x}_i \to \mathbf{x}_i + \mathbf{a} \tag{1.3}$$
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:
$$t \to t + t_0 \tag{1.4}$$
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:
$$\mathbf{x}_i \to R\mathbf{x}_i \tag{1.5}$$
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:
$$\mathbf{x}_i \to \mathbf{x}_i - \mathbf{v}t \tag{1.6}$$
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.
- $\Delta x = |\mathbf{x}_2 - \mathbf{x}_2|$ - is not invariant under boosts, as under $\mathbf{x}' = \mathbf{x} - \mathbf{v}t$ we have:

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

  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:

$$\Box\phi(\mathbf{x}, t) := (-\frac{1}{c^2}\partial_t^2 + \boldsymbol{\nabla}^2)\phi(\mathbf{x}, t) = 0 \tag{1.8}$$

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:

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

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 $\mathbf{x} \to \mathbf{x} + \mathbf{a}, t \to t + t_0$, then we have:

$$\phi(\mathbf{x}, t) \to \phi(\mathbf{x} + \mathbf{a}, t + t_0) = Ce^{i\mathbf{k}\cdot(\mathbf{x}+\mathbf{a}) - i\omega(t+t_0)} = Ce^{i\mathbf{k}\cdot\mathbf{a} - i\omega t_0}e^{i\mathbf{k}\cdot\mathbf{x} - i\omega t} \tag{1.10}$$

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

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

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

  so we have a new ω':

$$\omega' = \omega + \mathbf{k} \cdot \mathbf{v} \tag{1.12}$$

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

$$\omega + \mathbf{k} \cdot \mathbf{v} \neq c|\mathbf{k}| \tag{1.13}$$

  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:

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

with boosted solutions:

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

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

$$V = \frac{\omega + kv}{k} = \frac{\omega}{k} + v = c + v \tag{1.16}$$

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:

$$\mathbf{x}' = R\mathbf{x} \tag{1.17}$$

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

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

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

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

We then have the rotated wavenumber:

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

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:

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

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:

$$ds^2 = dx^i dx^i = dx^2 + dy^2 + dz^2 \tag{1.22}$$

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

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

ds is rotationally invariant:

$$ds^2 = dx^i dx^i = dx'^i dx'^i \tag{1.24}$$

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:

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

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

$$\eta = \mathrm{diag}(-1, 1, 1, 1). \tag{1.26}$$

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:

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

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:

$$x_\mu = \eta_{\mu\nu} x^\nu \tag{1.28}$$

so:

$$x_0 = -x^0, x_i = x^i \tag{1.29}$$

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:

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

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'Alambertian.

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

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

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:

$$ds^2 = dx^2 + dy^2 + dz^2 = dx^i dx^i \tag{2.1}$$

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

$$x'^i = R^i_j x^j \tag{2.2}$$

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

$$ds'^2 = dx'^i dx'^i = dx^i dx^i = ds^2 \tag{2.3}$$

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

$$D^2 = |\mathbf{x}_2 - \mathbf{x}_1|^2 \tag{2.4}$$

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

$$\nabla'^2 = \frac{\partial}{\partial x'^i}\frac{\partial}{\partial x'^i} = \nabla^2 \tag{2.5}$$

The idea was to now include time:

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

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

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

Note that we can write the line element equivalently as:

$$ds^2 = dx^\mu dx_\mu \tag{2.8}$$

where:

$$dx_\mu = \eta_{\mu\nu} dx^\nu \tag{2.9}$$

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

$$dx_0 = -dx^0, \quad dx_i = dx^i \tag{2.10}$$

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:

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

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:

$$x'^\mu = \Lambda^\mu_{\ \nu} x^\nu \tag{2.12}$$

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

Λ satisfies:

$$\Lambda^\mu_{\ \lambda}\Lambda^\nu_{\ \rho}\eta_{\mu\nu} = \eta_{\lambda\rho} \tag{2.13}$$

2.3 Solutions to the Wave Equation

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

$$\Box = -\frac{1}{c^2}\partial_t^2 + \nabla^2 = \partial_\mu\partial^\mu = \partial_\mu\partial_\nu\eta^{\mu\nu} \tag{2.14}$$

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

$$\Box_x = \Box_{x'} \tag{2.15}$$

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

$$\Box_x\phi(x) = 0 \tag{2.16}$$

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

$$\phi'(x') = \phi(x) \tag{2.17}$$

An important example is plane waves:

$$\phi(x^\mu) = e^{ik_\mu x6\mu} = e^{ik^\mu x^\nu \eta_{\mu\nu}} \tag{2.18}$$

Note that if ϕ is a solution:

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

thus:

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

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

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

[1] If 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:

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

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

$$k^0 = |\mathbf{k}|. \tag{2.23}$$

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

$$e^{ik'_\mu x^\mu} \tag{2.24}$$

where:

$$k'_\mu = \Lambda_\mu^{\ \nu} k_\nu \tag{2.25}$$

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

$$k'^\mu k'_\mu = 0 \tag{2.26}$$

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:

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & R & \\ 0 & & & \end{pmatrix} \tag{2.27}$$

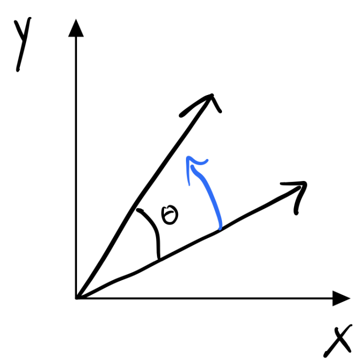
Note that for these:

$$x'^0 = \Lambda^0_{\ \nu} x^\nu = x^0 \tag{2.28}$$

$$x'^i = \Lambda^i_{\ \nu} x^\nu = R^i_j x^j \tag{2.29}$$

for $R^t R = RR^t = \mathbb{I}$, 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:

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

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

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

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

$$\gamma^2(1-\beta^2) = 1 \tag{2.32}$$

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

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{2.33}$$

We can write:

$$\beta = \tanh\zeta \tag{2.34}$$

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

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

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:

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

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:

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

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

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

so:

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

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:

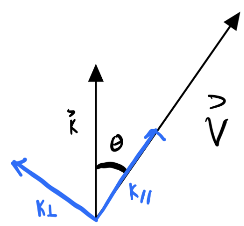
$$\phi = e^{ik_\mu x^\mu} \tag{2.40}$$

with:

$$k^2 = k_\mu k^\mu = 0 \tag{2.41}$$

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

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:

$$k'^\mu = \Lambda^\mu_{\ \nu} k^\nu \tag{2.43}$$

We have:

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

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

$$\mathbf{k}'_\perp = \mathbf{k}_\perp \tag{2.46}$$

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

A good consistency check is that:

$$k'^2 = k'_\mu k'^\mu = 0 \tag{2.47}$$

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

$$k'^0 = |\mathbf{k}'| \tag{2.48}$$

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:

$$\omega' = \gamma\omega(1 - \beta\cos\theta) \tag{2.49}$$

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

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

Since the denominator changes, so too must the angle.

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

$$\omega' = \gamma\omega \tag{2.51}$$

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:

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

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 be 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:

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

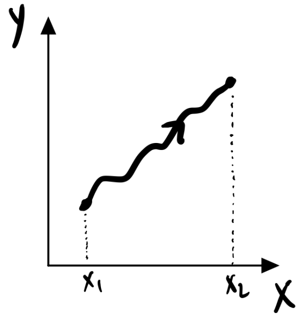
and then solve the EL equations to find the trajectories.

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:

$$x^\mu = x^\mu(\lambda) \tag{3.2}$$

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:



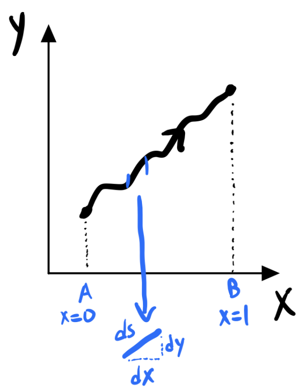
Then we could either write $y = y(x)$, or we could write $y = y(\lambda), x = x(\lambda)$ and make them depend on a common parameter:

$$x = \lambda, y = \lambda^2, \lambda \in [0, 1] \tag{3.3}$$

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:

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

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



Then:

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

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

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

So:

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

the length calculation is now symmetric in x, y. If we replace $\lambda \to f(\lambda)$, 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 $\lambda \to f(\lambda)$:

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

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

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

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 worldine as $x^\mu(\lambda)$. Other than the fact that the time component has a different sign:

$$ds^2 = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2 \tag{3.10}$$

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

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

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:

$$ds^2 = dt^2\left(-c^2 + \left(\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\right)^2\right) = -c^2 dt^2\left(1 - \boldsymbol{\beta}^2\right) \tag{3.12}$$

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

$$d\tau^2 = dt^2(1 - \boldsymbol{\beta}^2) \tag{3.13}$$

so:

$$ds^2 = -c^2 d\tau^2 \tag{3.14}$$

τ is called the proper time. In the frame where $\boldsymbol{\beta} = 0$, 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:

$$dt = \gamma d\tau \tag{3.15}$$

with:

$$\gamma^2 = \frac{1}{1 - \boldsymbol{\beta}^2} \tag{3.16}$$

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 $\frac{1}{1-\boldsymbol{\beta}^2}$ 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, $T = 2.2 \cdot 10^{-6}$s. In the muon frame, it can only travel $2.2 \cdot 10^{-6}\text{s} \cdot 3 \times 10^8 \text{ms}^{-1} = 0.66$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:

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

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:

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

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

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

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:

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

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

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

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:

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

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 Relavistic 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:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}^i} = \frac{\partial \mathcal{L}}{\partial x^i} \tag{3.23}$$

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}^i} = 0 \tag{3.24}$$

Thus we can define the (conserved) momentum:

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

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):

$$H = E = p_i \dot{x}^i - \mathcal{L} = \gamma m c^2 \tag{3.26}$$

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

You can easily convince yourself that:

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

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 $p^\mu = (\frac{E}{c}, \mathbf{p})$. $p^\mu \to p'^\mu = \Lambda^\mu{}_\nu p^\nu$ under Lorentz transformations. In this language, Eq. (3.27) has the nice inrepretation:

$$p_\mu p^\mu = -(mc)^2 \tag{3.28}$$

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:

$$\boldsymbol{\nabla} \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \tag{3.29}$$

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0 \tag{3.30}$$

$$\boldsymbol{\nabla} \times \mathbf{B} - \frac{1}{c}\mathbf{E} = \mu_0 \mathbf{J} \tag{3.31}$$

$$\boldsymbol{\nabla} \times \mathbf{E} + \dot{\mathbf{B}} = 0 \tag{3.32}$$

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.