Ph. 504 Relativity PSet 2: Lagrange, Intervals, and Lorentz Transformations

Dimitrios Kidonakis
KMHS Physics Club, Spring 2018

> "Since you are now studying the relativity theory, I will give you a problem. A spaceship travels in deep space. It left Earth with a cargo of water. It grosses 10^5 tonnes. It is bound for Andromeda. The solar sail is broken, the physicist is on deck, there are 144 passengers aboard, the anti-matter engines are nominal, the clock points to quarter past three in the afternoon. It is the month of May. How old is the captain?"
>
> — *Gustave Flaubert*

The difficulty of the problems are ranked with a 3-star system. No stars means standard exercise (5 min), one star means non-standard application but simple (15 min), two stars indicates a challenging problem (30 min), and three stars indicates a graduate-level problem (1 hr+). For comparison, all the problems in PSet 1 were one star. I have included hints for selected problems, located in the last pages, as well as full solutions to the problems.

1. Introduction to classical mechanics

The ideas of momentum, force, Lagrangians, and Hamiltonians are fundamental to all physics and are necessary to derive relativity. The reader is assumed to understand basic classical mechanics from the calculus standpoint, so this section will describe Lagrangians and Hamiltonians only. As none of the material presented in this section is required until §5, the anxious reader can start §2 immediately.

We will now derive physics completely from scratch. Before we can write the equations of classical mechanics down, we need to know what they depend upon. For a system of N particles, we need three spatial coordinates to describe the position of each particle at some instant – so we need $3N$ total numbers to describe the position of the system, or N position vectors. We say that this system has $3N$ ***degrees of freedom.*** We want to completely describe the system, i.e. be able to calculate the trajectories of all the particles based on information at on instant. How much information do we need to perform the calculation (Imagine that someone filmed the movement of billiard balls with an old film camera. They show you one frame with the instantaneous positions and velocities written for each ball. Would you be able to calculate the positions of the billiard balls in the next frame? What about the frame after that?)? First we would like to be able to calculate the positions of the particles after some instantaneous time dt. Then

$$r_i(t_0 + dt) = r_i(t_0) + \dot{r}_i(t_0)dt,$$ [1]

[1] If this is not obvious, than consider the limit definition of the derivative:

$$f'(t_0) = \lim_{t \to 0} \frac{f(t_0 + t) - f(t_0)}{t} = \frac{f(t_0 + dt) - f(t_0)}{dt}.$$

The fundamental reason why only these two quantities $(r, \dot{r})$ are sufficient is because ***the derivative is*** so both the position and the velocity at t_0 are needed to describe the positions of the particles in the next instant. Now we would like to calculate the instant after that. Say $t_1 = t_0 + dt$, so

$$r_i(t_1 + dt) = r_i(t_1) + \dot{r}_i(t_1)dt = r_i(t_0) + 2\dot{r}_i(t_0)dt + \ddot{r}_i(t_0)dt^2 = r_i(t_0) + 2\dot{r}_i(t_0)dt$$

and in general

$$r_i(t_n) = r_i(t_0) + n\dot{r}_i(t_0)dt.$$

Now it is clear that we only need the position and velocity at t_0, although if one includes terms $\mathcal{O}(dt)^2$ in fact one would really need ***all the derivatives*** $r^{(n)}(t_0)$ in order to determine the position of the particles at any time. Although the disregard of large powers of infinitesimals is a standard practice in calculus, it is surprising, at least to the author, that the calculation of trajectories only depends on position and velocity. This is a fundamental principle in physics. In fact it is called ***Newton's law of determinacy***. [1] It is equivalent to the curious fact that most differential equations in physics are of second order; consider, for example, Newton's second law. [2]

Now consider a function, in an ambiguous sense, that describes the motion of this set of particles, called $\mathcal{L}$, the Lagrangian. As we have just shown, this function only depends on $q_i, \dot{q}_i$ where q_i is the position of the ith particle. Let's consider the motion of this system from point A to point B. We know from experience that the path the system will take is the "path of least resistance;" an example of this is light moving in a straight line. We can think of L, the action, as a ***functional*** that describes the "resistance" of a path – it is a function that, unlike usual functions e.g. $f : \mathbb{R}^n \to \mathbb{R}^m$, outputs a number for a path. We write this as

$$L[x] : \text{path} \to \mathbb{R},$$

where $x(t)$ is a path. The action is found by taking the Lagrangian at each point of the path and summing it all up:

$$L[q] = \int_{t_0}^{t_1} \mathcal{L}(q, \dot{q}, t)dt.$$

The particular path of least resistance q satisfies the usual intuitive relative extremum condition: that if we consider a slightly different path, a ***variation***, then the value of L does not change. This is just the same as considering the extremum condition for functions: A slight change dx from the minimum results in a change of the order dx^2 of

a linear function. Refer to https://usamo.files.wordpress.com/2017/12/napkin-2017-12-11.pdf, page 263, for more. Note to self grad level problem, find the euler lagrange equations assuming $L(q, \dot{q}, \ddot{q})$

[1] Apparently, this fact should be "known from experience" according to Landau and Lifshitz, Vol 1. Arnold says that "it is hard to doubt this fact, since we learn it very early. One can imagine a world in which to determine the future of a system one must also know the acceleration at the initial moment, but experience shows us that our world is not like this." (Mathematical Methods of Classical Mechanics) These two books are in the top three best classical mechanics books, so these statements should be considered carefully. I think they essentially state that Nature is sufficiently smooth and well-behaved.

[2] ***If you are not convinced, Ostrogradsky showed that Lagrangians with higher derivatives are not stable.***

the function[1]. We can apply this idea to the present problem, and we get

$$\delta L[q] = \delta \int_{t_0}^{t_1} \mathcal{L}(q, \dot{q}, t)dt = \int_{t_0}^{t_1} \left(\frac{\partial \mathcal{L}}{\partial q}\delta q + \frac{\partial \mathcal{L}}{\partial \dot{q}}\delta \dot{q} \right) dt = 0,$$

which is analogous to the total derivative in the multivariable calculus.[2]

We compute the right part of the integral by parts, taking $u = \frac{\partial \mathcal{L}}{\partial \dot{q}}$ and $dv = \delta \dot{q} dt$ in the usual notation. As an exercise, check this computation. We get

$$\delta q \frac{\partial \mathcal{L}}{\partial \dot{q}} \Big|_{t_0}^{t_1} + \int_{t_0}^{t_1} \left(\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}} \right) \delta q dt = 0.$$

Now notice that the term outside the integral in the left hand side is 0 since the endpoints are fixed – $\delta q(t_0) = \delta q(t_1) = 0$. If we add a small kink in the path, δq changes, and furthermore we can change this as we wish – since we need the zero condition to hold over any path, the argument of the integral must be 0. Thus we get the ***Euler-Lagrange*** equations (sometimes just called the ***Lagrange*** equations)

$$\frac{\partial \mathcal{L}}{\partial q} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}}.$$

Exercises

1.1** The total derivative employed in varying the Lagrangian may not be clear, especially because you may have not mastered multivariable calculus yet. You will compute this step in a more intuitive way. As presented in the text, the variation was presented as an analogue to the total derivative. Here, you will consider another way of computing extrema of single-variable functions. Consider the neighborhood of the minimum of a function, say $f(x) = x^2 - 1$. In other words,

$$f(x+\epsilon) = (x+\epsilon)^2 - 1 = x^2 + 2x\epsilon + \epsilon^2 - 1, \quad f(x-\epsilon) = (x-\epsilon)^2 - 1 = x^2 - 2x\epsilon + \epsilon^2 - 1.$$

2. Special Relativity

So far we have shown three counter-intuitive relativistic effects, time dilation, length contraction, and loss of simultaneity, which are consequences of the invariance of the speed of light.

Our approach in deriving special relativity in the following sections is by examining invariants. Just as in the Newtonian physics, one person will measure the endpoints of a rod at different coordinate points from another person, but they will both agree on the magnitude of the difference between the coordinates of the endpoints (the length of the

[1]

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x)(x - x_0)^2 + \cdots$$

is the Taylor expansion of a function about the extremum x_0. Then by definition $f'(x_0) = 0$ and we have

$$f(x_0 + dx) = f(x_0) + \frac{1}{2}f''(x)(dx)^2 = f(x_0) + \mathcal{O}(dx)^2 = f(x_0).$$

[2] Refer to Exercise 1.1

rod), in the relativistic physics one takes the length of a moving rod in the Minkowski spacetime to be something everyone can agree upon. However, this length is not the usual Euclidean length, which is defined in three dimensions as

$$\sqrt{x^2 + y^2 + z^2},$$

but instead looks like the hyperbolic length

$$\sqrt{t^2 - x^2 - y^2 - z^2}.$$

Here is a more detailed definition.

3. The invariant interval

A derivative is the ***linear*** approximation to the function at some point. (Potential trip-up: the derivative of x^3 is certainly not linear) Two linear functions need not be proportional; $(ax+b)/(cx+d)$ may not be constant. However, the condition that the interval $= 0$ for light shows that $b = d = 0$, so that we have proved that $ds = ds'$.

Consider two events E_1, E_2 from two different frames K, K'. For simplicity, let E_1 be the origin in both frames. The interval in the K frame is

$$s^2 = t^2 - x^2 - y^2 - z^2$$

and the interval in the K' frame is

$$s'^2 = t'^2 - x'^2 - y'^2 - z'^2.$$

This is something everyone can agree on, so

$$s^2 = s'^2.$$

This is very useful when solving problems because you are usually given something like t, l, t' and you can solve for l' by equating intervals (let $l^2 = x^2 + y^2 + z^2$. Sometimes we will assume that motion is only along the x and x' axes, but this simplification has the same effect and will be used interchangeably).

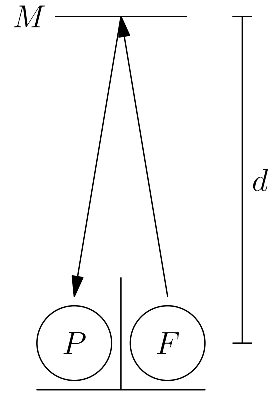
The interval can be interpreted as the length of a ***4-vector***, just as in Newtonian mechanics distance is the length of a displacement vector.

Exercises

2.1.*** A possible clock is shown in the figure below. It consists of a flashtube F and a photocell P shielded so that each views only the mirror M, located a distance d away, and mounted rigidly with respect to the flashtube-photocell assembly. The electronic innards of the box are such that when the photocell responds to a light flash from the mirror, the flashtube is triggered with a negligible delay and emits a short flash toward the mirror. The clock thus "ticks" once every $(2d/c)$ seconds when at rest.

(a) Suppose that the clock moves with a uniform velocity v, perpendicular to the line from PF to M, relative to an observer. Using the second postulate of relativity, show mathematically that the observer sees the clock ticking slowly.

(b) Suppose that the clock moves with a velocity v parallel to the line from PF to M. Verify that here, too, the clock is observed to tick more slowly, by the same time dilatation factor.



2.2.*** A rocket ship leaves the earth in the year 2100. One of a set of twins born in 2080 remains on earth; the other rides in the rocket. The rocket ship is so constructed that it has an acceleration g in its own rest frame. It accelerates in a straight-line path for 5 years (by its own clocks), decelerates at the same rate for 5 more years, turns around, accelerates for 5 years, decelerates for 5 years, and lands on earth. The twin in the rocket is 40 years old.

(a) What is the year on Earth?

(b) How far from Earth did the space ship travel?

4. Lorentz Transformations

In Newtonian physics, one would expect that the laws of physics don't change after rotating the coordinates. An even more boring statement is that the laws of physics shouldn't change with a translation parallel to one of the axes. In relativity, all these transformations are equally boring. But since there is an additional (temporal) dimension, one can rotated the plane t, x for example. This is not a pure spatial rotation like considered before, so it is called a ***boost***. To perform a boost from the K frame to the K' frame, apply the following rotation matrix on the 4-vector:

$$\begin{pmatrix} \cosh\psi & \sinh\psi & 0 & 0 \\ \sinh\psi & \cosh\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix}.$$

[1] From this we get the two equations

$$ct\cosh\psi + x\sinh\psi = ct', \quad ct\sinh\psi + x\cosh\psi = x'.$$

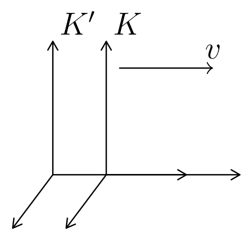
[1]The reason for using a hyperbolic rotation matrix will become clear later, in the velocity addition section.

We can check that this boost keeps the interval invariant. [1]

Example 2. Prove that boosts preserve lengths of 4-vectors.

$$
\begin{aligned}
c^2t^2 - x^2 - y^2 - z^2 =& c^2t'^2 - x'^2 - y'^2 - z'^2 \\
=& c^2t'^2 - x'^2 - y'^2 - z'^2 \\
=& (ct\cosh\psi + x\sinh\psi)^2 - (ct\sinh\psi + x\cosh\psi)^2 - y'^2 - z'^2 \\
=& (ct)^2\cosh^2\psi + x^2\sinh^2\psi + 2ctx\cosh\psi\sinh\psi \\
& - (ct)^2\sinh^2\psi - x^2\cosh^2\psi - 2ctx\sinh\psi\cosh\psi - y^2 - z^2 \\
=& (ct)^2(\cosh^2\psi - \sinh^2\psi) + x^2(\cosh^2\psi - \sinh^2\psi) - y^2 - z^2 \\
=& c^2t^2 - x^2 - y^2 - z^2
\end{aligned}
$$

In order to continue with the derivation, we need to find a way of eliminating the angles from the equations.
Consider the origin of the K system.



A rotation about the origin does not change the coordinates of the origin, so $x = 0$. Also, the origin of K as viewed in the K' frame is moving away at speed v, so $x' = vt'$. Substituting these values into the equations, we get:

$$ct\cosh\psi = ct', \quad ct\sinh\psi = vt',$$

and dividing the second equation by the first we get

$$\frac{v}{c} = \tanh\psi \implies \sinh\psi = \frac{v}{c}\cosh\psi.$$

Now using the classic hyperbolic identity $\cosh^2\psi - \sinh^2\psi = 1$, we get

$$\cosh^2\psi - \left(\frac{v}{c}\cosh\psi\right)^2 = 1 \implies \cosh^2\psi\left(1 - \frac{v^2}{c^2}\right) = 1 \implies \cosh\psi = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} = \gamma$$

as well as

$$\sinh\psi = \frac{v}{c}\gamma.$$

Now all that remains to do is sub these back into the original rotation equations, to get

$$ct' = \gamma\left(ct + \frac{xv}{c}\right), \gamma x' = \gamma(tv + x).$$

[1] The length of a vector is preserved under orthogonal matrices. Similarly, we define the length of a 4-vector to be its interval, so this rotation matrix is an orthogonal analogue.

These are the Lorentz transformation equations. Depending on the source, one will see these written slightly differently. Here are two other (quite trivially) different forms.

$$\begin{aligned} t' &= \gamma\left(t + \frac{xv}{c^2}\right) \\ x' &= \gamma(x + tv) \\ y' &= y \\ z' &= z \end{aligned} \tag{1}$$

and

$$\begin{aligned} t' &= \gamma(t + xv) \\ x' &= \gamma(x + tv) \\ y' &= y \\ z' &= z \end{aligned} \tag{2}$$

Remarks.

1. The above equations define a boost in the $x-$direction with ***rapidity*** ψ.

Exercise: derive the transformation equations associated with n temporal and m spatial dimensions.

5. Tensors and All That

Put

$$x^0 = ct, \quad x^1 = x, \quad x^2 = y, \quad x^3 = z$$

The components of a ***4-vector*** A are

$$A^0 = ct, \quad A^1 = x, \quad A^2 = y, \quad A^3 = z,$$

so that we have the notation $A = A^\mu = (A^0, A^1, A^2, A^3) = (ct, x, y, z) = (A^0, A^i)$, where Greek indices (such as μ, α, λ) run over 0,1,2,3 and Latin indices (such as i, j, k, l) run over 1,2,3. Also,

$$g(A, A) = (A^0)^2 - (A^i)^2,$$

so 4-vectors are ***Lorentz invariant***. Concrete definition of 4-vector: an object that transforms under a Lorentz

Exercises

4.1.* In this exercise, you will practice the formalism of tensors by constructing the electromagnetic field tensor which completely describes the electromagnetic field in spacetime. Note that you do not need to know any electromagnetic theory to complete this. ***Notation:*** $\vec{B}$ ***is the magnetic field,*** $\vec{E}$ ***is the electric field,*** $\vec{A}$ ***is the vector potential, and*** ϕ ***is the scalar potential. We denote the first component of*** $\vec{A}$ ***by*** A_x.

$$\vec{B} = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}, \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}, \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right).$$

$$\vec{E} = \left(-\frac{\partial A_x}{\partial t} - \frac{\partial \phi}{\partial x}, -\frac{\partial A_y}{\partial t} - \frac{\partial \phi}{\partial y}, -\frac{\partial A_z}{\partial t} - \frac{\partial \phi}{\partial z}\right).$$

Using covariant notation we have

$$\vec{B} = (-(\partial^2 A^3 - \partial^3 A^2), -(\partial^3 A^1 - \partial^1 A^3), -(\partial^1 A^2 - \partial^2 A^1))$$

and

$$\vec{E} = c(-(\partial^0 A^1 - \partial^1 A^0), -(\partial^0 A^2 - \partial^2 A^0), -(\partial^0 A^3 - \partial^3 A^0)).$$

The electromagnetic field tensor is

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

(a) What happens when the indices μ and ν are flipped?

(b) How do the diagonal components look like?

(c) Knowing the above two facts, how many elements of $T_{\mu\nu}$ do you need to calculate?

(d) Hence construct the electromagnetic field tensor in terms of the magnetic and electric fields.

PUEC 2017