Lecture Notes VI: Special Relativity

Cory Schillaci

July 25, 2013

1 The Principles of Relativity

Einstein's theory of special relativity rests on two principles. First is the principle of relativity, which says that physics should appear to function in the same way to all observers. Einstein did not introduce this idea. Classical physicists had embraced this principle because it seemed necessary for the existence of a satisfying theory. Suppose that physics depended on the reference frame from which it is described. Then we are forced to find a privileged observer, who can describe physics in some special way. All other observers will see the laws of physics modified.

Of course, there is no reason to believe that humans occupy this special position. Our planet is rotating on its axis and orbiting around the sun. The sun is orbiting about the center of the Milky Way galaxy. Naturally, the Milky Way is moving with respect to other galaxies and is unlikely to be at rest in this special reference frame either.

The principle of relativity makes perfectly good sense. It allows us to dispense with all of the above complications and be confident in our physical predictions from any one of the infinitely many equivalent reference frames.

The other is the constancy of the speed of light, which always travels at the velocity $c = 3.0 \times 10^8$m/s. Honestly, this should not be seen as its own principle. Classical electrodynamics[1] predicts the existence of waves which move at the speed of light, without specifying a natural frame in which this is true[2]. Classical physicists postulated that there must be some kind of fluid, which they called the ether, that carried these electromagnetic waves. The problem was that experiments, especially that of Michelson and Morley, ruled out the possibility that the ether moved with respect to the Earth. Going back to the first paragraph, it seems quite improbable that the Earth should somehow occupy such a special place.

[1] Electrodynamics is the study of the interactions between moving charges via electric and magnetic fields.

[2] For contrast, fluid dynamics predicts that air contains pressure waves (which we perceive as sound) that move a specific speed with respect to the air.

Einstein's insight was to assume that the laws of electrodynamics were complete, and that light followed these laws according to all observers. Although this was in some sense already predicted, we will see that the consequences are very strange. One can hardly fault 19th century physicists for not taking this prediction from electrodynamics seriously.

In summary, the two assumptions from which special relativity follows are

1. The principle of relativity: Physical laws work the same way for all observers.
2. Constancy of the speed of light: Electromagnetic waves (light) travel at the same speed according to all possible observers.

1.1 Galilean relativity

Classical mechanics already respected the principle of relativity in all inertial frames. As we discussed the first day, translations and rotations in time and space are good symmetries of the theory. Those translations basically correspond to redefining the origin of the coordinate system used to describe physical phenomena.

There is another symmetry, called *Galilean invariance*, which relates observers moving relative to one another. This means that a passenger in a car moving at a constant speed[3] on the highway should observe physics in the same way as someone standing on the side of the road trying to hitchhike.

Now the origins of the spatial reference frames are moving with respect to one another. Assume the hitchhiker uses the coordinates t, x, y and z while observing the car to move at a speed v in the $\hat{x}$ direction. If the passenger in the car wants to translate information into coordinates in which he is at rest, which we denote with primes, he must use the *Galilean transformations*:

$$\begin{aligned} t' &= t \\ x' &= x - vt \\ y' &= y \\ z' &= z \end{aligned} \tag{1}$$

These formulas assume that the frames correspond at $t = 0$.

Since this is a continuous symmetry, there is in fact an associated conservation law. Galilean invariance of classical mechanics implies uniform motion of the center of mass. In other words the center of mass of a system moves at a constant velocity. This conservation law is also a consequence of momentum conservation when mass is conserved, so it is usually not considered as a separate conservation law when learning Newtonian mechanics.

[3]Observers must be *inertial*, meaning that they are not accelerating.

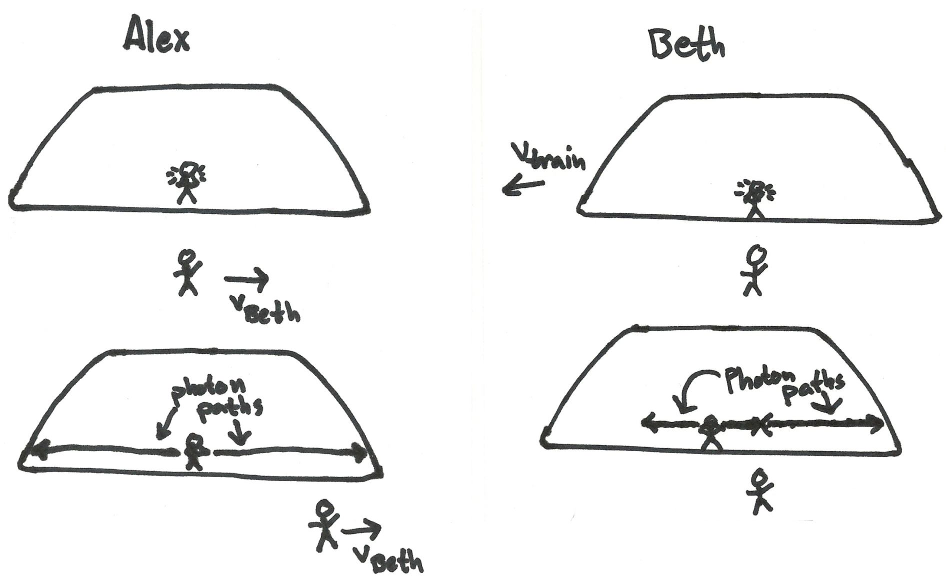

Figure 1: Alex's camera flash sends photons forward and back. On the left is the story from Alex's perspective. The right shows the events in Beth's frame.

Quick Question: According to the passenger, what is the velocity $\vec{v}$ of the hitch-hiker?

1.2 Consequences of the second principle

Now we will see some of the consequences of adding the second principle. Imagine that one observer, call him Alex, is riding in a maglev train[4] moving at a speed $v/c = \beta$ with respect to the ground. His friend Beth is waiting along the track to wave at him as he passes. Right as he passes, he uses his super fast reflexes to take a picture and the flash on his camera goes off. Let's consider what happens to the light from the flash. Assume he was right in the center of the train car for simplicity.

Remember that there is nothing special about Beth just because she is at rest with respect to the Earth. According to Alex, the train is stationary and Beth is blasting by along with the Earth. The principle of relativity tells us that both of them can assume that they are at rest, then describe physics using exactly the same rules. In particular, each of them will see light traveling at a speed c.

Figure 1 shows this story playing out according to each observer, with a special emphasis on when the photons from the flash strike the front and back of the train car. According to Alex, the light from his flash clearly hits both ends of the train at the same time. But what does Beth observe? Because the train is moving, the

[4]The speed record for a maglev train is about $5 \cdot 10^{-7}c$. As far as I can tell, the fastest that humans have travelled with respect to one another was during the Apollo 10 re-entry at roughly $3.7 \cdot 10^{-5}c$.

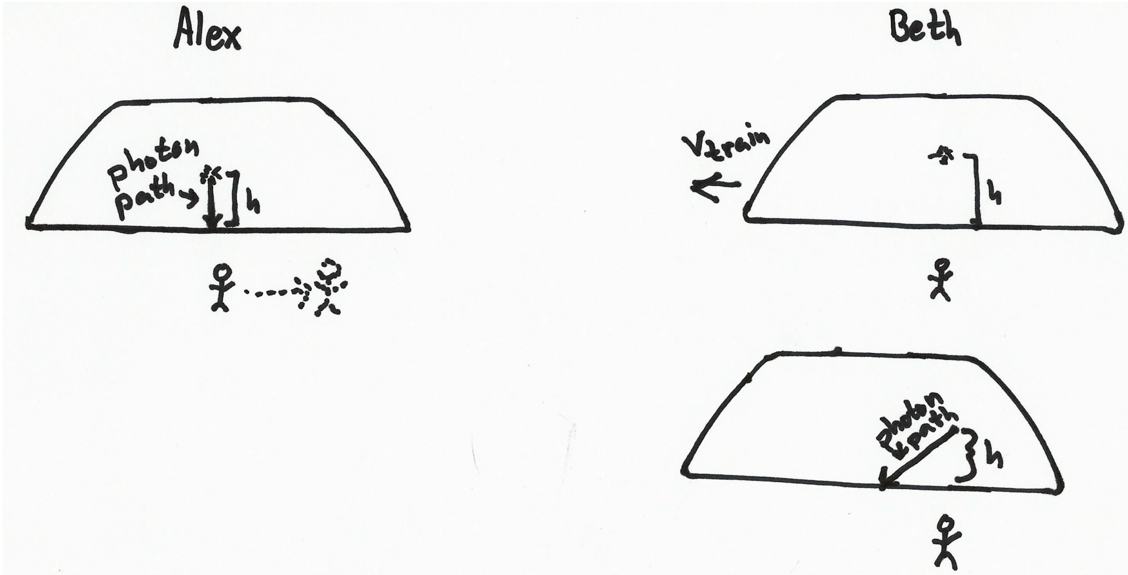

Figure 2: Alex's camera flash sends photons down to his feet. On the left is the story from Alex's perspective. The right shows the events in Beth's frame.

light hits the back of the train first! This implies that events which are simultaneous according to one observer are NOT necessarily simultaneous to another.

Quick Question: What would be different about this story if Alex had instead thrown two tennis balls (one forward and one backward) on a much slower train? Assume Galilean relativity.

Quick Question: Suppose instead we had two photographers stationed at the front and back of the train who simultaneously took a picture of Beth in their reference frame. What would Alex and Beth observe now?

This first example demonstrates a qualitative difference between Galilean and Special relativity. Consider now the photons from the flash that travel towards his feet on the floor of the train. Assume the camera was at a height h when the picture was taken (I won't prove it to you, but both Alex and Beth agree on this height). This situation is depicted in Figure 2.

Alex sees the light travel straight down, which takes an amount of time

$$\Delta t_A = \frac{h}{c}. \tag{2}$$

According to Beth the photons which hit near Alex's feet are traveling at an angle, and they have to travel a greater distance. The total distance she sees is given by the pythagorean formula as $d^2 = h^2 + (v\Delta t_B)^2$. The time it takes to travel this far is,

$$\Delta t_B = \frac{\sqrt{h^2 + (v\Delta t_B)^2}}{c}, \tag{3}$$

which can be solved for Δt_B to give

$$\Delta t_B = \frac{1}{\sqrt{1-\beta^2}}\frac{h}{c} = \frac{1}{\sqrt{1-\beta^2}} t_A. \tag{4}$$

According to Beth, the time between the flash going off and the light striking Alex's feet is greater than what Alex observes by a factor which we will encounter so frequently we give it a special name,

$$\gamma \equiv \frac{1}{\sqrt{1-\beta^2}}. \tag{5}$$

Because $0 \leq \beta^2 \leq 1$, we will always have $\gamma \geq 1$ (the equality holds only when $\beta = 0$).

This phenomenon is known as *time dilation*. If an observer B watches the clock that an observer A is holding, it appears to tick more slowly than any clocks at rest with respect to B. Alternately, the measured duration of any event is always shortest in its rest frame.

Quick Question: If A observes B's personal clock, do they appear to be faster or slower than the clocks at rest with respect to A? Is this consistent with our principles?

Now assume that the front of the train has a nice clean window, and some of the light that strikes it travels back to Alex. If we call the length of the train $2L$, Alex observes the duration of this process to be

$$\Delta t_A = \frac{2L}{c} \tag{6}$$

Beth's observations are more complicated because the ends of the train are moving. The time for the light to strike the front of the train is

$$\Delta t_1 = \frac{L + v\Delta t_1}{c} \tag{7}$$

while the return trip is shorter,

$$\Delta t_2 = \frac{L - v\Delta t_2}{c} \tag{8}$$

Solving these for the times and then combining them into the total time we get:

$$\Delta t_B = \Delta t_1 + \Delta t_2 = \frac{L}{c-v} + \frac{L}{c+v} = 2\gamma^2\frac{L}{c} = \gamma^2 \Delta t_A \tag{9}$$

and using the constancy of the speed of light we see that

$$\frac{L}{\Delta t_A} = \frac{1}{\gamma^2}\frac{L}{\Delta t_B} = c \tag{10}$$

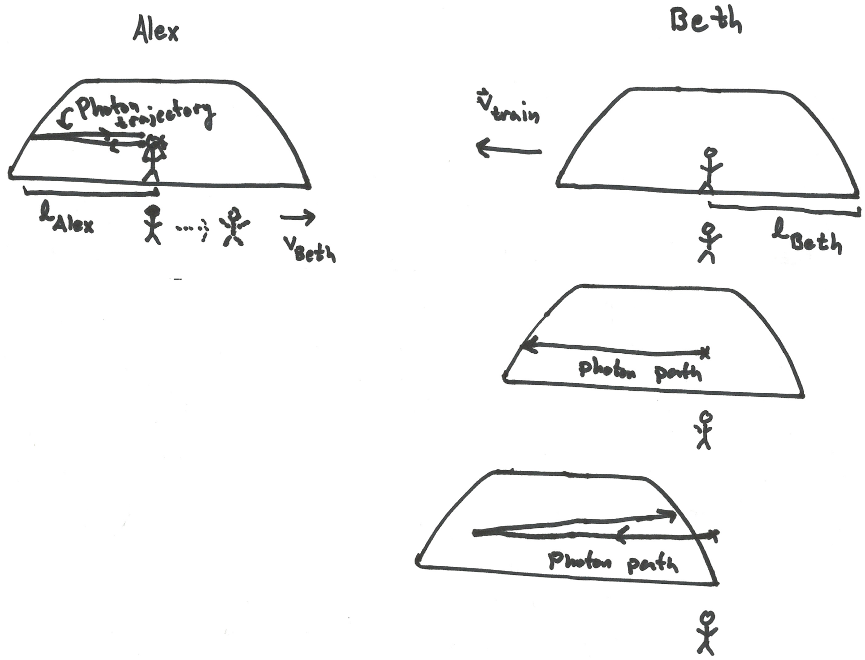

Figure 3: Alex's camera flash sends photons towards the front of the train, which then reflect back. On the left is the story from Alex's perspective. The right shows the events in Beth's frame. The photons are drawn at a small angle to make the situation clear, but the angle should be ignored.

One of these factors of γ comes from the time dilation effect of (4). With that adjustment, we are left with

$$L = \frac{1}{\gamma} L \tag{11}$$

This clearly can't be true (except for the very special case $\gamma = 1$). What went wrong? Apparently Alex and Beth measure the length of the train differently. If we allow that they might see the train as different lengths, we he have derived the formula for the effect called *Lorentz contraction.*

$$L_B = \frac{1}{\gamma} L_A \tag{12}$$

In other words, moving objects appear shorter than objects at rest! An object is longest in its own rest frame.

We can rephrase all of the above using the *Lorentz transformations* for observers with relative motion in the $\hat{x}$ direction,

$$\begin{aligned} t' &= \gamma\left(t - \frac{v}{c^2}x\right) \\ x' &= \gamma(x - vt) \\ y' &= y \\ z' &= z \end{aligned} \tag{13}$$

2 Minkowski geometry

You may have heard that time is the fourth dimension, or maybe read about the unified concept of spacetime. The Lorentz transformations of the previous section certainly suggest that time can no longer be thought of distinctly from spatial degrees of freedom. We can see this far more elegantly by thinking of special relativity geometrically.

Special relativity requires that an observer record four properties of an event: the time and the location in the three spatial dimensions. We can organize this into a *four-vector* describing the event,

$$x^\mu = \begin{pmatrix} V^0 \\ V^1 \\ V^2 \\ V^3 \end{pmatrix} = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{14}$$

To find when and where this occurred according to a second observer moving at velocity $\vec{v} = c\beta\hat{x}$ with respect to the first, we need to use the Lorentz transformations from (13). For the four-vector, this just amounts to matrix multiplication! In the new observer's reference frame[5],

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma(ct - \beta x) \\ \gamma(x - \beta ct) \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{15}$$

Quick Question: What would the Galilean transformation look like in this notation?

One question we haven't asked yet is how to define the distance between two events.

Quick Question: What is the Galilean meaning of the distance between points? Why doesn't it make sense to use this old definition in special relativity?

For a quantity to be a distance, it should be the same for all observers. This implies that we need a meaning of distance that doesn't change under the Lorentz transformation. Let's interpret the four-vector of (14) to be pointing from the origin to an event. Up to an overall sign, the only possible *Lorentz invariant* choice for the distance between the origin and the event is

$$|x^\mu|^2 = V \cdot V = -(ct)^2 + x^2 + y^2 + z^2. \tag{16}$$

This is the length of the vector. This distance is called the *spacetime interval*, and often denoted Δs^2. Note that we don't take the square root.

[5] I am assuming that the two reference frames line up at the origin, i.e. at $t = t' = x = x' = y = y' = z = z' = 0$.

Quick Question: The quantity in (16) is not always positive. Why is that ok here, but not in the case of quantum mechanics?

I should also mention that the time experienced by an observer, called the *proper time*, is given by $(\Delta\tau)^2 = -|x^\mu|^2$. This is just the length of their own path in their reference frame, where the spatial displacements are all zero. Since the length is Lorentz invariant, the length of a path is always the same to all observers.

Quick Question: How can I be sure that $(\Delta\tau)^2$ is always positive?

Let's go ahead and define a four-vector as any vector whose norm (as defined by (16)) is invariant under the Lorentz transformation. Another useful example is the *four-momentum*,

$$p^\mu = \begin{pmatrix} E/c \\ p_x \\ p_y \\ p_z \end{pmatrix}. \tag{17}$$

where $p = \gamma m v$.

Quick Question: If we define the mass of a particle to be $|p^\mu|^2 = -m^2c^2$, what is E in terms of the mass m and three-momentum $\vec{p}$?[6]

The four-momentum has some sweeping implications. Conservation of energy and momentum were distinct concepts in Newtonian mechanics. In relativistic mechanics they are both consequences of the conservation of the four-momentum due to translation invariance in spacetime. We also notice that Lorentz invariance requires the existence of rest energy, encapsulated in Einstein's famous equation $E = mc^2$. You can now see that this is the whole story only if the normal momentum is zero, which is possible only for an object at rest. Nothing requires that mass be conserved, and so it isn't. Nature still conserves energy, so any mass which "disappears" is converted into another form. We will see more about how this works when we discuss the fundamental particles.

[6]Actually, a more careful treatment would first define the four-momentum as a derivative of the position then identify the time component as the relativistic energy.