Special Relativity Notes (Reference)

Bryan Lu

7 January 2022

1 History (5-8 min)

Developments leading up to the theory of special relativity:

- *Light is an electromagnetic wave.* James Clerk Maxwell publishes his equations for electromagnetism (1864) and notices as a consequence that the speed of electromagnetic waves is very close to value of light ($c = 299792458$ m/s), and Heinrech Hertz (1887) is able to produce electromagnetic waves as predicted by Maxwell's theory.
- *Light waves propagate in a medium.* Known since the early 1800s that light propogates as a wave, and therefore must propogate in some medium. Proposed medium was called the *luminiferous aether*. However, various experiments to detect the presence of the aether (in particular, its motion) found no results – most notably the *Michelson-Morley experiment* (1887) which measured c in perpendicular directions and found no significant difference.
- *Strange phenomena.* Lorentz and Poincaré, creating a theory of the (motionless) aether (1892-1895, 1905) that mirrors very closely the results of the modern theory of special relativity, but working with an absolute frame of reference, the one of the aether.

In 1905, four papers by Albert Einstein were published, one of which proposed sweeping changes to classical mechanics as grounded in Newton's work, as it established a new framework for the nature of space and time itself. This paper rejected the theory of the aether altogether and would grow to become the foundation for the *special* theory of relativity that we know today.

2 Main Phenomena (12-15 min)

Einstein's conclusion was to accept the following two statements to build a new theory of dynamics, based on experimental evidence.

1. (Principle of Relativity) The laws of physics are invariant in all inertial frames of reference.
2. (Constancy of the Speed of Light) The speed of light is the same for all observers, regardless of the motion of the object or the observer.

Time Dilation

Consider a clock that measures time using light, i.e. it counts one "tick" as the time it takes for a light particle to traverse a certain distance. Suppose we have two identical clocks, which track time by stating that one tick is the amount of time it takes for a light particle to travel between two parallel reflecting mirrors a distance h apart

(from one mirror, to the other, and back). One of these clocks is set moving at speed v in the positive x-direction on a train, and another clock is stationary. Now, suppose we have Sally, who is on the train, and Joe, who is a stationary observer, observing these clocks. Both of the observers are in inertial reference frames. Let's consider the point of view of Joe.

Let's first consider Joe's clock. Here, the light travels straight upward, traversing a total distance of $2h$. **The time Joe's clock (in Joe's reference frame) measures as "one tick" is then** $t_J = \frac{2h}{c}$.

For Sally's clock, on the other hand, has the light take an angled path due to the horizontal motion of the train. The light's total speed is c, and the component of the speed in the x-direction is v, so the speed in the y-direction is $\sqrt{c^2 - v^2}$. Then the time Sally's clock (in Joe's reference frame) registers as "one tick" is $t_S = \frac{2h}{\sqrt{c^2 - v^2}} = \gamma t_J$, where $\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}$ is the *Lorentz factor*. (Note that as $v \leq c$, $\gamma \geq 1$.)

Therefore, the moving clock in Sally's frame as seen by Joe is ticks by a factor of γ slower than his! This is the phenomenon of **time dilation: relative to a stationary observer, moving clocks tick slower!**

Note that we can run the same analysis from Sally's point of view: Sally sees her clock tick γ times faster than Joe's clock!

Length Contraction

Let's modify the construction now. Here, we'll let Sally be the "stationary observer" in her frame of reference, and Joe will be moving relative to her. Consider a light beam traveling horizontally from the front of the train to the back and then returning to the front. **Suppose the length of the train is L_S in Sally's frame**. According to Sally, the time it takes for the light to traverse this path is $t_S = \frac{2L_S}{c}$.

In Joe's reference frame, the sides of the train are traveling in the positive x-direction with speed v. On the path from the front to the back of the train, the back of the train is moving toward the beam of light at speed v, so the time it takes to for the light to travel from the front to the back is $\frac{L_J}{c+v}$, where L_J is the length of the train in Joe's frame. On the second leg of the light's path, the front of the train is moving away at speed v, from the particle, so the time on this leg is $\frac{L_J}{c-v}$. The total time according to Joe is then

$$t_J = \frac{L_J}{c - v} + \frac{L_J}{c + v} = \frac{2cL_J}{c^2 - v^2}.$$

From the previous thought experiment, we have $t_J = \gamma t_S$, as the time interval measured by a moving observer is a factor of γ larger. This gives us

$$\frac{2cL_J}{c^2 - v^2} = \gamma \frac{2L_S}{c} \Rightarrow L_J = \frac{(c^2 - v^2)}{\sqrt{1 - v^2/c^2}} \cdot \frac{L_S}{c^2} = \sqrt{1 - \frac{v^2}{c^2}} L_S = \frac{L_S}{\gamma}.$$

Therefore, the length of the train that Joe observes is a factor of γ shorter than the actual length of the train as measured in Sally's frame! This is the phenomenon of **length contraction: moving lengths are shorter by a factor of γ in the frame of a stationary observer (in the direction of motion)!**

One thing to note, however: lengths still remain normal along perpendicular directions to the motion. Suppose the train is going to pass through a tunnel, where if height of the train in Sally's frame at rest is H_J, the height of the tunnel in Joe's frame is also H_J. Suppose without loss of generality that moving lengths decrease in the perpendicular direction relative to a stationary observer. Then, Joe will see the train's height decrease in his reference frame, so the train will fit in the tunnel, but Sally will see the tunnel's height shrink, so the train will not fit in the tunnel. Both of these things cannot be simultaneously true! As such, we see that length contraction only happens in directions parallel to the motion, but **not** in any perpendicular directions!

3 Lorentz Transformations (10-12 min)

Using the above phenomena, we would like to figure out how to describe transformations between different inertial frames of reference. After all, as indicated by our thought experiments, time and space are intimately connected, as moving through space corresponds to a slowing of time, and a shortening of observed lengths. As such, these transformations between different inertial reference frames have to have some "mixing" of the space and time coordinates.

Note that along with caring about where things are in space, we also have to care about where things are in time, so we would like to use vectors that represent the *spacetime* coordinates of an object. This vector must have *four* components, time and the three spatial dimensions, but if we're sticking with motion in one dimension as we are now, we can just think about the time t and position x.

When we derive the nature of the transformation that can transform between different inertial frames of reference, we assume that this transformation has to be *linear*. What I mean here is that if we have a displacement vector of sorts describing the differences in space and time between two events $\begin{bmatrix} \Delta t \\ \Delta x \end{bmatrix}$, and if this same displacement vector viewed in another reference frame is $\begin{bmatrix} \Delta t' \\ \Delta x' \end{bmatrix}$, then $\Delta t'$ can only depend linearly on Δt and Δx, and same for $\Delta x'$. If this were not the case, then objects moving at constant velocity would accelerate in others, introducing strange fictitious forces and things... and we don't want that at all.

With the property that the transformation is linear, we can declare that our transformation between frames can be modeled as a 2×2 matrix $\Lambda(v)$, acting on a vector of our coordinates:

$$\begin{bmatrix} \Delta t' \\ \Delta x' \end{bmatrix} = \begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix} \begin{bmatrix} \Delta t \\ \Delta x \end{bmatrix}$$

Clearly, our matrix can only be a function of the velocity v that we're trying to boost our second coordinates into, which helps somewhat... (This follows from isotropy and homogeneity of space.)

Here, let's use our experiments from above. Suppose our matrix represents the transformation from Sally's coordinates (moving) into Joe's coordinates (stationary). Let's use our time dilation thought experiment, and consider the photon in the clock in Sally's frame. Recall that in Sally's frame, the time it takes for the photon to return to her is $\Delta t = \frac{2h}{c}$, but Joe sees the photon return to Sally in time $\Delta t' = \frac{2h}{c} \gamma = \gamma \Delta t$, in which the photon has moved to the right by $\Delta x' = v \Delta t' = \gamma v \Delta t$ due to Sally having moved forwards on the train relative to Joe in that time. This allows us to conclude the form of the entries in the left-hand column:

$$\begin{bmatrix} \Delta t' \\ \Delta x' \end{bmatrix} = \begin{bmatrix} \gamma & a(v) \\ \gamma v & b(v) \end{bmatrix} \begin{bmatrix} \Delta t \\ \Delta x \end{bmatrix}$$

In order to conclude the form of the other two entries, notice that simply flipping v to $-v$ will give the boost that represents the transformation from Joe's coordinates (stationary) to Sally's coordinates (moving). This should also correspond to the matrix's inverse, as doing these two transformations one after the other should give the identity! Hence, we require that

$$\begin{bmatrix} \gamma & a(v) \\ \gamma v & b(v) \end{bmatrix} \begin{bmatrix} \gamma & a(-v) \\ -\gamma v & b(-v) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

The equations useful for us will be

$$\gamma^2 - \gamma v a(v) = 1 \quad \gamma a(-v) + a(v) b(-v) = 0$$

Solving the first:

$$\gamma v a(v) = \frac{1}{1 - \frac{v^2}{c^2}} - 1 = \frac{v^2}{c^2(1 - v^2/c^2)} \implies a(v) = \frac{\gamma v}{c^2}$$

Plutting into the second:

$$-\frac{\gamma^2 v}{c^2} + \frac{\gamma v}{c^2} b(v) = 0 \implies b(v) = \gamma$$

This gives us the linear transformation that takes a frame at rest and sends that frame moving v to the right, or in other words it takes an event in the moving coordinate system and sends it to the coordinates of an event in the stationary coordinate system.

Unfortunately, the dimensions of the matrix are sort of all over the place - we got entries that have units [m/s] and [s/m], as well as dimensionless constants. To rectify this, we fold in the speed of light into the time component of our vectors, in order to make everything have the same dimension:

$$\begin{bmatrix} \Delta t' \\ \Delta x \end{bmatrix} = \begin{bmatrix} \gamma & \frac{\gamma v}{c^2} \\ \gamma v & \gamma \end{bmatrix} \begin{bmatrix} \Delta t \\ \Delta x \end{bmatrix} = \begin{bmatrix} \gamma \Delta t + \frac{\gamma v}{c^2} \Delta x \\ \gamma v \Delta t + \gamma \Delta x \end{bmatrix} \implies \begin{bmatrix} c\Delta t' \\ \Delta x' \end{bmatrix} = \begin{bmatrix} \gamma c \Delta t + \frac{\gamma v}{c} \Delta x \\ \frac{\gamma v}{c} \cdot c \Delta t + \gamma \Delta x \end{bmatrix} = \begin{bmatrix} \gamma & \frac{\gamma v}{c} \\ \frac{\gamma v}{c} & \gamma \end{bmatrix} \begin{bmatrix} c\Delta t \\ \Delta x \end{bmatrix}$$

Our newly-derived matrix is called the **Lorentz transformation** $\Lambda(v)$[1], with all entries dimensionless:

$$\Lambda(v) = \begin{bmatrix} \gamma & \frac{\gamma v}{c} \\ \frac{\gamma v}{c} & \gamma \end{bmatrix}$$

In general, we consider **4-vectors**, denoted x^μ, with four components denoted like so:

$$x^\mu = \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix}$$

where the superscripts aren't exponents, they're indices. The placement of the index is (for right now) not super relevant, but you'll also see the index placed downstairs as a subscript. x^μ can refer to the vector as a whole, but also μ can take on values of $0, 1, 2, 3$ – so specific values of μ reference specific components of the vector. This is sort of a confusing consequence of Einstein summation convention, but I hope it's not too hard to get used to.

We can extend the Lorentz transformation to act on 4-vectors, noting that nothing happens to perpendicular directions as noted above, so extending the Lorentz transformation should naturally give

$$\Lambda(v\hat{\mathbf{i}}) = \begin{bmatrix} \gamma & \gamma\frac{v}{c} & 0 & 0 \\ \gamma\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

We can similarly do this for the other two perpendicular directions:

$$\Lambda(v\hat{\mathbf{j}}) = \begin{bmatrix} \gamma & 0 & \gamma\frac{v}{c} & 0 \\ 0 & 1 & 0 & 0 \\ \gamma\frac{v}{c} & 0 & \gamma & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \Lambda(v\hat{\mathbf{k}}) = \begin{bmatrix} \gamma & 0 & 0 & \gamma\frac{v}{c} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \gamma\frac{v}{c} & 0 & 0 & \gamma \end{bmatrix}$$

Velocity Addition

Now, let's study the problem of how to add parallel velocities in special relativity. Suppose that we have a rocket moving at speed u in the x-direction relative to some stationary observer (the lab frame). The rocket launches a torpedo at speed v relative to the rocket. Let's see if we can find the speed of the torpedo in the lab frame, w.

In a time t, the time-position vector of the torpedo in the rocket frame is $\begin{bmatrix} ct \\ vt \end{bmatrix}$. Applying the Lorentz

---

[1]If you look this up online, you'll see that most online resources have the off-diagonal terms negated. Usually people take this matrix with off-diagonal matrices negated as the standard "Lorentz transformation," but this version is what I initially learned/what I'm used to.

transformation, we find that in the lab frame the vector is
$$\begin{bmatrix} \gamma ct + \gamma \frac{u}{c} vt \\ \gamma ut + \gamma vt. \end{bmatrix}$$
The speed can be found by dividing the position by the time:
$$w = \frac{\gamma u + \gamma v}{\gamma + \gamma \frac{uv}{c^2}} = \frac{u + v}{1 + \frac{uv}{c^2}}.$$
This is the velocity addition formula in special relativity – note how it is not just $u + v$ anymore! Also, no matter how close u and v come to the speed of light, w never exceeds c, which is exactly as intended :)

4 Spacetime Diagrams and World Lines (INCOMPLETE)

One helpful tool for showing the motion of relativistic particles is the **spacetime diagram**, also known as the **Minkowski diagram**. The spacetime diagram shows the ct-x plane. An event, a point of fixed x and ct, is shown by a point on the diagram. The path of a particle on the spacetime diagram is called its **world line**.

The world line of light is always a 45-degree line in any spacetime diagram because the speed of light is constant in every frame. If we add another spatial dimension, light rays define a light cone. Only events within the light cone surrounding an event (a point on the spacetime diagram) can affect or be affected by that event.

We can show transformations between different reference frames on spacetime diagrams.

5 Loss of Simultaneity (somewhat incomplete)

We'll use all of the results so far to discuss a surprising result: the relativity of simultaneity! Events that may happen simultaenously in one frame may not be simultaneous in another. Here's how.

Suppose Sally is in the middle of a traincar of length $2L$ moving at speed v in the $+x$-direction, and Joe is standing stationary on the train platform. Let's suppose a flash of light is emitted at the center of the train just as Joe and Sally pass each other, and let's suppose the emission of the light is at the origin in spacetime for both Sally and Joe, without loss of generality.

According to Sally, then, the events of the pulses hitting the front and back of the train have spacetime coordinates of $x_F^\mu = \begin{bmatrix} L \\ L \end{bmatrix}$ and $x_B^\mu = \begin{bmatrix} L \\ -L \end{bmatrix}$, respectively. We can apply a Lorentz transformation, and see that these events transform to $x_F'^\mu = \begin{bmatrix} \gamma & \frac{\gamma v}{c} \\ \frac{\gamma v}{c} & \gamma \end{bmatrix} \begin{bmatrix} L \\ L \end{bmatrix} = \begin{bmatrix} L\gamma \left(1 + \frac{v}{c}\right) \\ L\gamma \left(1 + \frac{v}{c}\right) \end{bmatrix}$ and $x_B'^\mu = \begin{bmatrix} \gamma & \frac{\gamma v}{c} \\ \frac{\gamma v}{c} & \gamma \end{bmatrix} \begin{bmatrix} L \\ -L \end{bmatrix} = \begin{bmatrix} L\gamma \left(1 - \frac{v}{c}\right) \\ L\gamma \left(-1 + \frac{v}{c}\right) \end{bmatrix}$, respectively, in Joe's frame. Note that because the top coordinate in each of these vectors is not the same, we see that the events are not simultaneous in Joe's reference frame! Intuitively, this is because the train is moving forward, so a light pulse should hit the back of the train before the other light pulse hits the front. This is, in fact, what the calculations show.

We can see this also on a spacetime diagram!