An Overview of Special Relativity

Amal Bumbia

May 2025

Introduction and Preliminary Definitions

Special relativity is essential to understand as it pertains to a large swathe of physics topics. Unfortunately, it is not the most intuitive. The goal of these notes is to introduce the essentials in the least confusing way possible.

A basic observation is that whether you are in a moving vehicle or on the ground, the laws of physics remain the same. Such external parameters have no impact on the physical phenomena of a system. However, a person from within the moving vehicle and a person on the ground witnessing the same event will still disagree in some of their measurements and calculations regarding simple problems. The language of special relativity needs to be built up before we can discuss this further. Let's begin with our first key definition.

Key Definition

An *event* is an occurrence at a specific point of spacetime which we'll say for now is a four-dimensional set [x,y,z,t] (four-vector containing not only space but also a time dimension)

Key Definition

An *invariant* in this context is a quantity that does not change despite some particular transformation.

The exact values of the elements in our four-vector depend on initial conditions, but we can find invariants relating to our four-vector that can better aid us in describing the system and how it changes with reference frames. Something to keep in mind is that while we are proposing space and time as being part of the same entity, you can move back and forth in space but NOT in time. We'll come back to this later.

Key Definition

The motion of a free particle can be parameterized by a series of events, a *worldline*, which can get arbitrarily more detailed as the number of events measured increases.

Key Definition

At any event we can define a *light cone* representative of all possible paths light rays can take as they pass through the event.

To make sure we are comfortable with these ideas, let's introduce a visual.

Clock Lattice Visualizer

Consider a 3D lattice of clocks with equidistant spacing so that the position of an event can be described along any point relative to a defined spatial origin; an origin event occurs at t = 0 and all clocks on the lattice are synced to t = 0 at said origin event. When another event occurs at some spatial position of the lattice, its time is measured by the nearest clock. How do we sync these clocks? We set them with respect to a master clock so that they may measure the flow of an 'absolute' time.

Key Definition

A *reference frame* (RF) is just some framework where you can consistently sync all "clocks" that reside within it, meaning that an observer can accurately make and interpret measurements.

What setting are you measuring something relative to? Recall Newton's first law: *An object at rest remains at rest and an object in motion remains in motion unless acted upon by an opposing force. This is the law of inertia (resistance to motion).*

In an inertial RF (IRF), Newton's first law is true at all points, while in a non-inertial RF, fictitious forces are experienced (this happens in rotating or accelerating frames). Inertial frames move at a constant velocity relative to each other. We can define frames as locally inertial. We cannot define an IRF in the presence of a gravitational field. An RF at rest on Earth is non-inertial, non-rotating, and freely falling — basically, an ideal inertial frame isn't possible in practice, so we satisfy ourselves with locally inertial frames.

We don't exactly care about whether the reference frame is moving; we care about how it's moving and whether that motion can change the *observed* physics. We care to distinguish inertial and non-inertial frames. This leads us to the core idea:

The laws of physics are invariant between inertial reference frames.

Galilean Transform

Consider an event, say, a balloon popping high in the air. Two observers are viewing this event. One is running on the ground below and the other is on a tree next to the balloon. Each is viewing the event from a different RF. Both will say "Yes, the balloon popped at time t", but the position at which they perceived the event will differ.

We can pick which RF we want to define as our 'home frame' (I will use this language because while one can easily view the person in the tree as the 'stationary frame' and the person running as the 'moving frame', these labels are arbitrary). We can label the other frame as the 'other frame'. If we want to understand the position of our balloon-popping event where the running person, the 'other frame', saw it from with respect to the position the observer in our 'home frame' saw it from, we can use the *Galilean Transform*

$$x' = x - vt$$

where the v is the velocity of the observer in the 'other frame' and t is the time; if we recall from basic kinematics, $x = vt$, and in our case, x denotes the position in the home frame while vt is the change in position due to the velocity of the other frame. The time derivative of the above equation gives $v' = v - V$ in the direction of the other frame's motion (in our example, the runner is moving along the x direction, hence why the x position changes by the transform while the y and z positions remain the same between reference frames — so the y and z components of the velocity also remain the same). Upon taking another time derivative, we see that the acceleration as measured in both reference frames remains unchanged.

What we hope to answer...

1. How do we describe events by their position and time as measured relative to some origin?
2. How do those measurements change based on the RF an observer views the event from?
3. How can we relate the measurements between reference frames?

The Galilean Transform is the first step toward understanding these questions, but it doesn't always work. As we'll see, it's a special case of a more powerful set of transformations, the *Lorentz Transform.*

"But the Galilean Transform makes sense!'", you squawk like a malignant parakeet of naive disposition. Yeah, well it doesn't work. Light ruins that.

We've all heard that light speed is the fastest any particle can go — $c = 3E8$ is our limit. Later, we'll define c as 1 for a few reasons. But here's the thing, the Galilean Transform contradicts this. It says

$$v' = v - V$$

where V is the velocity of the other frame with respect to the home frame. How can this hold when light must be c in all frames? There can be only two possible conclusions. Either the Galilean Transform is wrong OR our understanding of light, and by extension, Maxwell's Equations, is deeply wrong. Choose your pick — electromagnetism as we know it, or...the Galilean Transform.

Like you, the great scientists of the time could simply not part with this simple transform. Instead, they considered the aether. I really don't feel like talking about the aether because truth be told, all we need to know is that the whole idea was wrong (read: Michelson-Morley experiment). After our academics finally decided that yes, the speed of light was constant and no, they did not have to scrap electromagnetism, it was obvious that the way they were relating reference frames needed to be tweaked to account for a constant light speed. The consequence of this was that we could no longer assume an absolute time between reference frames. If I shined a flashlight from point A to point B, then the times at points A and B would be different, and multiplying the difference between those times by the speed of light would give the distance traversed by the light beam. If the calculated distance matches the known distance between the points, then we can say that the 'clocks' at each point are synced, this time accounting for a constant c. By Moore, we can define two clocks in an IRF as being in sync if the time interval measured between the clocks as a light bean travels the distance is equal to their separation. This also gives us another fun tool — the ability to measure distance with time. Now you see why space and time can be considered part of the same entity.

Before we can get into the Lorentz Transform, we need to be able to measure the time interval between two events.

"Just subtract the ending time from the starting time!'", you squawk, rolling your beady little eyes. Eh, almost. We have three different types of time interval measurements to deal with.

Coordinate Time $\triangle t$

Consider two events, A and B. The time between both events can be measured by taking the difference between the time measured by the clock at event A and the clock at event B. The events occur in an IRF; note that we are using two synced clocks to determine the time interval between the events.

Definition

Coordinate time is the time measured between two events either by a pair of synced clocks at rest in a given IRF or by a single clock (if both events occur by the same clock).

$$\triangle t_{BA} = t_B - t_A$$

A key point to note is that coordinate time is frame dependent. We can illustrate this by considering two events occurring at the same time in different positions in a specific RF. Think of two fireworks exploding at different points in the sky. Now let's look at this same occurrence from the perspective of another inertial RF moving at a constant velocity, like a non-accelerating car. The time each event occurs as measured by this observer will no longer be equal. If both fireworks go off at some time t as measured by someone in the 'other frame', then someone in the car, our 'home frame' will perceive the clocks at each event as being out of sync since $\triangle t \neq 0$ from the perspective of the driving observer. If each observer syncs clocks in their own RF, they will perceive clocks in other RFs to be out of sync.

Also, coordinate differences between two events are frame dependent. This goes for both time and space. To put this into perspective, let's consider a very cool garden. We have a beehive in the center of the garden because honey is tasty. Somewhere nearby is a little cafe that makes its treats with honey from the beehive, and due to high demand, there's a straight path from the hive to the cafe. After all, we want the honey to get to the cafe as fast as possible and the shortest distance between two points is a line (in flat space at least). The rest of the garden is a series of winding paths and bridges weaving over ponds and patches of various flowers. All paths end up connecting the hive and cafe since those are the main attractions.

Because this garden is absurdly huge, all maps are overlayed with a coordinate grid so we can describe the positions of the hive, cafe, and flower patches with respect to a convenient system of our choosing. One day, the maps are misprinted, and the grid overlay is rotated by some small angle. Our previous coordinates are no longer accurate since the system is different. We can clearly see that the coordinate values depend on the grid being used. However, what doesn't change is the fact that the distances between the points haven't been altered just because the grid is different. We don't need to refer to the coordinate system to measure path lengths.

If space has this kind of unchanging property between fixed points regardless of the coordinate system, it stands that there must be a way to measure time intervals independent of reference frames.

Proper Time $\triangle T$

"Who cares about gardens! What other ways can you measure time intervals?", you squawk impatiently. Eat grass! Let me have fun.

Proper time is our frame-independent quantity obtained by using a clock present at both events, like how we could use a single ruler or tape to measure path length. Since the same clock will be there at each event as it occurs, there is no contradiction regardless of what RF an observer is in.

> **Definition**
>
> *Proper time* is the time interval between two events as measured by any clock present at both.

Though proper time is frame independent, it does depend on the worldline taken by the clock.

To illustrate this, we can consider the atomic clock experiment where two synced atomic clocks were started, one being on the ground and the other flying around on a plane. When brought back together, they were out of sync simply because they traversed different worldlines. Going back to the garden example, we established how there were multiple paths between the hive and cafe, each of varying length. The idea is similar.

Spacetime Interval $\triangle s$

"Well then what's the measurement independent of the worldline?", you squawk. Valid question.

There is only one worldline (a straight line) where the clock used to measure the time interval is an inertial clock traveling at a constant velocity. This measurement is powerful since it doesn't depend on frames or worldlines. It depends only on the separation between events. The spacetime interval is a special case of $\triangle t$ and $\triangle T$ since we are working with a specific worldline and within the RF of the inertial clock.

> **Definition**
>
> The *Spacetime Interval* is the proper time as measured by an inertial clock present at both events.

An odd quirk is that the spacetime interval is the longest proper time between two events. This is a result of spacetime being non-Euclidean. We can relate all three of our times as follows:

$$\triangle t \geq \triangle s \geq \triangle T$$

Metric Equation The metric equation links the frame independent $\triangle s$ to frame-dependent coordinate separations in any IRF:

$$\triangle s^2 = -\triangle t^2 + \triangle x^2 + \triangle y^2 + \triangle z^2$$

Also, noting that

$$\triangle s^2 = \sqrt{1 - v^2} \triangle t$$

allows us to see a few mildly interesting things. For instance, $\triangle s \leq \triangle t$. Also, if v approaches $c = 1$, there is a huge discrepancy between the time intervals, but if v is very small, they are approximately equal. To preserve causality (one event must happen before the event it's causing), c must be the speed limit. For small velocities, we can apply a binomial approximation such that $\sqrt{1 - v^2} \approx 1 - \frac{1}{2}v^2$.

Another representation is:

$$\triangle T = \int_{\text{worldline}} \sqrt{1 - v^2} dt$$

We can also categorize the spacetime interval based on its sign:

$\triangle s^2 > 0$; timelike
$\triangle s^2 = 0$; lightlike
$\triangle s^2 < 0$; spacelike

Lorentz Transform

If you glossed over everything prior to this point...well, my feelings are hurt quite frankly but whatever. This is the most important section so pay attention you parakeet.

The Galilean Transform, as we established, was incredibly flawed. The Lorentz Transform is basically relativistic gold since it'll allow you to relate the four vectors describing the same event(s) between different reference frames.

Lorentz Transform

$$\gamma = \frac{1}{\sqrt{1 - (\frac{v}{c})^2}} \geq 1$$

$$t' = \gamma t - \gamma v x = (t - vx) \frac{1}{\sqrt{1 - v^2}}$$

$$x' = \gamma x - \gamma v t = (x - vt) \frac{1}{\sqrt{1 - v^2}}$$

$$t = \gamma t' - \gamma v x' = (t' - vx') \frac{1}{\sqrt{1 - v^2}}$$

$$x = \gamma x' - \gamma v t' = (x' - vt') \frac{1}{\sqrt{1 - v^2}}$$

Consequences of the Lorentz Transform include time dilation....

Time Dilation

$$\triangle t = \gamma \triangle t'$$

...and length contraction — a displacement measured perpendicular to the relative motion of two IRFs will be identical in both IRFs, but parallel measurements are subject to length contraction where the length (distance between two events) in one frame is shorter than in the other

Length Contraction

$$L' = \frac{L}{\gamma}$$

and L is the largest possible length that can be observed.

Metric Tensor

You may have noticed that there happens to be a negative sign before the time component of the metric equation. This isn't accidental; it's a consequence of the 4D manifold that characterizes spacetime. We can rewrite the metric equation in terms of the metric tensor.

Spacetime coordinates using index notation:

$$x^\mu : [x^0 = ct, x^1 = x, x^2 = y, x^3 = z]$$

Space coordinates alone:

$$x^i : [x^1 = x, x^2 = y, x^3 = z]$$

Metric:

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

So we can see how the signature for Minkowski space is $(- + ++)$ as opposed to the Euclidean signature $(+ + ++)$.

Metric equation:

$$\triangle s^2 = \eta_{\mu\nu} \triangle x^\mu \triangle x^\nu$$

$$\triangle s = \int \sqrt{\eta_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda$$

$$\triangle T = \int \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda$$

There are a couple of things you need to know about indices that'll help them seem less confusing. Indices can be 'free' or 'dummy', and there can be both upper and lower indices. They won't bite you. It's just notation. Something you'll see a lot is the summation convention. If there is an index that's upper on one term and lower on another, it is summed over. These are the dummy indices. Free indices are preserved and appear in the final solution. We'll work with this more later.

The Lorentz Transform can be redefined as the matrices (λ) that satisfy $\eta = \lambda^T \eta \lambda$, or with all the indices, $\eta_{\rho\sigma} = \lambda^{\mu'}_{\rho} \eta_{\mu'\nu'} \lambda^{\nu'}_{\sigma}$. So we want $\lambda^{\mu'}_{\nu}$ such that $\eta_{\rho\sigma} = \eta_{\mu\nu}$. The set of Lorentz Transforms is a *non-abelian* group under matrix multiplication called the Lorentz Group.