Special Relativity Primer

W. B. Dowler
William.Dowler@cern.ch

June 12, 2004

Contents

1 Frames of reference **2**

2 Newton's Laws **3**

3 Newton's Flaws **4**

4 Einstein's Postulates **5**

5 Lorentz Transformations **6**

6 Velocity Transformation **7**

7 Length Contraction **9**
7.1 The Length Contraction Paradox 9

8 Time Dilation **10**
8.1 The Twin Paradox 10

9 Simultaneity **11**

1 Frames of reference

Before we can sensibly talk about relativity, we need to lay some groundwork. Any and all work done in science is built upon the idea that we can observe our environment and find rules to explain and predict what we see. These observations come as measurements, and these measurements are always measured with respect to some reference position.

If you are measuring something with a meter stick, you are measuring the length or position of something with respect to the end of the meter stick marked as 0 mm. If you are timing something, your time duration is measured with respect to the time you activated the stopwatch, which we call 0 s. You don't generally measure position with respect to the Sun, since that would require very complicated calculations and measurement devices, since we are on a planet that rotates about its axis as it orbits the sun. Similarly, you don't generally use the Big Bang as the moment where the time was 0 s. In other words, you are using the room you are in as your *frame of reference*, and you measure all things with respect to that frame.

There are two kinds of frames of reference, known as *inertial reference frames* and *non-inertial reference frames*. An inertial reference frame is one that does not exhibit *pseudo-forces*. A pseudo-force is what we call it when we see behaviour that appears to be a force when, in fact, there is no interaction.[1] For example, imagine that you are driving a hearse at well above the speed limit when a pedestrian crosses the street in front of you. You slam on the brakes, and the corpse and coffin in the back seat accelerate forward, slamming into the back of your seat. If you choose the hearse as your frame of reference, then the force which accelerated the body forward was a *pseudo-force*; nothing propelled it toward you, and nothing felt a reaction force. This is a non-inertial reference frame. If, however, you select the street as your reference frame, then there are no pseudo-forces, and the corpse was simply compelled by its inertia to continue forward as friction stopped the vehicle it was in. All reference frames follow this pattern: a reference frame is an inertial reference frame if it does not experience an acceleration, which a non-inertial reference frame is being accelerated.[2]

Special relativity is "special" because it doesn't deal with non-inertial reference frames. General relativity, which we won't discuss here, is more generally applicable because it can deal with either inertial or non-inertial reference frames.

[1] Every force requires an interaction between two bodies, thanks to Newton's Third Law, which will be discussed in the next section.

[2] Technically, you can only define acceleration between two reference frames, and then two frames with matching accelerations can then appear inertial when they are not. To be truly accurate, you must check the frames for pseudo-forces to determine if they are inertial or not, but this "rule of thumb" works in most cases.

2 Newton's Laws

Newton set forth four axioms[3] when describing his laws of motion.

Three of these axioms are well known, and well supported by observations. Cajori's translation of Newton's works lists them as:

- Every body continues in its state of rest, or of uniform motion in a right line, unless compelled to change that state by forces impressed upon it.

  In other words, objects only accelerate when they experience an unbalanced force. The converse is also true; any object that experiences an unbalanced force will accelerate.

- The change of motion is proportional to the motive force impressed; and is made in the direction of the right line in which that force is impressed.

  In other words, $\vec{F} = m\vec{a}$, where m refers to the *inertia* of a body, and not it's mass. This distinction is important. Newton's work showed no difference between mass and inertia, but Einstein's work requires it.

- To every action there is always opposed an equal reaction: or, the mutual action of two bodies upon each other are always equal, and directed to contrary parts.

  In other words, every action has an equal and opposite reaction. These first three axioms (which Newton called Laws) are sound when framed in terms of inertia instead of mass, and are still taught as Newton's three laws of kinematics. The Newtonian mechanics that develop from them form an extremely useful fiction. As long as the object isn't too big, too small, or too fast, we won't notice the differences without extremely sensitive equipment. Engineers, for instance, do not need to know about relativity yet. (Computer circuits are getting there.)

The fourth axiom is the one that needs to be changed. This one was implicit in his works, and often quoted in Newton's philosophical publications, but he didn't state it outright anywhere in his *Principia Mathematica* that I can find. (This axiom wasn't disputed during Newton's lifetime, so he didn't make a big deal out of it.) Hence, the wording here is my own.

- There is an absolute frame of reference, known as God's frame, which every other frame is moving with respect to. God's frame is the one which will never accelerate, while frames attached to other bodies do, thus allowing for a preferred, correct frame.[4]

[3] An axiom is something that is considered to be a self-evident truth. In other words, it's something that seems obvious to the person who proposed it, but is difficult or impossible to prove. In the experimental sciences, they are rarely set forth any more, because so many have been disproven. In math, well, everything's an abstract construct, so there's no problem including or omitting axioms as long as they are consistent with the other axioms being used.

[4] Apparently, Newton's God was a lazy slob who never moved, ever.

This idea was discarded around the turn of the twentieth century for reasons discussed in the next section.

One of the important facts that supported Newton's theory was that it led to the Galilean transformations. Galileo was the first to think about frames of reference, and he proposed a way for a person in one frame to predict the measurements made by a person in another frame. Normal variables are for measurements made in a single frame we'll call S. Variables with a $'$ are those made by an observer in the frame S', which is moving with respect to S with a speed v in the positive x direction.

$$x' = x - vt \tag{1}$$
$$y' = y \tag{2}$$
$$z' = z \tag{3}$$
$$t' = t \tag{4}$$

The idea that the flow of time would depend on an observer's reference frame was so counterintuitive and outright bizarre that the fourth equation above wasn't even written down in Newton's time. I've included it for comparison with later results.

3 Newton's Flaws

Newton's theories worked extremely well, for about 200 years. Then, a couple of problems crept up.

The first major problem was that Newton's laws above (and his law of gravity) failed to accurately predict the precession of Mercury's orbit. Mercury's orbit is very eccentric, meaning it's very elliptical. The long axis keeps rotating around the Sun each time Mercury makes an orbit. Essentially the entire orbit is rotating around the Sun. Newton's laws were unable to explain the amount of precession we saw in Mercury's orbit around the Sun, despite being able to accurately predict the precession of orbits of moons around other planets in the solar system. This problem cannot be solved without Einstein's general theory of relativity.

The second major problem was that Newton's equations seemed to be inconsistent with Maxwell's equations describing electricity and magnetism. The problem was in the use of the Galilean transformations. If one applied them to Maxwell's equations, one found that two observers would predict entirely different phenomena.

Imagine two infinitely long rods, each with the same uniformly distributed static charge. If the rods are at rest in frame S, then an observer in frame S will say that they repel each other, as like charges repel. However, an observer in a frame S' that is moving with respect to S will say that the charges are moving in his frame. In other words, we have parallel currents, so we'd predict a magnetic attraction as well as the electrostatic repulsion. If the observer moved fast enough,[5] the magnetic attraction would be strong enough to overcome the electrostatic repulsion. Two observers couldn't agree on whether the rods would be attracted to each other, or repelled by each other! This was a problem.

At the time, Newton's laws were a couple of centuries old, and some of Maxwell's equations were barely a couple of years old, while others were no older than a few decades. The vast majority of physicists understandably assumed that the problem was in electrostatic theory, and that the equations they'd found only applied exactly as written in God's frame. All other frames would have some sort of correction factor, and when this was found, we'd have finally found God's frame. Theologists were either excited or terrified by the notion of being able to pinpoint Heaven by looking for the body moving in that frame.

Some people, including young H. A. Lorentz, decided to figure out how the Galilean transformations would need to be adjusted to fix the problems. These transformations predicted some astounding phenomena that caused most people to reject them immediately. They'd be resurrected later when Einstein derived them using different logic.

4 Einstein's Postulates

Albert Einstein had very little formal education. He just couldn't focus on the material his teachers told him to focus on. He followed most developments in math and physics on his own, but he wasn't an "official" member of the community.

Einstein was working as a patent clerk when an odd revelation hit him. Maxwell's equations predicted the speed of light by the equation

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}} \tag{5}$$

where $\mu_0 = 4\pi \times 10^{-7}\,\mathrm{N/A^2}$ is the same constant we saw when dealing with magnetic theory, and ϵ_0 does a similar job for electrostatics.[6] Einstein realized that the derivation of the speed of light was in no way dependent upon a person's frame of reference. Therefore, he concluded, it must be the same for all observers. Since they already knew that observers in different reference disagree on the distance an object travels (with respect to themselves), Einstein realized that time must depend upon an observer's reference frame. After all, the speed of light was the ratio of the distance light travelled to the time elapsed as it travelled, and if the numerator in a constant ratio changes, then the denominator must change as well.

Einstein then discarded Newton's fourth axiom, and added two of his own:

- The speed of light in vacuum is invariant for all observers in all inertial reference frames.[7]

[5] Newton's mechanics have no speed limit, so someone could move "fast enough" to do just about anything

[6] Coulomb's constant k, when derived from theory, turns out to be $k = \frac{1}{4\pi\epsilon_0}$.

[7] The speed of light is exactly $299,792,458\,\mathrm{m/s}$. At first, it seemed incredible to me that

This was the one that really threw people for a loop. His second axiom was being bantered about by a few of the atheist philosophers, but they were in a clear minority.

- There is no absolute frame of reference. Instead, every individual in every inertial frame of reference has his or her experiences determined by the same set of equations.

  In other words, Montgomery Scott was right. Ye canna change the laws of physics! Everyone, everywhere, in every inertial frame is governed by the same laws.[8]

These new axioms led to the same set of transformations that Lorentz found from a different approach. Lorentz' definition is the kind that mathematicians love; it's absolutely formal, working strictly by the math and as few axioms as possible. In other words, Lorentz took Maxwell's equations, and used them to find the only set of transformations that can be used with them consistently. Einstein's derivation is much easier to follow, but it's a bit sloppier, adding in more constraints than mathematicians tend to allow. These were perfectly logical constraints, though. He included Newton's three axioms, his own first axiom listed above, the requirement that they should be very close to the Galilean versions when the relative speeds of the frames is low, and an additional constraint that everyone agree on which direction an object is travelling in its own reference frame.

These transformations won't be derived here, since that would be about eight uninteresting and uninformative pages. They can be found in Einstein's text, which has been reprinted by several publishing companies. It's called *Relativity: The Special and the General Theory*. Any University level text on the subject will have a derivation as well. This document will simply list them in the next section.

5 Lorentz Transformations

The new versions of the transformations are

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{6}$$

$$y' = y \tag{7}$$

this number is exact when we use our man-made arbitrary units of meters and seconds. Then I learned that it's only been exactly this value since 1984, when the SI international standards council defined a meter as the distance light travels in a vacuum in $\frac{1}{299{,}792{,}458}$ s. We're still a little fuzzy on how long a meter is or how much time elapses in a single second, but we know the speed of light exactly. This was the same meeting that redefined the Ampere, so that it's no longer based on the amount of current that produces a given force between two wires of known length, but so we have the exact value of $\mu_0 = 4\pi \times 10^{-7}\ \mathrm{N/A^2}$.

[8]When Einstein introduced his general theory of relativity, he was able to relax this, and find a set of laws that governs the non-inertial reference frames as well.

$$z' = z \tag{8}$$

$$t' = \frac{t - \frac{v}{c^2}x}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{9}$$

The new aspects are the frame dependence of time, and the inclusion of the factor

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{10}$$

If $v << c$, then $\gamma \approx 1$ and $\frac{v}{c^2} \approx 0\,\mathrm{m/s}$, leaving us with the Galilean versions. The bizarre implications of these equations will be dealt with below.

6 Velocity Transformation

We now know how we need to change the way we measure times and distances in relativity. How about the way we change velocities?

Let's assume that observers in frames S and S' are looking at an object in uniform motion. We're already using v for the relative speed of two reference frames, so we'll use u for the speed of the object as measured in the S frame, and u' for the speed as measured in the S' frame. Assume that the object is moving parallel to both the x and x' directions. Newton defined velocity as[9]

$$u = \frac{\Delta x}{\Delta t}$$

This definition is still good. We just need to make a slight modification to find u', the speed of the object as measured by an observer in S'. Well, if Einstein's second axiom holds, we just need to use this formula with the position and time as measured in S'. So,

$$u' = \frac{\Delta x'}{\Delta t'}$$

We now turn to the Lorentz transformations above. Remember that these are inertial frames, so v will never change. That means that γ is constant, so we might as well write our transformations as

$$x' = \gamma\,(x - vt)$$

$$t' = \gamma\left(t - \frac{v}{c^2}x\right)$$

We can then say that, since γ and c are constants,

$$\Delta x' = \gamma\,(\Delta x - v\Delta t)$$

$$\Delta t' = \gamma\left(\Delta t - \frac{v}{c^2}\Delta x\right)$$

[9]Technically, he defined it as $\vec{u} = \frac{d\vec{x}}{dt}$. If you've taken calculus, you know the difference. You can just replace Δ with d in what follows and use the same logic. Transform the components individually. Vector calculus on this stuff should wait until you've done some vector calculus and linear algebra.

As a result,

$$
\begin{aligned}
u' &= \frac{\gamma\left(\Delta x - v\Delta t\right)}{\gamma\left(\Delta t - \frac{v}{c^2}\Delta x\right)} \\
&= \frac{\Delta x - v\Delta t}{\Delta t - \frac{v}{c^2}\Delta x} \\
&= \frac{\frac{\Delta x}{\Delta t} - v}{1 - \frac{v\Delta x}{c^2\Delta t}}
\end{aligned}
$$

where we have cancelled out the factors of γ to reach the second step, and divided both numerator and denominator by Δt to reach the third step. When we note the definition of u, we reduce this to

$$
u' = \frac{u - v}{1 - \frac{uv}{c^2}} \tag{11}
$$

To check our result, we make sure that the speed of light really is invariant, or that $u' = c$ when $u = c$. We have

$$
\begin{aligned}
c' &= \frac{c - v}{1 - \frac{cv}{c^2}} \\
c' &= \frac{c - v}{1 - \frac{v}{c}} \\
c' &= c\frac{1 - \frac{v}{c}}{1 - \frac{v}{c}} \\
c' &= c
\end{aligned}
$$

So, everything looks good. This is our new velocity transformation. You can prove that any time $u < c$, you also have that $u' < c$, as follows:

$$
\begin{aligned}
u &< c \\
u\left(1 + \frac{v}{c}\right) &< c\left(1 + \frac{v}{c}\right) \\
u + \frac{uv}{c} &< c + v \\
u - v &< c - \frac{uv}{c} \\
u - v &< \left(1 - \frac{uv}{c^2}\right) c \\
\frac{u - v}{1 - \frac{uv}{c^2}} &< c \\
u' &< c
\end{aligned}
$$

This proves that, not only is the speed of light invariant, it's also the ultimate speed limit of the universe. Anything going slightly slower than an object will find that its speed is less than c. This then forces the speed of that object to be less than c in every reference frame. Nothing can ever exceed the speed of light.

7 Length Contraction

Imagine you have a rod of length L. This idea, combined with the Lorentz transformations, leads to the notion of length contraction, and to a limerick that shouldn't be shared with High School students.

Place this rod with one end at the origin in an inertial reference frame S' while you stay in frame S. (In other words, put it in the back of a truck, or on a conveniently passing train.) In frame S', its length is L. The length is determined by measuring the distance between the two ends of the rod, though. Remember, our two frames have a common origin, meaning that when $x = y = z = ct = 0\,\text{m}$, we also have $x' = y' = z' = ct' = 0\,\text{m}$.[10] Now, we measure the length of the rod in frame S. If we measure it at time $t = 0\,\text{s}$, then we know the length will just be the coordinate of the end of the rod that is *not* at $x = 0\,\text{m}$. For this, we turn to the first of the Lorentz transformations:

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}$$

and we not that $t = 0\,\text{s}$ and $x' = L$ (since the rod's length is unchanged in frame S', where it is at rest). Now, we solve for x and find that

$$x = L\sqrt{1 - \frac{v^2}{c^2}}$$

or, the length L' of a moving rod is given by

$$L' = L\sqrt{1 - \frac{v^2}{c^2}} \tag{12}$$

Since $0 \leq \sqrt{1 - \frac{v^2}{c^2}} \leq 1$, we know that $0\,\text{m} \leq L' \leq L$, so a moving rod will always be *smaller* than an stationary rod. This is length contraction, and it is the basis of the solution to the "two infinite rods" problem mentioned in the section about Newton's flaws. The charges on the two rods would be created by excess charged particles, either electrons or protons. The number of this particles on the rods does not change between reference frames, nor does the charge on the individual particles. Then *density* of the charges increases, however; you are packing the same total amount of charge into a shorter length, due to Lorentz contraction. This increases the electrostatic repulsion enough to overcome the magnetic attraction, and the rods repel in all frames.

While length contraction fixed one problem, it seemed to cause another.

7.1 The Length Contraction Paradox

Imagine your friend is carrying a ladder that is 20 m long. Imagine that you have a barn that is 15 m deep. Now, imagine that your friend can run very, very fast.

[10] The only reason to multiply t by c here is to make sure all values have the units of distance. If $ct = 0\,\text{m}$, then $t = 0\,\text{s}$, since $c \neq 0\,\text{m/s}$.

Relativity predicts that, if your friend runs fast enough, the ladder will shorten enough to fit in the barn. Let's assume he runs at $\frac{\sqrt{3}}{2}c$. Then, we find that $L' = \frac{L}{2}$. So, you can kick back and watch your friend run fast enough that the ladder becomes 10 m long, fitting comfortably into the barn. You can even close the door with room to spare, provided to have enough time to be somewhere else when the ladder stops and becomes 5 m longer than the barn itself.

However, there seems to be a problem when we apply Einstein's second axiom. Things should work just as nicely from your friend's perspective. From his perspective, he's carrying a 20 m ladder on his shoulder while a 7.5 m deep barn is bearing down at him at $\frac{\sqrt{3}}{2}c$. This seems to be a problem for your friend.

The problem is solved when we realize that, as found in the velocity transformation section above, nothing can travel faster than light, including information. In other words, there is a time delay between when the front of the ladder hits the back of the barn and when the back of the ladder "learns" it has to stop. When the front of the ladder hits the back of the barn, the back of the ladder will have two things rushing to meet it. The front of the barn is 12.5 m away and approaching at $\frac{\sqrt{3}}{2}c$, while the information that it needs to stop is 20 m away and approaching at c. The entire ladder will get into the barn if and only if the front of the barn gets to the back of the ladder before information about the impact does. In other words, it needs to take less time, so that

$$\frac{12.5\,\mathrm{m}}{\frac{\sqrt{3}}{2}c} < \frac{20\,\mathrm{m}}{c}$$

You can check that this condition is met.

8 Time Dilation

We can use the same logic on time intervals that we used on space intervals above to determine what happens to time as measured by a moving observer. If we measure a time interval with no different in spatial coordinates, we find that

$$t' = t\sqrt{1 - \frac{v^2}{c^2}} \tag{13}$$

where t is the time interval measured in frame S and t' is the time interval measured in frame S'. In other words, a moving reference frame measures a smaller time interval than one in your own inertial reference frame. We say that "moving clocks run slow."

This also seems to lead to a paradox.

8.1 The Twin Paradox

Imagine two identical twins. One stays on Earth (let's call her Mary Kate), and the other (let's call her Ashely) flies to Alpha Centauri at a significant fraction of light speed, only to (unfortunately) instantaneously turn around on arrival and come back at the same speed. From Mary Kate's perspective, Ashley has had less time elapse, so Ashely should be younger than she is. It would seem that Einstein's second axiom above would also predict that Ashley would say Mary Kate is younger. Take a good look at that axiom: you must be in an *inertial* frame of reference. When Ashley turned around and came back, she not only upset those who thought we'd finally gotten rid of her, but she accelerated. She was no longer in an inertial reference frame, and the rules provided no longer apply. General relativity provides the appropriate rules; Mary Kate is the younger twin, and the age she predicts with special relativity is the correct one. Ashley's predictions are only correct until she turns around and comes back.

9 Simultaneity

Many of our perceptions depend upon our reference frame. Our determination of which events are simultaneous is one of them.

Simultaneous events are those that we perceive at the same time. However, time is dependent upon our frame of reference. Therefore, we formally define simultaneity are those events whose existence we are informed of at the same time. If you are standing exactly in the middle of two traffic lights and you say that they change at the same time, then the light from them reached you at the same time as you measure it. For you, they are simultaneous. A person directly below one light, who is also standing motionless, will say that the closer light changed first, and they were not simultaneous. This is because the light (information) from the nearer traffic light had less distance to travel, and reached this person sooner.

A person who is moving with respect to the traffic lights may or may not agree that they are simultaneous. The light from each traffic light need only reach him at the same time as he measures it for the events to be considered simultaneous. Therefore, there is no absolute definition of which events are simultaneous, and which ones are not.