Special Relativity

Andi Gu

A Thought Experiment

We begin with the axiom that the speed of light c is a universal constant. This forms the underpinning for everything we derive.

Imagine two frames of reference A and B, with B moving at a speed v in the $+\hat{x}$ direction relative to A. We look to investigate how length and time behave for B, in comparison to A. We put a 'light' clock in B's system that works as follows: it shoots a beam of light at a mirror at a height h, and when it returns, this counts as two units of time. In A's frame of reference, this is how B's clock looks:

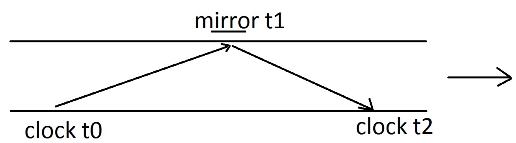

Figure 1: Adapted from StackExchange

A sets up an identical clock for himself that is stationary in his frame of reference. We calculate how long each clock takes to 'tick'. For A, this is simple – it is $\Delta t = \frac{h}{c}$. For B, it is a bit more complicated since the mirror travels a bit in between ticks and thus the light must travel longer. If the tick takes $\Delta t'$:

$$\begin{aligned}\Delta t' &= \frac{\sqrt{(v\Delta t')^2 + h^2}}{c} \\ c^2\Delta t'^2 &= v^2\Delta t'^2 + h^2 \\ \Delta t' &= \frac{h}{\sqrt{c^2 - v^2}}\end{aligned}$$

The ratio between the two times is:

$$\begin{aligned}\frac{\Delta t'}{\Delta t} &= \frac{h}{\sqrt{c^2 - v^2}}\frac{c}{h} \\ &= \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}\end{aligned}$$

The factor $\frac{v}{c}$ appears quite often in special relativity and so we call it β.

$$= \frac{1}{\sqrt{1-\beta^2}}$$

This ratio gets the special name γ – it notes the fact that one second in a moving frame is longer than a second in a stationary frame by a factor of γ.

We do something similar for length. We set up the Michael-Morsley experiment.

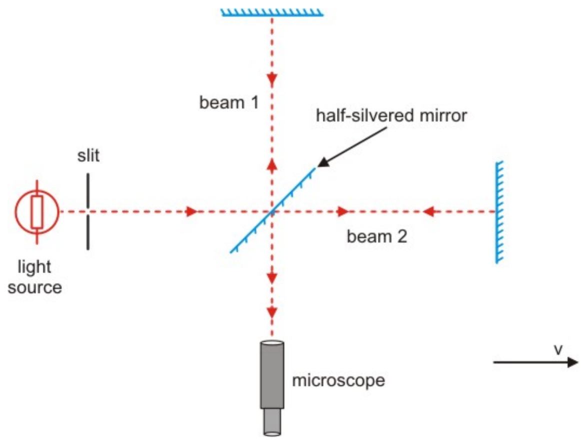

Figure 2: Adapted from here

Imagine we set this up in the moving B's frame of reference. The time it takes for the beams to return must be the *same*, since B otherwise B would be able to tell he was moving. In B's frame of reference, the length of each arm is L. In A's frame of reference, the transverse arm is L as well (since it is perpendicular to the motion[1]), but we allow for the other arm to be potentially a different length L' and find what it must be in order for the light to return to the mirror at the same time in A's frame of reference. We call the time for the transverse light to return Δt_1.

$$\begin{aligned}
\sqrt{L^2 + \left(\frac{v\Delta t_1}{2}\right)^2} &= \frac{c\Delta t_1}{2} \\
L^2 + \frac{(v\Delta t_1)^2}{4} &= \frac{(c\Delta t_1)^2}{4} \\
\Delta t_1 &= \sqrt{\frac{4L^2}{c^2 - v^2}} \\
&= \frac{2L}{\sqrt{c^2 - v^2}}
\end{aligned}$$

We call the time for the parallel light to hit the mirror Δt_2 and the time to return Δt_3. Our

[1]This reason is only an intuition and perhaps needs to be further explicated

restriction is thus $\Delta t_1 = \Delta t_2 + \Delta t_3$.

$$\begin{aligned} L' + v\Delta t_2 &= c\Delta t_2 \\ \Delta t_2 &= \frac{L'}{c-v} \end{aligned}$$

Similarly:

$$\begin{aligned} \Delta t_3 &= \frac{L'}{c+v} \\ \Delta t_2 + \Delta t_3 &= \frac{2L'c}{c^2 - v^2} \end{aligned}$$

We apply our restriction and see how L' is related to L:

$$\begin{aligned} \Delta t_1 &= \Delta t_2 + \Delta t_3 \\ \frac{2L}{\sqrt{c^2 - v^2}} &= \frac{2L'c}{c^2 - v^2} \\ L' &= L\frac{\sqrt{c^2 - v^2}}{c} \\ &= \frac{L}{\gamma} \end{aligned}$$

So length contracts in moving frames. More concretely, a meter stick that B holds looks to be shorter by a factor of $\frac{1}{\gamma}$ in a stationary frame of reference.

Deriving the Lorentz Transform

First, a note on events and observation. When we talk of an event occurring at (x, t), and (x', t'), we assume the (x, t) we speak of is the result of *instant* observation. What this means is that if a star explodes 10 light years away at $t = 0$, we will observe it at a time $10c$ – however, we say the event occurred at $t = 0$, because we adjust for observation time. Disparities between the coordinates (x, t) and (x', t') will never be due to observation delays such as this – they are due entirely to time and length contractions.

First, to derive $x' = \gamma(x - vt)$. This is quite intuitive. An event happens at (x, t). At this time, B is at $x_B = vt$, and the event is thus a distance (in A's frame of reference) $x - vt$ away from B. However, for B's shortened meter stick, this will be farther by a factor of γ (it takes more 'meter sticks' to cover the gap since B's meter stick is shorter than that of A).

The transform $t' = \gamma(t - \frac{vx}{c^2})$ is more counter-intuitive at first, but can be explained nonetheless.

Clocks

There is a unique notion of time in special relativity that is not immediately obvious. For each frame of reference, each point in space has a clock running. Clocks always agree (i.e. a clock at $x = 2$ and $x = 1000$ read the same thing forever) *within* any frame of reference. We achieve this synchronization with the following strategy. At $t = 0$, send a light beam out from $x = 0$. When a point at x sees the light beam, it sets its time to $\frac{x}{c}$ and starts. This is key to understanding the time Lorentz transform.

We synchronize the clocks for A and B and see how their clocks differ. The light is beamed out at $x = 0$, $t = 0$. Let E be the event that a point in space x_E sees the light – it is important to note that this event E defines time for A and B, since it sets the clock in motion. For A, this occurs at $x_E = x_E$ and $t_E = \frac{x_E}{c}$, as expected. For B, this occurs at $x'_E = \gamma(x_E - vt_E) = \gamma\big(x_E - \frac{vx_E}{c}\big)$. The clock in his reference sets itself to $t_E = \frac{x'_E}{c} = \gamma\big(\frac{x_E}{c} - \frac{vx_E}{c^2}\big) = \gamma\big(t_E - \frac{vx_E}{c^2}\big)$. So we have the remarkable result that where a clock in A gets initialized as t_E, a clock at that point for B gets initialized as $\gamma\big(t_E - \frac{vx_E}{c^2}\big)$.

We observe now that the difference between times at any fixed pair of points remains constant over time, since they run at the same rate (even though, for B's clocks, A may not see them to be synchronized). In fact, the difference between two times is a linear function of the distance between them. We proceed to quantify this difference. Say we take a snapshot of B's clocks at $x' = 0$ and $x = x_1$ when the clock at $x = x_1$ has just been started. The clock at $x = x_1$ reads $\gamma(\frac{x_1}{c} - \frac{vx_1}{c^2})$ as found previously. The clock at $x' = 0$ reads $\frac{x_1}{\gamma c}$ (this was demonstrated in the first part of the note – time in the origin[2] of B runs slower than that of A by a factor of $\frac{1}{\gamma}$). Thus, the ratio between the time difference is:

$$\begin{aligned}
\frac{\gamma(\frac{x_1}{c} - \frac{vx_1}{c^2}) - \frac{x_1}{\gamma c}}{x_1 - \frac{vx_1}{c}} &= \frac{\gamma}{c} - \frac{x_1}{\gamma c}\frac{c}{x_1(c-v)} \\
&= \frac{\gamma}{c} - \frac{1}{\gamma(c-v)} \\
&= \frac{\gamma}{c}\left(1 - \frac{1-\beta^2}{1-\beta}\right) \\
&= \frac{\gamma}{c}(1 - 1 + \beta) \\
&= -\frac{\gamma v}{c^2}
\end{aligned}$$

Thus, the time at a point x is proportional to its distance from $x' = 0$ (or $x = vt$) by a factor of $-\frac{\gamma v}{c^2}$. The time at $x' = 0$ is $\frac{t}{\gamma}$, so the time at other points is:

$$\begin{aligned}
t'(x,t) &= \frac{t}{\gamma} - (x - vt)\frac{\gamma v}{c^2} \\
&= \gamma\left((1-\beta^2)t - \frac{xv}{c^2} + \frac{v^2}{c^2}t\right) \\
&= \gamma\left(t - \frac{xv}{c^2}\right)
\end{aligned}$$

We have thus proved the Lorentz transform for time.

Intuition

Why is the clock at x not *smaller* by a factor of $\frac{1}{\gamma}$? After all, we showed that time runs slower for B by a factor of $\frac{1}{\gamma}$ – this is true, clocks run slower. However, B's clocks at two different points

[2]This rule holds only for the origin of B, however. This is important to note. B's clocks do not run more slowly at fixed x, this is because B is also 'dragging' along clocks, so the clock we observe at x at one time is a 'different clock' than what we observe at a later time.

x_1 and x_2 do *not* agree in general. As mentioned previously, B is essentially dragging clocks along (B's clocks do not stay fixed).

Velocity

Take an object that is moving at a velocity u'_x in the frame of reference of an observer moving at a velocity v. Then, $\Delta x' = u'_x \Delta t'$. How does this object behave for a stationary observer? First, $\Delta x = \gamma(\Delta x' + v\Delta t')$, and $\Delta t = \gamma(\Delta t' + \frac{v\Delta x'}{c^2})$. So the velocity for the stationary observer is:

$$\begin{aligned} u_x &= \frac{\Delta x}{\Delta t} \\ &= \frac{\Delta x' + v\Delta t'}{\Delta t' + \frac{v\Delta x'}{c^2}} \\ &= \frac{u'_x \Delta t' + v\Delta t'}{\Delta t' + \frac{u'_x v \Delta t'}{c^2}} \\ &= \frac{u'_x + v}{1 + \frac{u'_x v}{c^2}} \end{aligned}$$

On the other hand, for velocity in transverse directions such as $\hat{y}$ or $\hat{z}$, the distance travelled is unaffected but time is. $u_y = \frac{1}{\gamma} u'_y$, and likewise for u'_z – this is because B's clocks run slower. It is important to note that this only applies if $u'_x = 0$, otherwise, $\Delta t \neq \gamma \Delta t'$.

Revisiting Addition of Velocities

Addition of velocities is not so simple in 3D space. We first solve the problem of finding the rest frame velocity of an object moving at $\vec{\mathbf{u}}'$ in a frame moving at $\vec{\mathbf{v}}$. The component of $\vec{\mathbf{u}}$ parallel to $\vec{\mathbf{v}}$ (which we denote with $\|$) is the typical formula, $\frac{\vec{\mathbf{u}}'_\| + \vec{\mathbf{v}}}{1 + \frac{\vec{\mathbf{u}}' \cdot \vec{\mathbf{v}}}{c^2}}$. The component of $\vec{\mathbf{u}}$ perpendicular to $\vec{\mathbf{v}}$ (which we denote with $\perp$) is $\frac{1}{\gamma_v} \frac{\vec{\mathbf{u}}'_\perp}{1 + \frac{\vec{\mathbf{u}} \cdot \vec{\mathbf{v}}'}{c^2}}$. Thus:

$$\begin{aligned} \vec{\mathbf{u}} &= \frac{\vec{\mathbf{u}}'_\| + \vec{\mathbf{v}}}{1 + \frac{\vec{\mathbf{u}}' \cdot \vec{\mathbf{v}}}{c^2}} + \frac{1}{\gamma_v} \frac{\vec{\mathbf{u}}'_\perp}{1 + \frac{\vec{\mathbf{u}} \cdot \vec{\mathbf{v}}'}{c^2}} \\ &= \frac{1}{1 + \frac{\vec{\mathbf{u}}' \cdot \vec{\mathbf{v}}}{c^2}} \left(\vec{\mathbf{u}}'_\| + \vec{\mathbf{v}} + \frac{\vec{\mathbf{u}}'_\perp}{\gamma_v} \right) \\ &= \frac{1}{1 + \frac{\vec{\mathbf{u}}' \cdot \vec{\mathbf{v}}}{c^2}} \left(\vec{\mathbf{u}}'_\| + \vec{\mathbf{v}} + \frac{\vec{\mathbf{u}}' - \vec{\mathbf{u}}'_\|}{\gamma_v} \right) \end{aligned}$$

Let $\alpha_v = \frac{1}{\gamma_v}$.

$$= \frac{1}{1+\frac{\vec{\mathbf{u}}'\cdot\vec{\mathbf{v}}}{c^2}}\left(\alpha_v\vec{\mathbf{u}}' + \vec{\mathbf{v}} + (1-\alpha_v)\frac{\vec{\mathbf{v}}\cdot\vec{\mathbf{u}}'}{\|\vec{\mathbf{v}}\|^2}\vec{\mathbf{v}}\right)$$
$$= \frac{1}{1+\frac{\vec{\mathbf{u}}'\cdot\vec{\mathbf{v}}}{c^2}}\left(\vec{\mathbf{v}} + \vec{\mathbf{u}}' + \frac{(1-\alpha_v)}{\|\vec{\mathbf{v}}\|^2}((\vec{\mathbf{v}}\cdot\vec{\mathbf{u}}')\vec{\mathbf{v}} - (\vec{\mathbf{v}}\cdot\vec{\mathbf{v}})\vec{\mathbf{u}}')\right)$$
$$= \frac{1}{1+\frac{\vec{\mathbf{u}}'\cdot\vec{\mathbf{v}}}{c^2}}\left(\vec{\mathbf{v}} + \vec{\mathbf{u}}' + \frac{1-\sqrt{1-\beta_v^2}}{\|\vec{\mathbf{v}}\|^2}(\vec{\mathbf{v}}\times(\vec{\mathbf{v}}\times\vec{\mathbf{u}}'))\right)$$
$$= \frac{1}{1+\frac{\vec{\mathbf{u}}'\cdot\vec{\mathbf{v}}}{c^2}}\left(\vec{\mathbf{v}} + \vec{\mathbf{u}}' + \frac{\gamma_v}{c^2(1+\gamma_v)}(\vec{\mathbf{v}}\times(\vec{\mathbf{v}}\times\vec{\mathbf{u}}'))\right)$$

We can find the magnitude of the velocity by squaring the sums of the parallel and perpendicular velocities:

$$\|\vec{\mathbf{u}}\| = \frac{1}{1+\frac{\vec{\mathbf{u}}'\cdot\vec{\mathbf{v}}}{c^2}}\sqrt{(\vec{\mathbf{u}}'_{\|}+\vec{\mathbf{v}})^2 + \alpha_v^2(\vec{\mathbf{u}}'_{\perp})^2}$$
$$= \frac{1}{1+\frac{\vec{\mathbf{u}}'\cdot\vec{\mathbf{v}}}{c^2}}\sqrt{\left\|\vec{\mathbf{u}}'_{\|}\right\|^2 + 2\vec{\mathbf{u}}'_{\|}\cdot\vec{\mathbf{v}} + \|\vec{\mathbf{v}}\|^2 + \alpha_v^2\|\vec{\mathbf{u}}'_{\perp}\|^2}$$
$$= \frac{1}{1+\frac{\vec{\mathbf{u}}'\cdot\vec{\mathbf{v}}}{c^2}}\sqrt{\|\vec{\mathbf{u}}'\|^2 + \|\vec{\mathbf{v}}\|^2 + 2\vec{\mathbf{u}}'_{\|}\cdot\vec{\mathbf{v}} - (1-\alpha_v^2)\|\vec{\mathbf{u}}'_{\perp}\|^2}$$
$$= \frac{1}{1+\frac{\vec{\mathbf{u}}'\cdot\vec{\mathbf{v}}}{c^2}}\sqrt{\|\vec{\mathbf{u}}'+\vec{\mathbf{v}}\|^2 - (1-\alpha_v^2)\|\vec{\mathbf{u}}'_{\perp}\|^2}$$

Momentum

We take conservation of momentum as a fundamental rule, and see what this means for how momentum transforms for a moving observer.

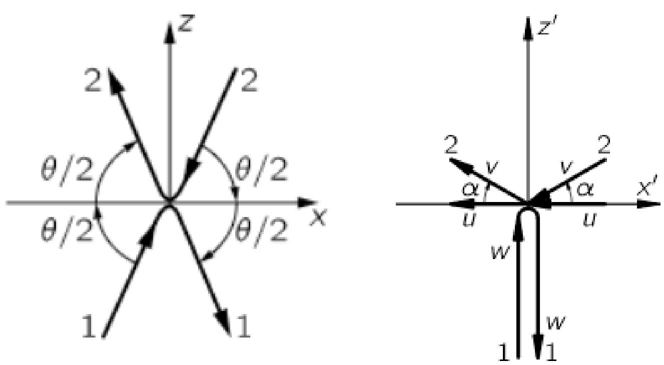

Figure 3: Adapted from Feynman's Lecture notes

Imagine the stationary observer sees a collision like the one on the left, with two identical particles moving at the same speed colliding head-on.

Now, imagine an observer moving at the horizontal speed of particle 1 (the lower one) in the $+\hat{x}$ direction, so that they observe a collision like the one on the right. Say the vertical velocity of particle 1 is w, and the components of the velocity of particle 2 are u_x and u_z. To find u_z, we need to transform to a reference frame such that $u'_x = 0$, and then apply our formula for transverse velocity. Of course, we do this simply by moving in the horizontal direction of particle 2 at a speed u_x, in which case, by the symmetry by the situation, the vertical velocity of particle 2 in this frame becomes w (we see the collision on the right, but flipped upside down). Thus, $u_z = w\sqrt{1 - \frac{u_x^2}{c^2}}$.

Now, the masses of the two particles are not necessarily the same in the case of an observer that sees a collision like the one on the right. Label the mass of the lower particle m_1 and the mass of the upper particle m_2. Then:

$$m_1 w = m_2 w\sqrt{1 - \frac{u_x^2}{c^2}}$$
$$m_2 = m_1 \gamma$$

If we take the limit as w goes to zero, we get $m_u = \gamma m_0$ (this is essentially when the particles collide nearly head-on, so that the vertical velocity of both particles is barely altered in the collision).

Now, imagine a head on collision of two particles so that they stick together. Both particles are moving at each other at a speed v, and have a tiny upwards velocity of $\mathrm{d}u$. Their mass is essentially $m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}$, and their initial combined momentum is $2m\,\mathrm{d}u$, which must also be their final momentum. Once the two particles combine, they form an object that moves upwards at a tiny speed of $\mathrm{d}u$ – but its mass must be $2m$. This is certainly non-intuitive, because now the object is at rest, but still has a mass that is 'increased' by the γ factor. Therefore, the mass of an object when two objects combine is the sum of the two masses in motion.

We use this to find its consequences for the energy of a mass m moving at a velocity v:

$$\begin{aligned} K &= \int_0^s F \cdot \mathrm{d}s \\ &= \int_0^s \frac{\mathrm{d}p}{\mathrm{d}t}\,\mathrm{d}s \\ &= \int_0^{mv} v\,\mathrm{d}p \end{aligned}$$

At this point, we can find v in terms of p:

$$\begin{aligned}
p &= \frac{m_0 v}{\sqrt{1-\frac{v^2}{c^2}}} \\
&= \frac{m_0}{\sqrt{\frac{1}{v^2}-\frac{1}{c^2}}} \\
\frac{1}{v^2}-\frac{1}{c^2} &= \frac{m_0^2}{p^2} \\
v &= \sqrt{\frac{c^2p^2}{m_0^2c^2+p^2}} \\
&= \frac{pc}{\sqrt{m_0^2c^2+p^2}}
\end{aligned}$$

So we return to our integral:

$$\begin{aligned}
K &= \int_0^p \frac{pc}{\sqrt{m_0^2c^2+p^2}}\,\mathrm{d}p \\
&= \left(c\sqrt{p^2+m_0^2c^2}\right|_0^p \\
&= c\sqrt{p^2+m_0^2c^2}-m_0c^2 \\
&= m_0c\sqrt{\frac{v^2c^2}{c^2-v^2}+c^2}-m_0c^2 \\
&= m_0c^2\sqrt{\frac{v^2+c^2-v^2}{c^2-v^2}}-m_0c^2 \\
&= \gamma m_0c^2-m_0c^2 \\
&= mc^2-m_0c^2
\end{aligned}$$

This motivates a definition for the rest energy of an object as m_0c^2.

Momentum-Energy Four Vector

We investigate how momentum and energy transform. Let $\beta_x = \frac{x^2}{c^2}$ Say that in a rest frame, we have a particle with momentum $p_x = \frac{m_0 v}{\sqrt{1-\beta_u^2}}$ and energy $E = \frac{m_0}{\sqrt{1-\beta_u^2}}$, and a frame is moving at speed u in the $\hat{\mathbf{x}}$ direction.

$$\begin{aligned}
v' &= \frac{v-u}{1-\frac{uv}{c^2}} \\
&= \frac{c^2(v-u)}{c^2-uv} \\
p'_x &= \frac{m_0 v'}{\sqrt{1-\frac{v'^2}{c^2}}} \\
&= \frac{m_0 v'}{\sqrt{1-\frac{c^2(v-u)^2}{(c^2-uv)^2}}} \\
&= \frac{m_0 v'(c^2-uv)}{\sqrt{c^4-2uvc^2+u^2v^2-c^2v^2+2uvc^2-u^2c^2}} \\
&= \frac{m_0 c^2(v-u)}{\sqrt{c^4+u^2v^2-(u^2+v^2)c^2}} \\
&= \frac{m_0 c^2(v-u)}{\sqrt{(c^2-u^2)(c^2-v^2)}} \\
&= \frac{1}{\sqrt{1-\beta_u^2}}\left(p_x-\frac{uE}{c^2}\right) \\
E' &= \frac{m_0 c^2(c^2-uv)}{\sqrt{(c^2-u^2)(c^2-v^2)}} \\
&= \frac{m_0 c^2(1-\frac{uv}{c^2})}{\sqrt{(1-\beta_u^2)(1-\beta_v)^2}} \\
&= \frac{1}{\sqrt{1-\beta_u^2}}(E-up_x)
\end{aligned}$$

As shown previously, vertical momentum is invariant under a Lorentz transformation. This is because transverse velocity is smaller by a factor $\frac{1}{\gamma}$, but mass is increased by a factor γ. Thus p_x and E transform like x and t.

Forces

Force is defined as $\vec{\mathbf{F}} = \frac{\partial \vec{\mathbf{p}}}{\partial t}$. Let the particle move with velocity v and the frame move with speed u. Say $\gamma = \frac{1}{\sqrt{1-\beta_u^2}}$.

$$\begin{aligned}
F'_x &= \frac{\partial p'_x}{\partial t'} \\
&= \frac{\frac{\partial}{\partial t}\big(\gamma\big(p_x-\frac{uE}{c^2}\big)\big)}{\frac{\partial t'}{\partial t}}
\end{aligned}$$

Assuming $\frac{\partial E}{\partial t} = \vec{\mathbf{F}} \cdot \vec{\mathbf{v}} = F_x v$

$$\begin{aligned}
&= \gamma \frac{F_x - \frac{uvF_x}{c^2}}{\gamma\left(1 - \frac{uv}{c^2}\right)} \\
&= F_x \\
F'_y &= \frac{\partial p'_y}{\partial t'} \\
&= \frac{F_y}{\gamma(1 - \frac{uv}{c^2})} \\
F'_z &= \frac{F_z}{\gamma(1 - \frac{uv}{c^2})}
\end{aligned}$$

If $v = 0$, we see $F'_y = \frac{F_y}{\gamma}$ and $F'_z = \frac{F_z}{\gamma}$.