Scratch notes on Special Relativity

Arthur J. Parzygnat

Last updated: February 22, 2018

Abstract

In this note, we derive Lorentz transformations from the postulate of the constancy of the speed of light as well as the universality of the laws of physics in different inertial frames. We discuss other consequences of this assumption such as time dilation, length contraction, and the dependence of the notion of simultaneity for different inertial observers. In particular, we also discuss simultaneity in some detail including the ordering of spacetime events and causality. Most of this note is expository with the exception of a few examples in the last section on black holes and the Schwarzschild radius.

Contents

1 Newton's laws and Galilean transformations

Newton's law says that the force on a (point-like) massive object is proportional to its acceleration

$$F \propto a. \tag{1.1}$$

The proportionality constant can be used to define the mass of the object through

$$F = ma. \tag{1.2}$$

Conversely, this law says that if the forces are known, the acceleration can be determined. An inertial frame of reference is one in which Newton's law holds. More accurately,[1] an inertial frame of reference is one in which the trajectory of a particle that is not subject to any forces is in a straight line at constant velocity [5]. Let us consider two intertial frames, $\mathcal{O}$ and $\mathcal{O}'$, and a body moving in space. In what follows, we will compare the descriptions of a moving body and how to transform between these two descriptions.

Let us describe the coordinates of this moving body in these frames using (x, y, z, t) and (x', y', z', t'), respectively. We will further assume that time is absolute, which means that t and t' are related by a constant shift. Because of this, it suffices to assume $t = t'$. This assumption states that the time passed in $\mathcal{O}$ coincides with the time passed in $\mathcal{O}'$. It suffices to suppose that $\mathcal{O}$ is at rest (the existence of such a frame is the assumption that space is absolute) and that $\mathcal{O}'$ is moving with respect to $\mathcal{O}$ at a velocity $v_x\hat{x} + v_y\hat{y} + v_z\hat{z}$. Let us denote the velocity of the moving body as $\vec{u}$ and $\vec{u}'$ in the respective frames. Similarly, let $\vec{a}$ and $\vec{a}'$ denote the accelerations. The Galilean-Newtonian postulate states that the accelerations in both frames will be the same, i.e. $\vec{a} = \vec{a}'$, i.e. $\frac{d\vec{u}}{dt} = \frac{d\vec{u}'}{dt}$ since $t = t'$ by the absolute time assumption. Hence, by integration

$$\vec{u} = \vec{u}' + \vec{v} \tag{1.3}$$

for some constant of integration $\vec{v}$. We can use our assumptions to say what this constant is. If we take a particle that is not moving in the frame $\mathcal{O}'$, i.e. $\vec{u}' = \vec{0}$, then from the perspective of $\mathcal{O}$, it is moving at velocity $v_x\hat{x} + v_y\hat{y} + v_z\hat{z}$. Hence $\vec{v} = v_x\hat{x} + v_y\hat{y} + v_z\hat{z}$. Since velocity is the derivative of position with respect to time, another integration yields

$$\vec{r} = \vec{r}\,' + t\vec{v} + \vec{s} \tag{1.4}$$

for some constant of integration $\vec{s}$. The frames $\mathcal{O}$ and $\mathcal{O}'$ can be chosen so that if the particle is at $\vec{0}$ in $\mathcal{O}$, the position of the particle is at $\vec{0}'$ in $'$, and that these two coincide at time $t = 0$. This forces $\vec{s} = \vec{0}$. Therefore,

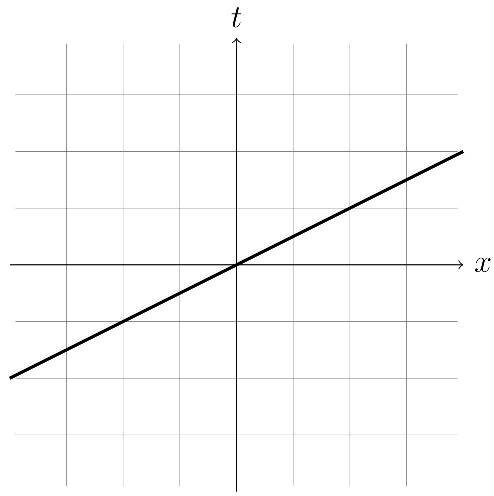
$$\vec{r} = \vec{r}\,' + t\vec{v}. \tag{1.5}$$

This can be used to define a linear transformation $\mathbb{R}^4_{\mathcal{O}'} \leftarrow \mathbb{R}^4_{\mathcal{O}}$ taking us from the coordinates used by $\mathcal{O}$ to the coordinates used by $\mathcal{O}'$. In matrix form, it is given by

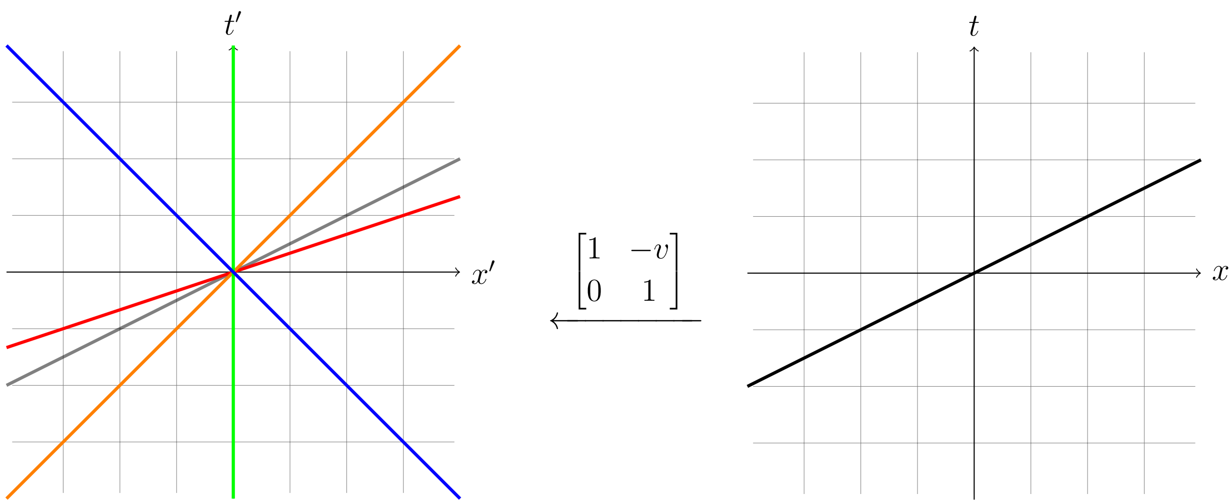
$$\begin{bmatrix} x' \\ y' \\ z' \\ t' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -v_x \\ 0 & 1 & 0 & -v_y \\ 0 & 0 & 1 & -v_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix}. \tag{1.6}$$

Example 1.7. Suppose observer $\mathcal{O}$ rolls a ball in the $\hat{x}$ direction and it moves at a constant velocity $u_x = 2$ (in some choice of units). In spacetime, ignoring the y and z coordinates, this would look like

[1] This isn't a very mathematically precise definition. Unfortunately, one definition I am aware of already assumes that spacetime is of a particular form. Furthermore, it assumes you know what a spacetime metric is as well as vector fields and covariant derivatives. I don't want to spend time defining these, so we will take the more physical definition for granted.



because the ball moves 2 units on the x axis for every 1 unit of time. Now consider an observer $\mathcal{O}'$ moving at different velocities $v \in \{-1, -1, 2, 3\}$ with respect to $\mathcal{O}$ along the x direction. The observed position as a function of time in this new frame is depicted in the following figure



since the observed velocities u' for $\mathcal{O}'$ are given by

$$v = -1 \implies u' = 3, \quad v = 1 \implies u' = 1, \quad v = 2 \implies u' = 0, \quad \text{and } v = 3 \implies u' = -1. \tag{1.8}$$

Notice that if $\mathcal{O}'$ is going faster than the object, then this object appears to be moving backwards.

Let us call each of these individual transformations along a particular direction a *Galilean boost* so that boosts along the three axes are given by

$$G_x := \begin{bmatrix} 1 & 0 & 0 & -v_x \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad G_y := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -v_y \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \text{and} \quad G_z := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -v_z \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{1.9}$$

Notice that the order in which G_x, G_y, and G_z are multiplied is immaterial. The result of multiplying all three gives us a general Galilean boost in an arbitrary direction.

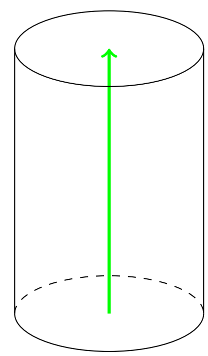
Also notice that if there is a third observer $\mathcal{O}''$ moving at a velocity v' with respect to $\mathcal{O}'$ in the same direction, then the velocity of $\mathcal{O}''$ as observed by $\mathcal{O}$ would be the composition of the linear transformations

$$\begin{bmatrix} 1 & v' \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & v \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & v+v' \\ 0 & 1 \end{bmatrix} \tag{1.10}$$

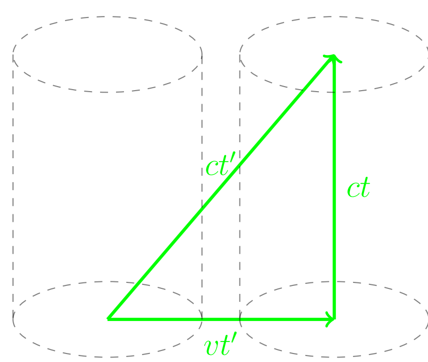
provided that we have ignored the y and z components (which remain unchanged). The reason a minus sign does not appear here is because we are applying the transformation from above in reverse. This result makes sense because of the intuitive algebra of adding velocities.

2 The constancy of the speed of light in inertial frames

It is an observed phenomenon that the speed of light c is the same in all inertial frames [1]. This is in stark contrast to the addition of velocities obtained from Galilean relativity because one would think that the velocity of light in a faster moving frame would appear to be moving faster as viewed by a stationary observer. Hence, we will make the *assumption* that the speed of light is constant in *all* inertial frames. We will first see how different frames perceive time. Since the speed of light is constant, we can devise a clock by counting the number of reflections inside of a cylindrical box with mirrors on its ends [3].



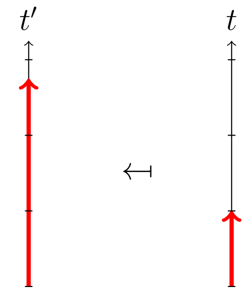
Suppose that each observer has such a clock. Now, suppose observer $\mathcal{O}'$ moves at velocity v in some fixed direction with respect to $\mathcal{O}$. Let t be the time that passes by in frame $\mathcal{O}$ for the light to go from the bottom of their box to the top of their box. Because the speed of light is c, the distance traversed is ct, the distance from the bottom to the top of the box. Let t' be the time that passes for $\mathcal{O}'$ during that traversal of light from the bottom to the top of the box when $\mathcal{O}'$ is looking at the light in the box carried by $\mathcal{O}$. Because the cylinder carried by $\mathcal{O}$ is also moving with velocity v (just in the negative direction), the distance, ct', that light has to travel to go from the bottom of the box to the top is actually greater than ct.



By Pythagorean's theorem,

$$c^2t'^2 = v^2t'^2 + c^2t^2 \qquad \Rightarrow \qquad t' = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}t \tag{2.1}$$

is the time that passes for $\mathcal{O}'$ during the passage of time t for $\mathcal{O}$. The reason for the formula written this way can be explained by considering what happens to a unit of time in $\mathcal{O}$ as viewed by $\mathcal{O}'$ as shown in the figure.



Hence, it appears that clocks in $\mathcal{O}$ appear to move slower than the clocks in $\mathcal{O}'$ from the perspective of $\mathcal{O}'$. Confusingly (though not contradictory), this argument is interchangeable and $\mathcal{O}$ will reach the same conclusion as $\mathcal{O}'$, namely that $\mathcal{O}$ will see the clocks in $\mathcal{O}'$ to move slower than the clocks in $\mathcal{O}$. The coefficient above appears so frequently that it is given a notation

$$\gamma := \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}. \tag{2.2}$$

Note that this factor depends on the relative speed between the two frames.

The above result, known as *time dilation* is mind-blowing—time can no longer be considered to be absolute. This has many other significant consequences. For example, distances are perceived differently by different inertial observers. To see this, we should first define how we will measure distances in each frame. For time, we agreed that we would count the number of times that light inside a cylindrical box reflects off the mirrors on its ends. For distance, we should use the constancy of the speed of light to define a scale of distance. For example, we could set up a scale where the basic unit of distance is given by the distance that light travels in one unit of time. We could, for example, choose this unit to be exactly the time it takes for light to go from the bottom of the box to the top. Regardless of our choice, to measure the distance from some point to another point, we would send light out from the initial point towards the destiny, wait for it to come back, and use our clock to count the number of times our clock ticked before this signal returned to us. Then, we would divide by two (since the light travels to and fro) to obtain the time t it takes to arrive at the destiny. We would then *define* the distance x between these points *in our frame* to be given by $x = ct$. Distance in $\mathcal{O}'$ is defined in precisely the same way so that $x' = ct'$. However, in order to translate from $\mathcal{O}$ to $\mathcal{O}'$, we have already discovered that $t' = \gamma t$, which together with $x' = ct'$, implies that

$$x' = \gamma x. \tag{2.3}$$

This is the readjustment $\mathcal{O}'$ needs to make when translating distance measurements made in $\mathcal{O}$ to their frame. This means that a given distance x in $\mathcal{O}$ as viewed by $\mathcal{O}'$ corresponds to the distance $x' = \gamma x$ in the frame of $\mathcal{O}'$. Since $\gamma \geq 1$, this means that from the perspective of $\mathcal{O}'$, the ruler used by $\mathcal{O}$ seems short compared to the ruler used by $\mathcal{O}'$. For example, if $\mathcal{O}$ says that the distance between two points (in the direction of relative motion for the two frames) is 100 meters, $\mathcal{O}'$ will disagree and say that the distance is actually less than 100 meters as measured by them. The distance that $\mathcal{O}'$ measures for this distance x is

$$x = \frac{1}{\gamma} x'. \tag{2.4}$$

This effect is known as *length contraction*.

The following example illustrating the significance of this in a comical way comes from [3].

Example 2.5. A bomb that has an internal clock in it is set to detonate in 20 seconds and has been placed in Minagi's rocket ship, but it can be immediately disarmed if Minagi can reach planet X, which is 6×10^{10} meters (20 light seconds) away as measured by the stationary frame. Will Minagi be able to disarm the bomb if he travels at 0.8 times the speed of light towards planet X?

To answer this question, let us calculate the time it would take in the stationary frame to reach planet X:

$$t = \frac{\text{distance}}{\text{velocity}} = \frac{20c}{0.8c} = 25 \text{ seconds.} \tag{2.6}$$

One might, incorrectly, argue that Minagi will explode on his way to planet X. But how much time would pass by in Minagi's frame as he is traveling? For the stationary frame, Minagi's clocks are slowing down and the time that passes by for Minagi is actually given by

$$t' = \sqrt{1 - 0.8^2}(25) = 15 \text{ seconds.} \tag{2.7}$$

Therefore, Minagi will be safe! Dually, one can think of this in terms of length contraction. The distance Minagi has to travel is less in his moving frame than the distance travelled in the stationary frame, and the distance Minagi claims that he needs to travel is

$$x = \sqrt{1 - 0.8^2}(20c) = 12 \text{ light seconds} \tag{2.8}$$

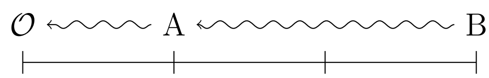
instead of 20 light seconds. Hence, the time it would take Minagi to reach planet X, which from his perspective is coming towards him at 0.8 times the speed of light, is

$$t' = \frac{12c}{0.8c} = 15 \text{ seconds.} \tag{2.9}$$

3 Simultaneity and order in spacetime

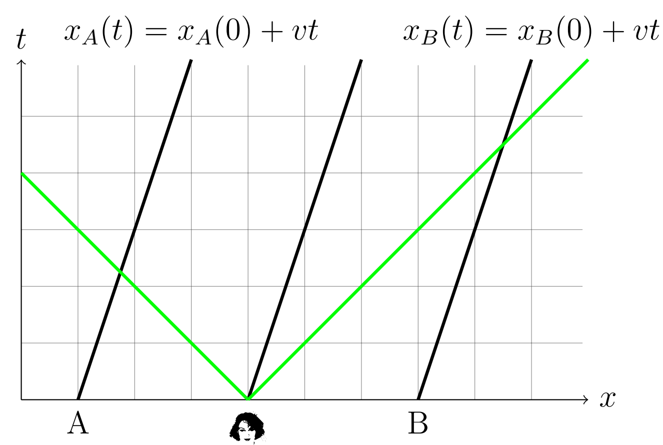
An <u>*event*</u> is a point in spacetime, which is assumed to be an absolute concept. Note that different inertial observers will associate different *coordinates* that they use to *quantify* the position and time at which an event occurs. To see why this is so, if I draw something on the board and ask you to quantify its location, you would specify an origin from which you will make your measurements, then you will choose two directions and a scale for each of them to count how many units in each direction you need to move to arrive at this drawing. In special relativity, we also have to specify when things occur and each frame will have their associated notion of time. In order to specify the coordinates of time, a given observer $\mathcal{O}$ first needs to decide when two events are *simultaneous.*

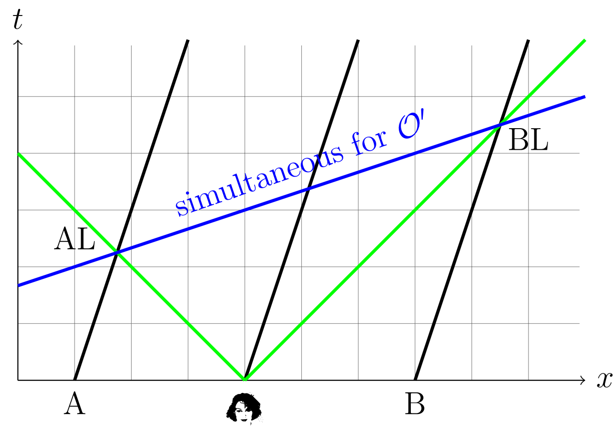
It is easy to decide when two things occur simultaneously at the location of the observer, but what about at larger distances? To make this decision, let us first consider two objects, stationary with respect to the observer, both located the same distance d away from the observer (not necessarily in the same direction). Imagine these two objects are bombs. When the detonation takes place, it will take the same amount of time for the light from each bomb to reach the observer. Hence, the explosions can be said to be *simulatenous* if and only if the signals reach the observer simultaneously (this is not circular because we have already established that simultaneity in the vicinity of the observer is well-defined). This now defines global simultaneity provided that the distance to each point is known. For example, if bomb A is 1 light second away from the observer and bomb B is 3 light seconds away from the observer, and the observer sees bomb A explode 1 second before bomb B, the observer can deduce that actually bomb B exploded 1 second *earlier* than bomb A.



Hence, *given* a reference frame $\mathcal{O}$, we can define a relation $\leq$ on the set of events. We say that an event $(x_1, y_1, z_1, t_1) \leq (x_2, y_2, z_2, t_2)$ if and only if $t_1 \leq t_2$. This relation is symmetric, transitive, and every event can be compared, but it is not reflexive because two different events can occur at the same time.

Let A and B be two specified locations in a frame $\mathcal{O}'$ that is moving in the direction from A to B with velocity v with respect to a stationary frame $\mathcal{O}$. Furthermore, suppose that the observer $\mathcal{O}'$ is situated midway between A and B in their frame and that A and B are stationary in the frame $\mathcal{O}'$. The time it takes for light to travel from the observer to each of these points and back is the same. Hence, light arrives at these two locations at the same time for $\mathcal{O}'$. Does $\mathcal{O}$ agree with this conclusion [1]? Namely, does $\mathcal{O}$ agree that the light signal arrives at these two locations at the same time in their reference frame? To answer this question, it helps to draw a spacetime diagram depicting the trajectory of the points A, B, and the origin of the light signal, drawn as , as viewed by the frame $\mathcal{O}$. The units of this graph have been chosen so that $c = 1$, or equivalently, in such a way so that the time axis is actually in units of ct instead of t. The straight line connecting the intersection of the light rays with A and B describes the line of events occurring simultaneously for $\mathcal{O}'$. Furthermore, every line parallel to this one describes all the events that occurring simultaneously for $\mathcal{O}'$.



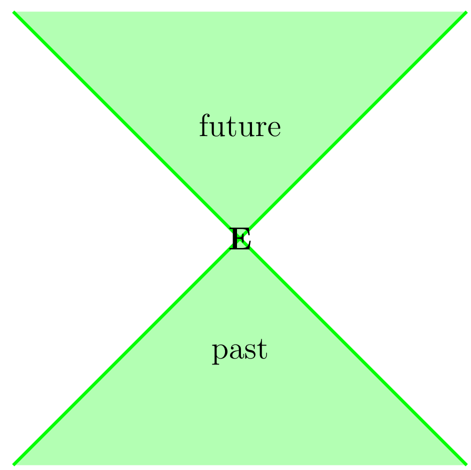


Let's denote the event that the light signal reaches A and B by AL and BL, respectively. For $\mathcal{O}$, event AL occurs *before* event BL. One might therefore become suspicious of this result for the following reason. If an observer $\mathcal{O}''$ moved at a velocity v with respect to $\mathcal{O}'$, then a similar argument would show that $\mathcal{O}''$ observes event BL to occur *before* AL (because $\mathcal{O}'$ is moving at velocity $-v$ with respect to $\mathcal{O}''$). In other words, there seem to be frames in which all possible orders of events are allowed. What happens to order? What happens to causality? Notice that earlier, we defined an order on the space of events *with respect to* a given frame. What we would like to have is a notion of when events are in the future or past of an event *irrespective* of any reference frame.

To address this, given any event E in spacetime, the *(causal) past* of E is the set of events F for which at least one of the following conditions hold:

(a) there exists an inertial reference frame where E and F occur at the same position with E occurring after F

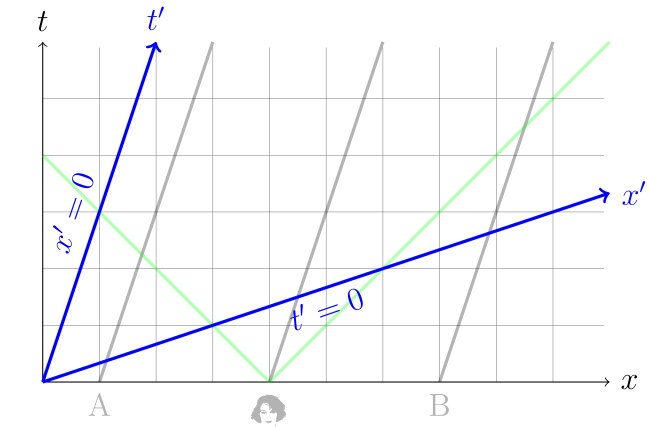
(b) a light signal from F reaches E in finite time.



A similar definition is made for *(causal) future*. We will be able to make this definition more precise after discussing Lorentz transformations and the invariant distance. These definitions are well-defined because the speed of light is constant in all inertial reference frames. However, what about non-inertial reference frames? One can still define this using four-vectors and the Minkowski metric, but instead of doing this, let's move on to discuss Lorentz transformations and some of the other consequences of special relativity.

4 Lorentz transformations

By the observation of the relativity of simultaneity, we can describe the transformation needed to go from one inertial reference frame to another. The $t' = 0$ line is determined being the unique line that is parallel to the line connecting the two simultaneous events from before and intersects $t = 0$ at $x = 0$ for the frame $\mathcal{O}$. This means that the two clocks in the two frames both begin at zero. The $x' = 0$ line is determined by assuming that the origins of the two frames agree at $t = 0$. The trajectory of the origin must be parallel to the trajectory of the observer $\mathcal{O}'$ as viewed by $\mathcal{O}$.



Therefore, we can assume that the transformation $\mathbb{R}^2_{\mathcal{O}'} \leftarrow \mathbb{R}^2_{\mathcal{O}}$ is described by a linear transformation of the form

$$\begin{bmatrix} x' \\ t' \end{bmatrix} = \begin{bmatrix} a & -b \\ -d & e \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} = \begin{bmatrix} ax - bt \\ -dx + et \end{bmatrix}, \quad (4.1)$$

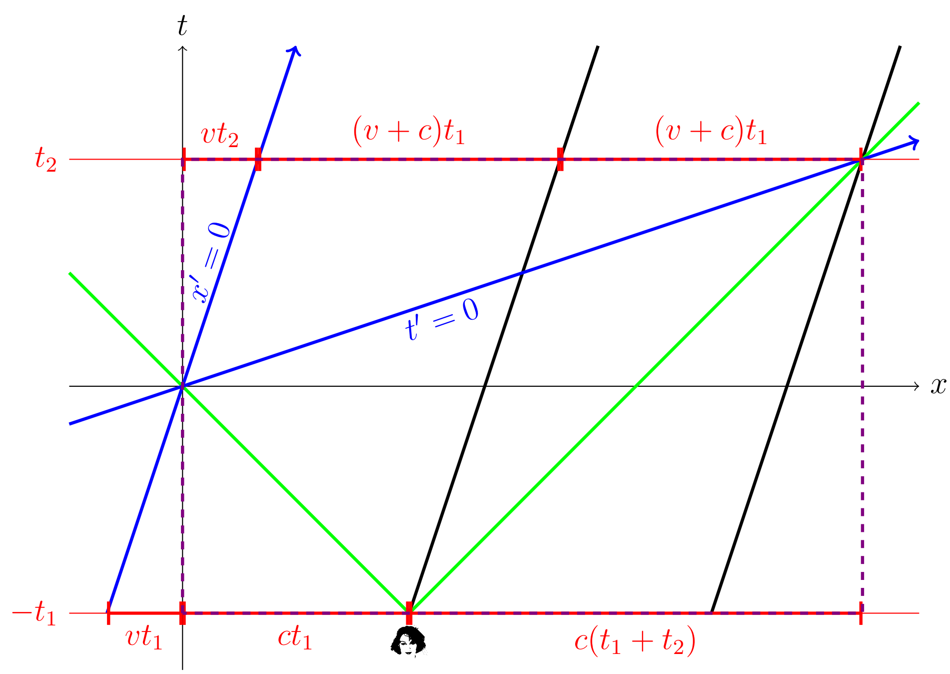
where a, b, d, e are some constants to be determined momentarily and where we have ignored the y and z coordinates because they are perpendicular to the direction of motion (and will be unchanged). This transformation will describe what changes must be made to the measurements done in the frame $\mathcal{O}$ to arrive at the corresponding measurements that $\mathcal{O}'$ would have made.

As explained in the above figure, if an object is placed at position $x' = 0$ at time $t' = 0$ then the object will have trajectory $x = vt$ as viewed in $\mathcal{O}$. Hence,

$$0 = x' = ax - bt \qquad \Rightarrow \qquad v = \frac{x}{t} = \frac{b}{a}. \quad (4.2)$$

Furthermore, when $t = 0$, then $x' = \gamma x$ by our earlier result (2.3). Hence, $a = \gamma$ and therefore $b = av = \gamma v$. This shows that $x' = \gamma x - \gamma vt$.

To obtain the lower part of the linear transformation in (4.1), note that when $x = 0$, we get time dilation $t' = \gamma t$ so that $e = \gamma$. To obtain the other entry in the matrix, we use some basic geometry and consider a similar scenario as before with equally spaced objects and light signals[2]



[2] It is an unfortunate coincidence that $t_1 = t_2$ in this figure. If the slope of the line of simultaneity were smaller, t_2 would have been less than t_1 and if the slope were greater, t_2 would have been greater than t_1. Indeed, imagine that the velocity was zero. Then t_2 would have been zero.

The top and bottom of the dashed rectangle are equal so that

$$vt_2 + 2(v+c)t_1 = ct_1 + c(t_1 + t_2), \tag{4.3}$$

which we can solve for t_1 in terms of t_2 as

$$t_1 = \frac{c-v}{2v} t_2. \tag{4.4}$$

Using this result, the slope of the $t' = 0$ line is given by

$$\frac{t_2}{vt_2 + 2(v+c)t_1} = \frac{t_2}{vt_2 + 2(v+c)\left(\frac{c-v}{2v}\right)t_2} = \frac{v}{c^2}. \tag{4.5}$$

But, the slope of this line is precisely $\frac{t}{x}$ when $t' = 0$ (this line of simultaneity) and the relation to the other coefficients is therefore given by

$$0 = t' = -dx + et \qquad \Rightarrow \qquad \frac{v}{c^2} = \frac{t}{x} = \frac{d}{e}. \tag{4.6}$$

Using the fact that $e = \gamma$ from before, we conclude that $d = \frac{\gamma v}{c^2}$. Hence, the matrix describing the transformation is

$$\mathbb{R}^2_{\mathcal{O}'} \xleftarrow{\begin{bmatrix} \gamma & -\gamma v \\ -\frac{\gamma v}{c^2} & \gamma \end{bmatrix}} \mathbb{R}^2_{\mathcal{O}} \tag{4.7}$$

One can show that the perception of distances along directions orthogonal to the direction of motion are unchanged so that the full linear transformation describing the change of coordinates from $\mathcal{O}$ to $\mathcal{O}'$ is given by

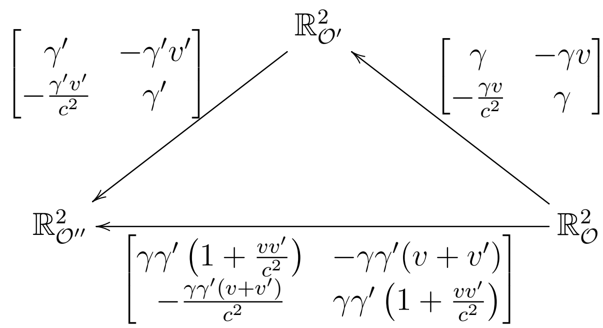
$$\begin{array}{ccc} \mathbb{R}^4_{\mathcal{O}'} & \xleftarrow{\begin{bmatrix} \gamma & 0 & 0 & -\gamma v \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{\gamma v}{c^2} & 0 & 0 & \gamma \end{bmatrix}} & \mathbb{R}^4_{\mathcal{O}} \\ \begin{bmatrix} x' \\ y' \\ z' \\ t' \end{bmatrix} & \longleftarrow\!\shortmid & \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} \end{array} \tag{4.8}$$

Using similar arguments, one can obtained similar transformations if $\mathcal{O}'$ is moving with a velocity along the y and z directions. These give the three *Lorentz boosts* and are analogous to the Galilean boosts introduced earlier

$$L_x := \begin{bmatrix} \gamma_x & 0 & 0 & -\gamma_x v_x \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{\gamma_x v_x}{c^2} & 0 & 0 & \gamma_x \end{bmatrix}, \; L_y := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \gamma_y & 0 & -\gamma_y v_y \\ 0 & 0 & 1 & 0 \\ 0 & -\frac{\gamma_y v_y}{c^2} & 0 & \gamma_y \end{bmatrix} \text{ and } L_z := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma_z & -\gamma_z v_z \\ 0 & 0 & 0-\frac{\gamma_z v_z}{c^2} & \gamma_z \end{bmatrix}, \tag{4.9}$$

respectively. One can also write down what a Lorentz boost is in an arbitrary direction, though this is not so simple. Instead, we will consider successive boosts.

First, let us consider multiple boosts in the same direction described by an observer $\mathcal{O}'$ moving at velocity v with respect to $\mathcal{O}$ and an observer $\mathcal{O}''$ moving at velocity v' with respect to $\mathcal{O}'$, all in the same direction. For simplicity, let us only focus on boosts along the x direction and ignore the y and z coordinates since they do not change in this case. The associated coordinate transformation that takes us from $\mathcal{O}$ coordinates to $\mathcal{O}''$ coordinates is the the composite linear transformation



(4.10)

since it is given by matrix multiplication.

$$\begin{bmatrix} \gamma' & -\gamma' v' \\ -\frac{\gamma' v'}{c^2} & \gamma' \end{bmatrix} \begin{bmatrix} \gamma & -\gamma v \\ -\frac{\gamma v}{c^2} & \gamma \end{bmatrix} = \begin{bmatrix} \gamma\gamma'\left(1+\frac{vv'}{c^2}\right) & -\gamma\gamma'(v+v') \\ -\frac{\gamma\gamma'(v+v')}{c^2} & \gamma\gamma'\left(1+\frac{vv'}{c^2}\right) \end{bmatrix}. \tag{4.11}$$

However, we know that this matrix represents the net coordinate transformation, which must be of the form

$$\begin{bmatrix} \gamma'' & -\gamma'' v'' \\ -\frac{\gamma'' v''}{c^2} & \gamma'' \end{bmatrix}. \tag{4.12}$$

Comparing these two matrices, we obtain the following relations

$$\gamma'' = \gamma\gamma'\left(1+\frac{vv'}{c^2}\right) \qquad \& \qquad -\gamma'' v'' = -\gamma\gamma'(v+v'). \tag{4.13}$$

Plugging the first into the latter gives

$$-\gamma\gamma'\left(1+\frac{vv'}{c^2}\right) v'' = \gamma\gamma'(v+v') \qquad \Rightarrow \qquad v'' = \frac{v+v'}{1+\frac{vv'}{c^2}}, \tag{4.14}$$

a highly counter-intuitive formula! Using this we can also check that

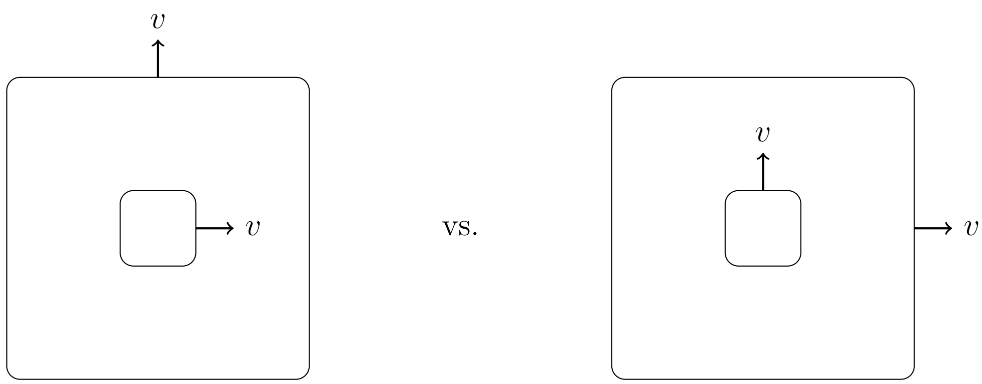
$$\gamma'' = \frac{1}{\sqrt{1+\frac{v''^2}{c^2}}} \tag{4.15}$$

though I'll let you check that out for yourself. The formula for v'' says that summing relativistic velocities is not the same as summing non-relativistic velocities—we cannot just add the velocities using ordinary addition. Instead, the addition is skewed by a factor (depending on the velocities) so as to guarantee that the net velocity is still bounded from above by c, the speed of light. Also notice that if $v' = -v$, then $v'' = 0$, and in fact, we have

$$\begin{bmatrix} \gamma' & -\gamma' v' \\ -\frac{\gamma' v'}{c^2} & \gamma' \end{bmatrix} \begin{bmatrix} \gamma & -\gamma v \\ -\frac{\gamma v}{c^2} & \gamma \end{bmatrix} = \begin{bmatrix} \gamma & \gamma v \\ \frac{\gamma v}{c^2} & \gamma \end{bmatrix} \begin{bmatrix} \gamma & -\gamma v \\ -\frac{\gamma v}{c^2} & \gamma \end{bmatrix} = \begin{bmatrix} \gamma^2(1-\frac{v^2}{c^2}) & 0 \\ 0 & \gamma^2(1-\frac{v^2}{c^2}) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{4.16}$$

so that boosting by the same speed but in the opposite direction is the inverse of the original boost as we might have expected.[3] Fortunately, because matrix multiplication is associative, this new definition of summing velocities is commutative and associative provided that the velocities are in the same direction.

But... what happens if we boost in different directions? This brings us to our next surprise of special relativity: two Lorentz boosts in different directions induce a rotation! We will illustrate this in the simplest case possible. Consider boosts in the x and y directions of the same magnitude v. We will compare $L_y L_x$ with $L_x L_y$.

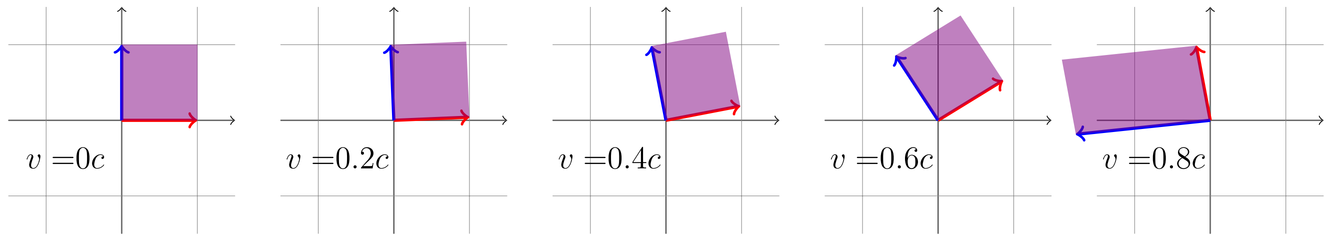


To do this, we will calculate the transformation $L_y^{-1} L_x^{-1} L_y L_x$, which says to boost in x, then boost in y, then boost backwards along x and finally backwards along y. This expression is known as the *group commutator* of L_y and L_x (it appears in many other contexts in mathematics). For Galilean transformations, boosting in two different directions does not depend on the order in which the boosts are applied. Hence, the end result of the group commutator is the identity transformation (as you would intuitively expect). Is this also true for Lorentz transformations? A lengthy calculation shows us that we obtain (we have ignored the z direction since it remains unchanged in this calculation)

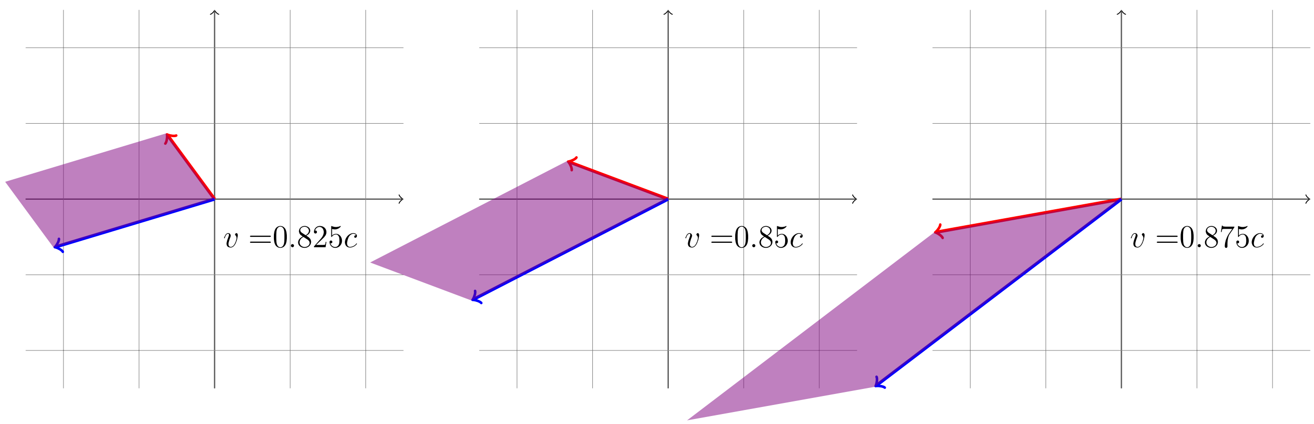
$$\begin{aligned} L_y^{-1} L_x^{-1} L_y L_x &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \gamma & \gamma v \\ 0 & \frac{\gamma v}{c^2} & \gamma \end{bmatrix} \begin{bmatrix} \gamma & 0 & \gamma v \\ 0 & 1 & 0 \\ \frac{\gamma v}{c^2} & 0 & \gamma \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \gamma & -\gamma v \\ 0 & -\frac{\gamma v}{c^2} & \gamma \end{bmatrix} \begin{bmatrix} \gamma & 0 & -\gamma v \\ 0 & 1 & 0 \\ -\frac{\gamma v}{c^2} & 0 & \gamma \end{bmatrix} \\ &= \begin{bmatrix} \gamma & 0 & \gamma v \\ \frac{\gamma^2 v^2}{c^2} & \gamma & \gamma^2 v \\ \frac{\gamma^2 v}{c^2} & \frac{\gamma v}{c^2} & \gamma^2 \end{bmatrix} \begin{bmatrix} \gamma & 0 & -\gamma v \\ \frac{\gamma^2 v^2}{c^2} & \gamma & -\gamma^2 v \\ -\frac{\gamma^2 v}{c^2} & -\frac{\gamma v}{c^2} & \gamma^2 \end{bmatrix} \\ &= \begin{bmatrix} \gamma^2 \left(1 - \frac{\gamma v^2}{c^2}\right) & -\frac{\gamma^2 v^2}{c^2} & \gamma^2 v(\gamma - 1) \\ \frac{\gamma^3 v^2}{c^2}(2 - \gamma) & \gamma^2 \left(1 - \frac{\gamma v^2}{c^2}\right) & \gamma^3 v \left(\gamma - \frac{v^2}{c^2} - 1\right) \\ \frac{\gamma^3 v}{c^2} \left(1 + \frac{v^2}{c^2} - \gamma\right) & \frac{\gamma^2 v}{c^2}(1 - \gamma) & \gamma^3 \left(\gamma - \frac{2v^2}{c^2}\right) \end{bmatrix}, \end{aligned} \tag{4.17}$$

which is certainly not the identity. The images of the standard unit square in $\mathbb{R}^2$ under this transformation for various values of v are shown below.

[3]but we should be careful about what we "expect" when it comes to special relativity!



For a velocity that is 20% the speed of light, there is very little deviation from the identity, but some is indeed visible. Notice how drastic the situation becomes when the velocity is very close to the speed of light. For example, even below 90% the deviation becomes significant.



We can make some sense of the apparent rotation for smaller velocities if we expand out the terms in orders of $\frac{v}{c}$. If we approximate this up to second order in $\frac{v}{c}$, we obtain

$$L_y^{-1} L_x^{-1} L_y L_x \approx \begin{bmatrix} 1 & -\frac{v^2}{c^2} & \frac{v}{2}\frac{v^2}{c^2} \\ \frac{v^2}{c^2} & 1 & \frac{v}{2}\frac{v^2}{c^2} \\ 0 & 0 & 1 \end{bmatrix} \approx \begin{bmatrix} \cos\left(\frac{v}{c}\right) & -\sin\left(\frac{v}{c}\right) & \frac{v}{2}\frac{v^2}{c^2} \\ \sin\left(\frac{v}{c}\right) & \cos\left(\frac{v}{c}\right) & \frac{v}{2}\frac{v^2}{c^2} \\ 0 & 0 & 1 \end{bmatrix}, \tag{4.18}$$

which, when restricted to the xy plane, is a rotation in the xy plane counter-clockwise by $\frac{v}{c}$ radians. Remember, this is only an approximation and only applies to small velocities. The expression (4.17) is exact and the images indicate that for velocities significantly closer to the speed of light, the transformation is far from a rotation. In the images above, one sees that the transformation is still approximately close to being a rotation for $v = 0.6c$ but has shifted dramatically from a rotation when $v = 0.8c$.

5 Energy, mass, and Einstein's equation

Although we will not derive the result here, one can show that demanding the conservation of energy and momentum, together with keeping a similar form to Newtonian equations for collisions (see Chapter 6 of [1]), the mass of a massive particle depends on its velocity v via

$$m(v) = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}} = \gamma m_0, \tag{5.1}$$

where m_0 is the mass of the same particle as measured in its rest frame. In other words, although the faster you go, the thinner you get, you also get heavier! Furthermore, and more simply, by expanding out this expression in terms of $\frac{v}{c}$, one obtains

$$m(v) = m_0 + \frac{1}{2} m_0 \frac{v^2}{c^2} + \mathcal{O}\left(\frac{v^4}{c^4}\right). \tag{5.2}$$

Multiplying by c^2, we see that

$$m(v)c^2 = m_0 c^2 + \frac{1}{2} m_0 v^2 + v^2 \mathcal{O}\left(\frac{v^2}{c^2}\right). \tag{5.3}$$

The second term in this expression is the kinetic energy of a particle of mass m_0 moving at velocity v. In the absence of a potential, this is normally taken to be the energy of the particle. The additional term $m_0 c^2$ suggests that there is more energy intrinsic in a particle than just its kinetic energy. This term is called the *rest mass* of the massive particle. Furthermore, we define the total energy of a massive particle of mass m moving at velocity v subject to no potential to be

$$E := \gamma m_0 c^2. \tag{5.4}$$

The meaning of this additional rest energy term is important. It is possible to transmute this mass to energy in other forms, such as photons or other particles.[4]

Example 5.5. The mass of a proton is m_p and the mass of an electron is m_e. The Hydrogen atom is composed of a proton and an electron, and yet its mass is m_H, which is known to be less than $m_p + m_e$. The difference between these masses is given by

$$m_p + m_e - m_H = 2.4826 \times 10^{-35} \text{ kg}. \tag{5.6}$$

The Bohr radius of the electron (the average distance between the electron in its lowest energy state and the proton in the nucleus of the Hydrogen atom) is r_e. The potential energy due to the Coulombic interaction is $-\frac{ke^2}{r_e}$, where k is Coulomb's constant and e is the magnitude of the electric charge of the electron (and proton). The Bohr radius r_e and Coulomb's constant k are related via

$$k = \frac{\hbar^2}{m_e e^2 r_e}. \tag{5.7}$$

Hence, the mass associated with this potential energy is given by

$$\frac{ke^2}{r_e c^2} = \frac{\hbar^2}{m_e r_e^2 c^2} = 4.85087 \times 10^{-35} \text{ kg}. \tag{5.8}$$

If we also assume (semi-classically) that the electron is orbiting the proton (which is so heavy in comparison to the electron that we can assume the orbit is circular) at the Bohr radius, then this means the electron also has kinetic energy. If we assume that the speed is small compared to the speed of light, we can equate $\frac{ke^2}{r_e^2} = F = m_e a$, where a is the centripetal acceleration of the

[4]Given certain inputs, the possible particles that are allowed to be produced from this internal energy and other particles is dictated by quantum field theory and the interactions that are possible in these theories.

electron. The centripetal acceleration is related to the transverse velocity via $a = \frac{v^2}{r_e}$. Hence, the speed of the electron is

$$v = \sqrt{\frac{k}{mr_e}}e. \tag{5.9}$$

Using the relationship (5.7) between k and r_e, the speed is

$$v = \frac{\hbar}{m_e r_e}. \tag{5.10}$$

Therefore, the kinetic energy of the electron is given by

$$\frac{1}{2}m_e v^2 = \frac{\hbar^2}{2m_e r_e^2} \tag{5.11}$$

so that the mass that the kinetic energy contributes is

$$\frac{\hbar^2}{2m_e r_e^2 c^2} = 2.42543 \times 10^{-35} \text{ kg} \tag{5.12}$$

In the above calculation, we assumed that $F = m_e a$, which is only valid for small velocities (in general, the correct formula in special relativity is $F = \frac{dp}{dt}$, where p is the momentum). Now, to see that this assumption is valid, note that

$$v = \frac{\hbar}{m_e r_e} = \alpha c, \tag{5.13}$$

where α is the fine structure constant

$$\alpha = \frac{\hbar}{m_e c r_e} \approx \frac{1}{137}. \tag{5.14}$$

Hence, the velocity is much smaller than the speed of light and our Newtonian approximation for the force equation is a valid one. Since the kinetic energy adds a positive contribution to the mass of the Hydrogen atom and since the potential energy reduces the mass, special relativity dictates that the mass of the Hydrogen atom should be

$$m_H \stackrel{?}{=} m_p + m_e + \frac{1}{2}m_e v^2 - \frac{ke^2}{r_e c^2} \tag{5.15}$$

(we can ignore the gravitational attraction between the two because it is negligible compared to the Coulombic interaction). The sum of the kinetic energy and potential energy is given by

$$\frac{1}{2}m_e v^2 - \frac{ke^2}{r_e c^2} = \frac{\hbar^2}{2m_e r_e^2} - \frac{\hbar^2}{m_e r_e^2} = -\frac{\hbar^2}{2m_e r_e^2} = -2.42543 \times 10^{-35} \text{ kg}. \tag{5.16}$$

This agrees very closely with the above difference of masses. In units of energy, this is approximately −13.6 eV (electron volts). The meaning of this is that when an electron and proton bond to form a Hydrogen atom in its ground state, a photon of energy 13.6 eV is released from forming this bond. This loss of energy corresponds to a loss of mass to the proton and electron pair that have formed the Hydrogen atom. Special relativity is used to keep track of this change.

Remark 5.17. One should be a bit careful in the above analysis because the dynamics of the electron at this distance scale is dictated by quantum mechanics. In the above calculation, we have assumed that the electron orbits the proton at a fixed radius, the Bohr radius. In reality, the Bohr radius is the *average* distance between the electron and the proton in the ground state. Quantum effects do not associate a precise location to this electron and the energy difference should actually be obtained from Schrödinger's equation applied a quantum version of the Coulombic potential. Does our result change if we apply the methods of quantum mechanics?

By using the ground state wave function u_0 for the Hydrogen atom, one can show that the expectation value of the total energy (which is equal to the energy since u_0 is an energy eigenstate of the Hamiltonian) is

$$\langle H \rangle_{u_0} = -\frac{\hbar^2}{2m_e r_e^2 c^2}. \tag{5.18}$$

One can also calculate the kinetic and potential energies and get the same answers as we got in the semi-classical calculation above. Namely, one can show that

$$\left\langle \frac{P^2}{2m_e} \right\rangle_{u_0} = \frac{\hbar^2}{2m_e r_e^2 c^2} \qquad \& \qquad \left\langle -\frac{ke^2}{R} \right\rangle_{u_0} = -\frac{\hbar^2}{m_e r_e^2 c^2}, \tag{5.19}$$

where P is the momentum operator and R is the radial operator defined as $R := \sqrt{X^2 + Y^2 + Z^2}$, with X, Y, Z the position operators in the three different directions. It appears as though we were a bit lucky with our semi-classical calculation since the answers were exactly the same.

Similar analyses (not involving quantum mechanics) can be used to answer many interesting questions. For example, when carbon C and oxygen O form carbon monoxide, heat escapes and lowers the resulting mass of CO (see Problem 1-10 in [1]). A surprising result occurs due to the electrostatic repulsion of like charges. For example, if 1 gram of electrons could be confined to a sphere of 10 cm radius, the mass associated with the potential energy of this configuration would be on the order of 10 million tons (see Problem 1-8 in [1]).

6 Some peculiarities with packing mass and black holes

This section discusses some of the strange properties of combining ideas from gravity, special relativity, and quantum mechanics. I still have to think more deeply about Examples 6.1 and 6.6. They should not be taken too seriously.

Example 6.1. Let us play a strange game. Consider two (neutral charge) masses m_1 and m_2 that are separated by some distance d. By special relativity arguments, the net mass of the configuration is given by

$$m_1 + m_2 - \frac{Gm_1 m_2}{dc^2} \tag{6.2}$$

because gravity is an attractive force. Is it possible for the distance to be small enough for this net mass to become zero? Let's find out:

$$m_1 + m_2 - \frac{Gm_1 m_2}{dc^2} = 0 \iff d = \frac{G}{c^2}\frac{m_1 m_2}{m_1 + m_2}. \tag{6.3}$$

In the special case that $m_1 = m_2 = m$, this becomes

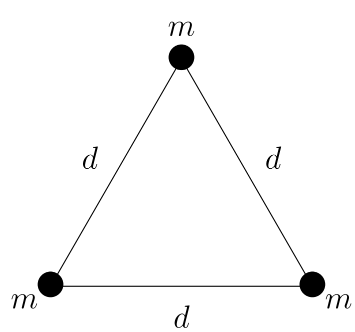
$$d = \frac{Gm}{2c^2}. \tag{6.4}$$

What does this mean? Is it possible to bring two neutral masses together so that their effective mass as viewed by someone who is far away suddenly becomes zero? Does the gravitational force suddenly become repulsive when the masses come even closer together so that their net mass is *negative*? Is any of this possible? Although quantum mechanics prevents such a possibility from occurring,[5] we can try to argue why this cannot happen from a purely classical and information theoretic level. Compare this distance to the Schwarzschild radius r associated with a mass m

$$r = \frac{2Gm}{c^2}. \tag{6.5}$$

They are on the same order of magnitude with the Schwarzschild radius four times as large. The meaning of the Schwarzschild radius is as follows. If a particle (massive or massless) is within the Schwarzschild radius of a mass m, then no amount of energy can be given to this particle so that it can escape from within this region in finite time. In other words, it is the region associated with m that acts as a black hole. The spherical shell of this Schwarzschild radius is known as the *event horizon* of the black hole. To understand the full classical effects of black holes, one should use Einstein's theory of general relativity. One of the consequences of this is that an outside observer, one who is stationary or out infinitely far away from the gravitational effects of the black hole, will see that a particle moving towards the horizon slows down dramatically as it approaches the horizon. In fact, the outside observer will conclude that the infalling particle will only arrive at the horizon in the infinite future. It is quite remarkable that nature creates a black hole before two massive particles can come close enough together so that their net mass becomes negative! Or does it?

Example 6.6. Let's keep playing the game from above, but this time consider three massive objects all separated by the same distance.



The net mass in this case is

$$3m - \frac{3Gm^2}{dc^2}. \tag{6.7}$$

Equating these gives

$$d = \frac{Gm}{c^2}, \tag{6.8}$$

[5]This is due to many factors including the Pauli exclusion principle, the strong force, etc.

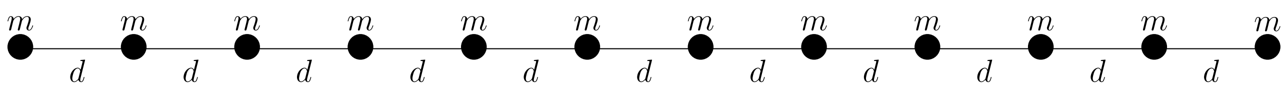
which is a larger distance than the case of just two masses. However, it is still smaller than the Schwarzschild radius. Therefore, one might ask if it is possible in *any* way to configure masses in $\mathbb{R}^3$ so that their net mass is negative but so that they are separated by a distance so that they lie outside of each other's event horizons. For example, with four equally spaced masses (so that they are on the vertices of a tetrahedron), the net mass is

$$4m - \frac{6Gm^2}{dc^2}. \tag{6.9}$$

Equating this to zero gives

$$d = \frac{3Gm}{2c^2}, \tag{6.10}$$

which is getting closer to the Schwarzschild radius for m. How about if we had $N+1$ masses all lined up along a line?



The gravitational potential energy would be given by

$$\begin{aligned}\frac{Gm^2}{d}\left(N + \frac{N-1}{2} + \frac{N-2}{3} + \cdots + \frac{1}{N}\right) &= \frac{Gm^2}{d}\sum_{i=1}^{N}\frac{N+1-i}{i} \\ &= \frac{Gm^2}{d}\left(-N + (N+1)\sum_{i=1}^{N}\frac{1}{i}\right)\end{aligned} \tag{6.11}$$

The net mass would be

$$(N+1)m - \frac{Gm^2}{dc^2}\left(-N + (N+1)\sum_{i=1}^{N}\frac{1}{i}\right). \tag{6.12}$$

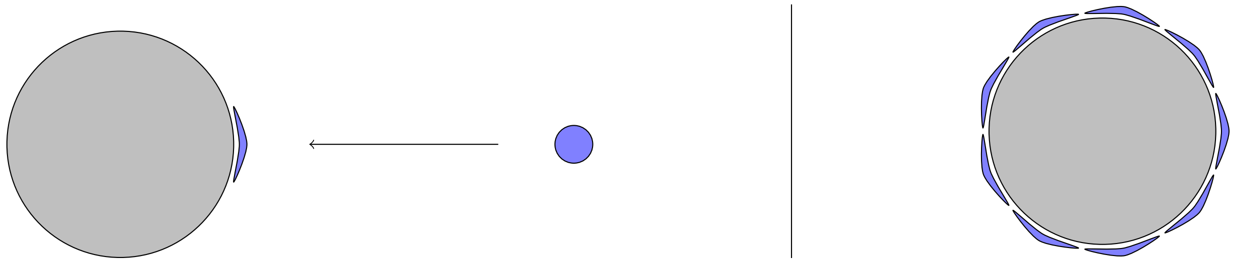
Setting this equal to zero as before and solving for d gives

$$d = \frac{Gm}{c^2}\left(\frac{-N}{N+1} + \sum_{i=1}^{N}\frac{1}{i}\right). \tag{6.13}$$

The amazing thing about this formula is that it asymptotically increases without bound as N increases. Therefore, there exists an N for which $d > r$, the Schwarzschild radius. This happens when $N = 10$, i.e. when there are 11 particles as in the picture above. In order for $d > 2r$, one needs $N = 82$, i.e. 83 particles. I'm not sure what this all means—I'll let you think about it.

Jacob Bekenstein asked questions that, I believe, are some of the most crucial towards our understanding of quantum gravity, a physical theory that describes both the subatomic scale but also the high gravity scale. Such situations are common in the vicinity of black holes, but also in neutron stars. He is also responsible for some highly surprising relationships between entropy, energy, and thermodynamics.

Example 6.14. Take my word for it but when matter falls into a black hole, as viewed by an outside observer, it falls in very slowly and asymptotically approaches the horizon [2], [4].



Hence, all the information that falls towards a black hole eventually stays near the horizon as viewed by a stationary observer. Classically, one could then argue that the thin layer around the horizon could contain an infinite amount of information. Bekenstein, motivated by this puzzle, and later Hawking, showed that the entropy of a black hole is given by

$$S = \frac{k_B A c^3}{4\hbar G} = \frac{k_B A}{4\ell_P^2}, \tag{6.15}$$

where A is the area of the black hole horizon, ℓ_P is the Planck length, and k_B is Boltzmann's constant. This formula says that the entropy of a black hole is naturally measured in $\frac{1}{4}$ units of Planck area. This is incredibly surprising. If we scaled a bag of rice down to a size at which it would become a black hole, this entropy formula says that the amount of rice you can stuff inside the bag is actually only enough to coat the outside surface area of the bag.

Bekenstein also went on to formulate his famous theoretical upper bound on information storage in space. The Bekenstein bound, as it is now called, is one of the most counter-intuitive information-theoretic consequences of the statistical mechanics of black holes. It states that the entropy S and energy E contained inside of a sphere of radius R are related by the following upper bound on the entropy

$$S \leq \frac{2\pi k_B r E}{\hbar c}. \tag{6.16}$$

Notice that in the case of a black hole of radius $r = \frac{2Gm}{c^2}$ and energy $E = mc^2$, this says

$$S \leq \frac{2\pi k_B r E}{\hbar c} = \frac{2\pi k_B r m c^2}{\hbar c} = \frac{2\pi k_B r^2 c^3}{2G\hbar} = \frac{k_B A}{4\ell_P^2}. \tag{6.17}$$

In other words, a black hole saturates this bound, which makes intuitive sense because adding any more mass/information to a black hole would increase its mass and therefore its radius. One cannot add any more information/mass into a black hole without increasing its size.

If you want to learn more about the bizarre world of black holes and information theory, I encourage you to read Jacob Bekenstein's papers.

Acknowledgements

I thank Keith Conrad for some helpful comments and I thank Jonathan Ben-Benjamin for several discussions.

References

[1] Anthony Philip French, *Special relativity*, CRC Press, 1968.

[2] James Lindesay and Leonard Susskind, *An Introduction To Black Holes, Information And The String Theory Revolution: The Holographic Universe*, World Scientific Publishing Company, 2004.

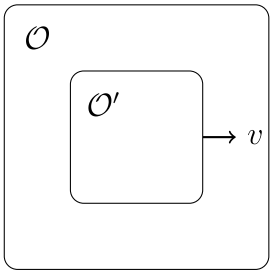
[3] Hideo Nitta, Masafumi Yamamoto, Keita Takatsu, and Trend-Pro Co., Ltd., *The Manga Guide to Relativity*, 1st ed., No Starch Press, 2011.

[4] Leonard Susskind, *Inside Black Holes*. Video lecture available at http://www.youtube.com/watch?v=yMRYZMv0jRE.

[5] Wikipedia contributors, *Inertial frame of reference*, Wikipedia, The Free Encyclopedia, January 26, 2018. Available as https://en.wikipedia.org/w/index.php?title=Inertial_frame_of_reference&oldid=822423025 (accessed February 4, 2018).

A reference frame is a choice of coordinates. An inertial frame of reference is a reference frame in which the trajectory of a particle that is not subject to any forces is in a straight line at constant velocity.

Galilean boosts are used to transform from one inertial frame and observer $\mathcal{O}$ to another $\mathcal{O}'$:



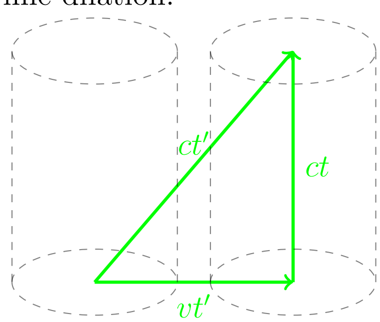
$$\begin{bmatrix} x' \\ t' \end{bmatrix} = \begin{bmatrix} 1 & -v \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix}$$

Addition of velocities is given by successive Galilean boosts:

$$\begin{bmatrix} 1 & v' \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & v \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & v'+v \\ 0 & 1 \end{bmatrix}$$

In special relativity, we assume that the speed of light is the same in all inertial frames. This has many consequences, including time dilation, length contraction, relativity of simultaneity, and the equivalence between mass and energy.

Time dilation:



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

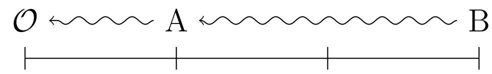
Length contraction:

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

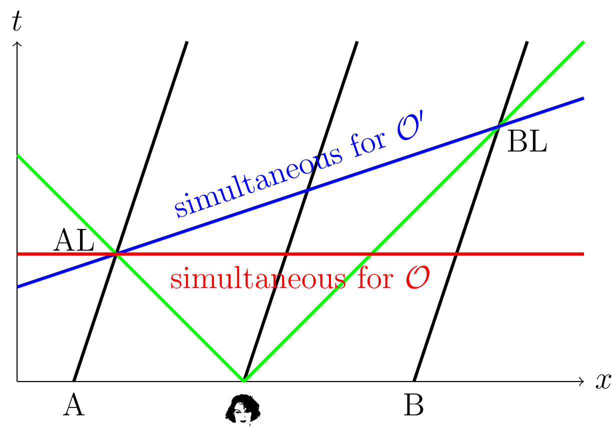
It is convenient to define

$$\gamma := \left(1-\frac{v^2}{c^2}\right)^{-1/2}$$

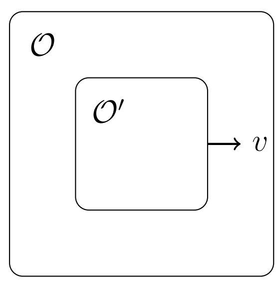
Simultaneity is defined using light signals, which propagate at a finite speed. For example, consider an observer $\mathcal{O}$ and a location A one light second away and a location B three light seconds away from $\mathcal{O}$. If light from A reaches $\mathcal{O}$ one second before the light from B, $\mathcal{O}$ concludes that B occurred one second earlier than A.



Simultaneity is relative: if $\mathcal{O}'$ is moving at velocity v with respect to $\mathcal{O}$ and AL and BL occur simultaneously for $\mathcal{O}'$, then AL occurs before BL for $\mathcal{O}$.



Lorentz boosts are used to transform from one inertial frame and observer $\mathcal{O}$ to another $\mathcal{O}'$:



$$\begin{bmatrix} x' \\ t' \end{bmatrix} = \begin{bmatrix} \gamma & -\gamma v \\ -\frac{\gamma v}{c^2} & \gamma \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix}$$

As $v \to 0$, this agrees with Galilean boosts.

Addition of velocities is given by successive Lorentz boosts:

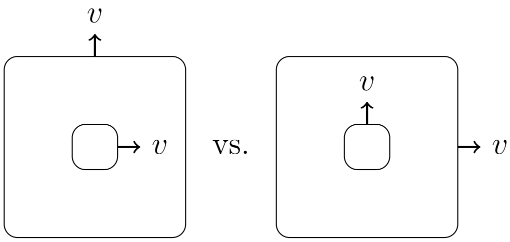
$$\begin{bmatrix} \gamma' & -\gamma' v' \\ -\frac{\gamma' v'}{c^2} & \gamma' \end{bmatrix} \begin{bmatrix} \gamma & -\gamma v \\ -\frac{\gamma v}{c^2} & \gamma \end{bmatrix} = \begin{bmatrix} \gamma'' & -\gamma'' v'' \\ -\frac{\gamma'' v''}{c^2} & \gamma'' \end{bmatrix}$$

where

$$v'' = \frac{v+v'}{\sqrt{1+\frac{vv'}{c^2}}}.$$

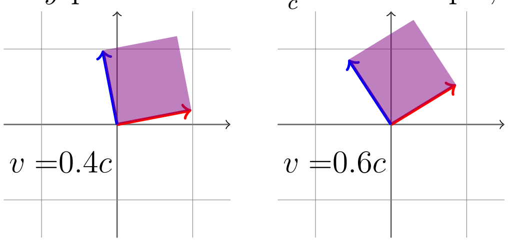
Arthur Parzygnat "The physics and mathematics of special relativity" UConn Math Club 21/2/2018

Lorentz boosting in two perpendicular directions, say x and y, induces a rotation



$$L_y^{-1}L_x^{-1}L_yL_x \approx \begin{bmatrix} \cos\left(\frac{v}{c}\right) & -\sin\left(\frac{v}{c}\right) \\ \sin\left(\frac{v}{c}\right) & \cos\left(\frac{v}{c}\right) \end{bmatrix}$$

in the xy plane for small $\frac{v}{c}$. For example,



Mass increases with velocity via

$$m(v) = \gamma m_0 \approx m_0 + \frac{1}{2}m_0\frac{v^2}{c^2} + \cdots,$$

where m_0 is the rest mass. The energy is

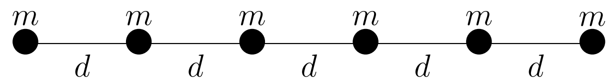
$$E = mc^2 = \gamma m_0 c^2 \approx m_0c^2 + \frac{1}{2}m_0v^2 + \cdots$$

so that mass and energy are equivalent. The two Newtonian laws of conservation of mass and conservation of energy have become one law thanks to special relativity: the conservation of energy.

For example, the mass of the Hydrogen atom is equal to the mass of the proton plus the mass of the electron *minus* the energy/c^2 of the photon released when the electron enters the ground state of the Hydrogen atom.

Each mass m has an associated radius, known as its *Schwarzschild radius*, given by $r = \frac{2mG}{c^2}$. Once another particle is within this radius, there is no escape beyond this region. The spherical shell whose radius is r is called the *horizon*. An object of mass m that is contained within $2r$ is called a *black hole*.

Consider $N+1$ masses m separated by a distance d (significantly greater than the Schwarzschild radius of m) along a straight line.

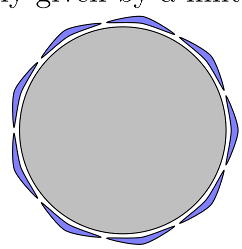


Because gravity is attractive, the net mass, according to special relativity, is

$$(N+1)m - \frac{Gm^2}{dc^2}\left(-N + (N+1)\sum_{i=1}^{N}\frac{1}{i}\right),$$

which can become negative as N increases! Can this happen? If so, what does it mean?

To an observer far away from the black hole's horizon, it takes an infinite amount of time for an object to reach the horizon due to time dilation effects. Length contraction also flattens out the object. Hence, classically, an infinite amount of entropy (information) is allowed in a thin region just outside the horizon. Bekenstein and Hawking showed that the entropy is actually given by a finite quantity



$$S = \frac{k_B A}{4\ell_P^2}$$

where $\ell_P = \sqrt{\frac{\hbar G}{c^3}}$ is the Planck length, approximately 1.6×10^{-35} meters and k_B is Boltzmann's constant. Note that $\lim_{\hbar\to 0} S = \infty$.

Bekenstein also showed that the entropy S and energy E stuffed into a spherical region of radius r is bounded from above by

$$S \leq \frac{2\pi k_B r E}{\hbar c}.$$

This bound is saturated for black holes.

Detailed notes & references can be obtained at:
https://arthur-parzygnat.uconn.edu/talks/