# PHYS 23: Special Relativity
## Connor Neely, Fall 2022

# Chapter 1

# Consequences of Einstein's Postulates

## 1.1 Einstein's Postulates

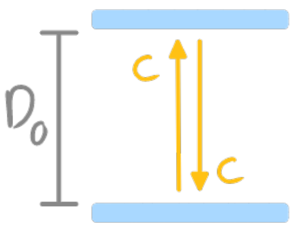
All of special relativity is derived from two postulates.

1. The laws of physics are the same in all inertial (i.e., constant-velocity) frames of reference.
2. The speed of light $c$ is the same in all frames of reference.

Both of these claims have mounds of experimental evidence to back them up. They seem innocent enough at first glance, but together they lead to a rich and wildly intuitive theory of how the universe works.

## 1.2 Time Dilation

Suppose we're on a spaceship equipped with a light clock. This light clock consists of two mirrors spaced a distance $D_0$ apart, with a beam of light bouncing back and forth between them.

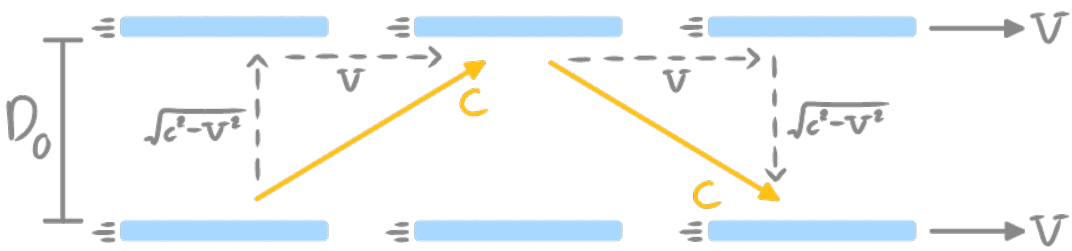



This clock "ticks" every time the light beam hits the bottom mirror. From the equation $d = rt$, we can find the time $\Delta t_s$ between ticks:

$$2D_0 - c\Delta t_s \implies \Delta t_s = \frac{2D_0}{c}.$$

Now imagine another observer watches this spaceship (and the clock on it) fly past Earth at some speed $V$. From the perspective of this observer, the light clock looks a little different.

The light, of course, is still bouncing between the mirrors; however, in order to "keep up" with the mirrors, the light now appears to travel *diagonally* rather than straight up and down. But remember, Einstein's postulates say that light must always travel at $c$, no matter the perspective. No faster, no slower. Just $c$.



Is the time between ticks for the observer on Earth the same as it is for the observer on the spaceship? To determine this, note that the beam of light bounces back and forth between the mirrors with vertical speed $\sqrt{c^2 - V^2}$. Using the same $d = rt$ relationship, we can find the time $\Delta t_m$ between these new ticks:

$$2D_0 = \sqrt{c^2 - V^2} \cdot \Delta t_m \implies \Delta t_m = 2D_0 \cdot \frac{1}{\sqrt{c^2 - V^2}}.$$

We can see right away that this expression doesn't match the one for $\Delta t_s$, but how do the two compare? To find out, let's do a bit of algebraic manipulation.

$$\begin{aligned}\Delta t_m &= 2D_0 \cdot \frac{1}{\sqrt{c^2 - V^2}} \\ &= \frac{2D_0}{c} \cdot \frac{1}{\sqrt{1 - (V/c)^2}} \\ &= \Delta t_s \cdot \frac{1}{\sqrt{1 - (V/c)^2}}\end{aligned}$$

The expression $\sqrt{1 - (V/c)^2}$ is clearly less than one for $V < c$, making the above fraction greater than one. Therefore, $\Delta t_m > \Delta t_s$. Moving light clocks literally take longer to tick than stationary ones—even if they're otherwise identical! Logical consistency requires that all accurate clocks, regardless of type, tick at the same speed, so we can make a generalization: moving clocks tick slowly by a factor of $\sqrt{1 - (V/c)^2}$.

This effect is called *time dilation*—from the Earth's perspective, time itself runs slowly on the moving spaceship. More precisely, if an undilated amount of time $t_0$ passes in our "rest" frame, then we perceive a moving object to have experienced the dilated time

$$t = t_0 \sqrt{1 - \left(\frac{V}{c}\right)^2}.$$

Of course, the observer on the spaceship perceives themself to be the one at rest, so to them it is really our clock that is moving and thus running slowly.

**Summary**

A clock that appears to move at a speed $V$ ticks slowly by a factor of $\sqrt{1 - (V/c)^2}$. That is, if it takes a time $t_0$ for a stationary clock to tick, the time it takes for a moving clock to tick is

$$t = t_0 \sqrt{1 - \left(\frac{V}{c}\right)^2}.$$

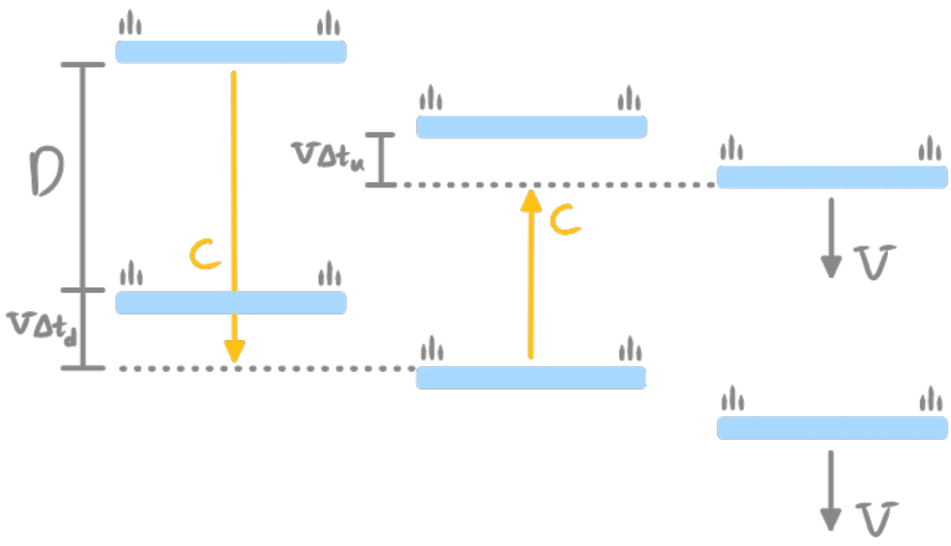
## 1.3 Length Contraction

In our derivation of time dilation, we made an implicit assumption that the mirrors of our light click lay parallel to the direction of motion. What if this were not the case? More specifically, what would happen if your light clock instead moved downward rather than to the right?

Notice how (in the figure on the next page) I named the distance between the mirrors $D$ rather than $D_0$. Given our unexpected conclusion about the passage of time, in the realm of relativity, we can no longer make any seemingly-obvious assumptions about the properties of moving objects. More to the point, we will show that the distance $D$ between the moving mirrors is not $D_0$.

Let $\Delta t_d$ and $\Delta t_u$ be the amount of time the light spends moving downward and upward, respectively. Using the figure above, we get the following $d = rt$ relationships.

$$(D + v\Delta t_d) = c\Delta t_d \implies \Delta t_d = \frac{D}{c - V}$$
$$(D - v\Delta t_u) = c\Delta t_u \implies \Delta t_u = \frac{D}{c + V}$$




Since the total time $\Delta t_m = \Delta t_d + \Delta t_w$,

$$\begin{aligned}\Delta t_m &= \Delta t_d + \Delta t_u \\ &= \frac{D}{c-V} + \frac{D}{c+V} \\ &= \frac{2Dc}{c^2 - V^2} \\ &= \frac{2D}{c} \cdot \frac{1}{1-(V/c)^2}\end{aligned}$$

Lastly, let $\Delta t_s = 2D_0/c$ be the time it takes a stationary light clock to tick. By time dilation,

$$\begin{aligned}\Delta t_m &= \Delta t_s \cdot \frac{1}{\sqrt{1-(V/c)^2}} \\ \frac{2D}{c} \cdot \frac{1}{1-(V/c)^2} &= \frac{2D_0}{c} \cdot \frac{1}{\sqrt{1-(V/c)^2}} \\ D &= D_0\sqrt{1-\left(\frac{V}{c}\right)^2}\end{aligned}$$

This effect is called *length contraction*—from a stationary point of view, space itself becomes compressed for moving objects. Note, however, that this only happens in the direction of motion, which is why length contraction was not a factor in our other setup.

**Summary**

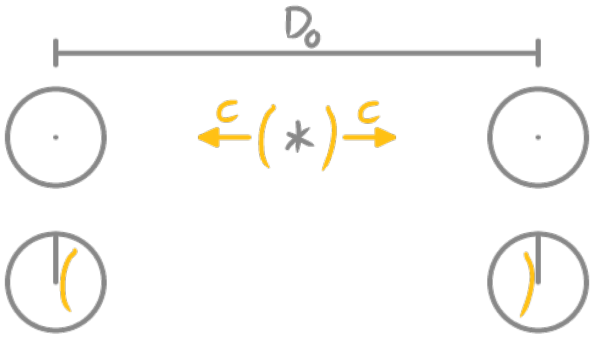
A length that appears to move at a speed $V$ is compressed by a factor of $\sqrt{1-(V/c)^2}$ in the direction of motion. That is, if the length is $D_0$ when it is at rest, it becomes

$$D = D_0\sqrt{1-\left(\frac{V}{c}\right)^2}$$

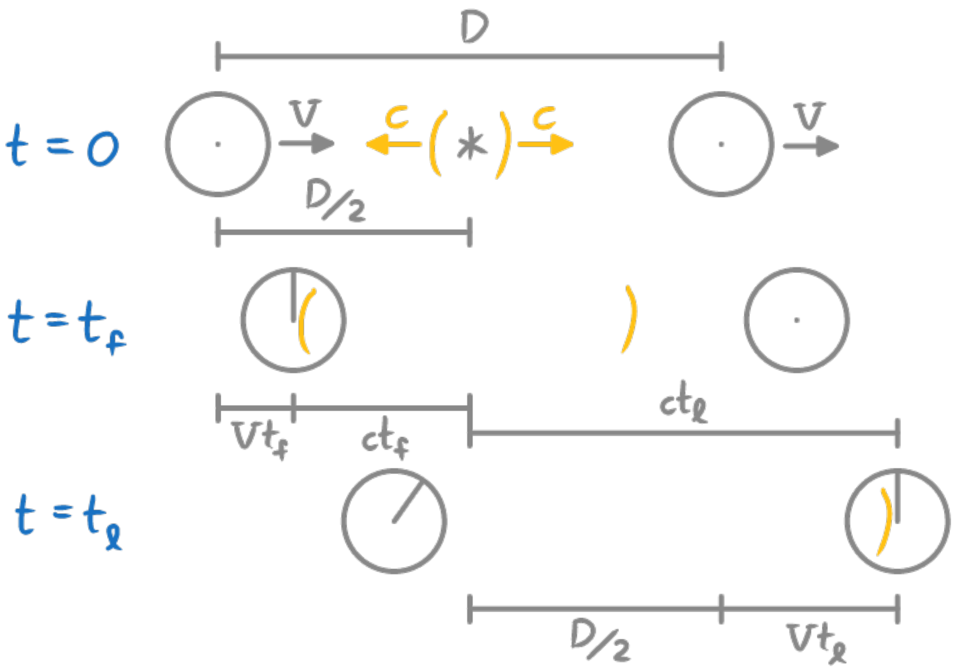
when it moves in a direction parallel to iself.

## 1.4 Relative Simultaneity

Let's imagine, again, that we're on a spaceship, this time equipped with two analog clocks spaced a distance $D_0$ apart. In order to synchronize the clocks, a pulse of light is emitted from their midpoint so that it will reach each clock at the exact same time. The clocks begin running when they detect the pulse.




Clearly, from our perspective, the clocks will begin running simultaneously. But what if we instead perceive the clocks to be moving, one following immediately after the other?




One immediate change we see is the distance between the two clocks. By length contraction, the distance between the moving clocks is $D = \sqrt{1 - (V/c)^2}$. Not only that, but the beam of light will obviously reach the following (left) clock first, and *then* the leading (right) one—these events, although simultaneous in one reference frame, are not simultaneous in another! This is the relativity of simultaneity. Now, let's quantify it.

Our aim, specifically, is to determine the difference $\Delta t$ between the times displayed by the clocks once they're both activated. If the initial pulse is emitted at time $t = 0$, the light reaches the following clock at $t = t_f$ and the leading clock at $t = t_l$. We can use the $d = rt$ relationships in the figure above to solve for these times.

$$\begin{aligned} ct_f + Vt_f &= \frac{D}{2} \\ (c+V)t_f &= \frac{D}{2} \\ t_f &= \frac{D}{2(c+V)} \end{aligned} \qquad\qquad \begin{aligned} ct_l &= \frac{D}{2} + Vt_l \\ (c-V)t_l &= \frac{D}{2} \\ t_l &= \frac{D}{2(c-V)} \end{aligned}$$

From our perspective, the time difference between the two events is $\Delta t_0 = t_l - t_f$.

$$\begin{aligned}\Delta t_0 &= t_l - t_f \\ &= \frac{D}{2(c-V)} - \frac{D}{2(c+V)} \\ &= \frac{DV}{c^2 - V^2} \\ &= \frac{DV}{c^2\left(1-(V/c)^2\right)}\end{aligned}$$

Since $D = D_0\sqrt{1-(V/c)^2}$, this becomes

$$\begin{aligned}&= \frac{D_0 V\sqrt{1-(V/c)^2}}{c^2\left(1-(V/c)^2\right)} \\ &= \frac{D_0 V}{c^2\sqrt{1-(V/c)^2}}\end{aligned}$$

Finally, we must consider that the clocks are moving, and therefore tick slowly due to time dilation. Therefore, the difference between the displayed times is

$$\begin{aligned}\Delta t &= \Delta t_0\sqrt{1-\left(\frac{V}{c}\right)^2} \qquad (1.1)\\ &= \frac{D_0 V\sqrt{1-(V/c)^2}}{c^2\sqrt{1-(V/c)^2}} \qquad (1.2)\\ &= \frac{D_0 V}{c^2} \qquad (1.3)\end{aligned}$$

Note that this quantity is additive rather than multiplicative. For example, if the following clock displays $t = \frac{D_0 V}{c^2}$, then the leading clock will display $t = 0$.

**Summary**

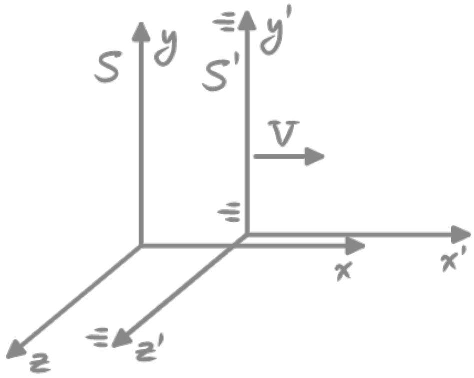
Events that occur simultaneously in one reference frame do not necessarily occur simultaneously in another frame. This is quantified by how, for two clocks moving at a speed $V$, the leading clock lags behind the other by an additive amount of $D_0 V/c^2$, where $D_0$ is the rest length between the clocks.

## 1.5 The Lorentz Transformation

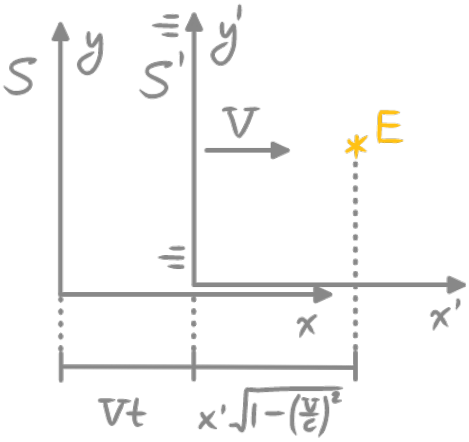
So far we've used two simple postulates to derive three basic rules about the nature of spacetime: time dilation, length contraction, and relative simultaneity. Here, we'll bring all of these rules together into one mathematical framework known as the Lorentz transformation in an effort to formalize everything we've talked about so far. We'll then use it to see how to transform velocities between reference frames.

Let's begin with some terminology. Consider two frames of reference called $S$ and $S'$. By convention, we assume that the $S$ frame is at rest while the $S'$ frame moves in the positive $x$ direction with speed $V$.

Note that, in the figure on the next page, the $x$ and $x'$ axes should coincide—I just separated them slightly to distinguish the different frames. For the remainder of this section, I will also omit the $z$ and $z'$ axes for clarity. (Also, the primes on these coordinates have nothing to do with derivatives. It's simply a way to denote the "transformed" version of each cooordinate.)




Suppose an event $E$ occurs at a point in spacetime with $S$ coordinates $(t, x, y, z)$ and $S'$ coordinates $(t', x', y', z')$. We'll use the three rules we've derived so far to determine the relationship between these sets of coordinates, starting with the spatial coordinates.




In the figure above, we are observing $E$ from within the $S$ frame. We can decompose the $x$-coordinate of $E$ into two parts: the origin-to-origin distance between $S$ and $S$, and the distance from the $S'$ origin to $E$.

We know that the origin-to-origin distance is $Vt$, since the $S'$ frame has been traveling at a speed $V$ for a time $t$. the distance between the $S'$ origin and $E$ appears to be $x'$, but recall that cooordiantes in the moving $S'$ frame are contracted, giving us $x'\sqrt{1-(V/c)^2}$ for our distance. Putting these two together, the $x$-coordinate of $E$ is

$$x = Vt + x'\sqrt{1-\left(\frac{V}{c}\right)^2}.$$

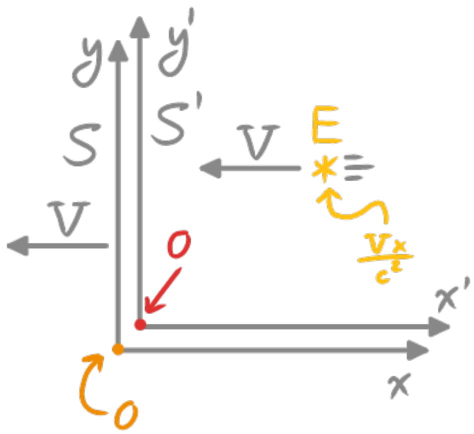
Solving for $x'$ gives the spatial transformation

$$x' = \frac{x - Vt}{\sqrt{1-(V/c)^2}}.$$

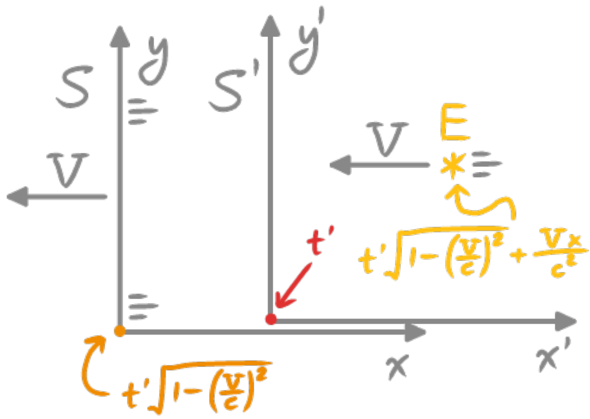
The $y$ and $z$ axes lay perpendicular to the direction of motion, so they aren't impacted by relativistic effects. Their transformations are simply $y' = y$ and $z' = z$.

To transform between the temporal coordinates $t$ and $t'$, consider the same scenario from within the $S'$ frame. The $S$ frame, along with the spatial location of $E$, is moving in the negative $x'$ direction at a speed $V$. Also, the moment the two origins coincide, clocks stationed at each origin are synchronized with each other. These are, in turn, used to synchronize all other clocks in their respective frames.

An observer at the origin of the $S'$ frame notices that a clock at the spatial position of $E$ ticks $Vx/c^2$ ahead of a clock at the origin of the $S$ frame. (Remember—leading clocks lag.)




As the $S$ frame continues to move, the event $E$ occurs at a time $t'$.




Time in the moving $S$ frame is dilated, but we must be careful when we dilate—the phenomenon is only valid for clocks that have been previously synchronized, which in this case are the two clocks at the origins. So, the clock at the origin of the $S$ frame reads $t'\sqrt{1-(V/c)^2}$. Finally, leading clcoks lag, meaning a clock present for the event $E$ would read a time that is $Vx/c^2$ ahead of the origin clock. So, in the $S$ frame, the event $E$ occurs at the time

$$t = t'\sqrt{1-\left(\frac{V}{c}\right)^2} + \frac{Vx}{c^2}.$$

Solving for $t'$ gives the temporal transformation

$$t' = \frac{t - Vx/c^2}{\sqrt{1-(V/c)^2}}.$$

Taken together, all these relationships comprise the Lorentz transformation. This gives another, more formal way to analyze the behavior of spacetime, albeit less immediately intuitive.

The Lorentz transformation tells us how to transform positions in spacetime. Transforming velocities is easy—we can just differentiate the Lorentz transformation equations!

Specifically, we'd like to relate the velocities $\left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right)$ and $\left(\frac{dx'}{dt'}, \frac{dy'}{dt'}, \frac{dz'}{dt'}\right)$ in the $S$ and $S'$ frames, respectively. To do this, we simply take advantage of the chain rule for derivatives:

$$v'_x = \frac{dx'}{dt'} = \frac{dx'/dt}{dt'/dt} = \frac{v_x - V}{1 - Vv_x/c^2}$$

$$v'_y = \frac{dy'}{dt'} = \frac{dy'/dt}{dt'/dt} = \frac{v_y\sqrt{1-(V/c)^2}}{1 - Vv_x/c^2} \qquad v'_z = \frac{dz'}{dt'} = \frac{dz'/dt}{dt'/dt} = \frac{v_z\sqrt{1-(V/c)^2}}{1 - Vv_x/c^2}$$

Take great care when using this velocity transformation! It is assumed that the $S'$ frame moves to the right with respect to the $S$ frame.

**Summary**

Let there be two reference frames $S$ and $S'$. The $S'$ frame moves with speed $V$ in the positive $x$ direction of the $S$ frame, and the frames' origins coincide at $t = t' = 0$. If an event occurs with $S$ coordinates $(t, x, y, z)$, then in the $S'$ frame these coordinates are given by the Lorentz transformation:

$$\begin{aligned} t' &= \gamma \left( t - \frac{Vx}{c^2} \right) \\ x' &= \gamma\,(x - Vt) \\ y' &= y \\ z' &= z \end{aligned}$$

where $\gamma = \frac{1}{\sqrt{1-(V/c)^2}}$ is the Lorentz factor. Differentiating these equations gives the velocity transformation:

$$v'_x = \frac{dx'}{dt'} = \frac{dx'/dt}{dt'/dt} = \frac{(v_x - V)}{1 - Vv_x/c^2}$$

$$v'_y = \frac{dy'}{dt'} = \frac{dy'/dt}{dt'/dt} = \frac{v_y}{\gamma\,(1 - Vv_x/c^2)} \qquad v'_z = \frac{dz'}{dt'} = \frac{dz'/dt}{dt'/dt} = \frac{v_z}{\gamma\,(1 - Vv_x/c^2)}$$

# Chapter 2

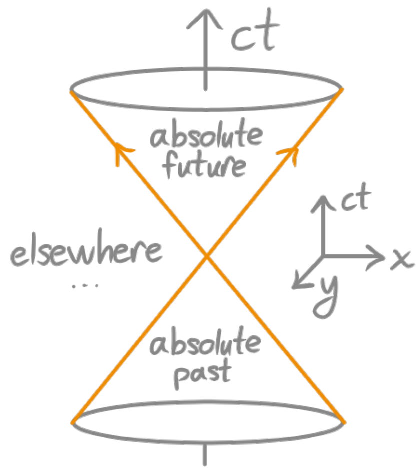
# Conservation Laws in Spacetime

## 2.1 Spacetime

The Lorentz transformation has shown us how space and time transform in such a way that they cannot be considered separately. This leads naturally into the notion of spacetime, the geometry of which is slightly different from the Euclidean geometries we're used to seeing in the world.

A visual will be useful to us in understanding this "pseudo-Euclidean" geometry, which we'll call Minkowski space (after its inventor). Of course, we'd need to work in four dimensions to fully do it justice, but we can communicate the key ideas by working with just one or two spatial dimensions.

Let's start with two. We might model "(1+2)-dimensional" spacetime by putting two spatial dimensions in the $xy$-plane and one temporal dimension on the vertical axis. If we consider the paths of all possible beams of light passing through the origin $(0, 0, 0)$, we get the double-cone below. All massive objects travel more slowly than the speed of light and so are trapped within this "light cone".




Note that from here on out we will refer to time as $ct$ rather than just $t$. There's nothing fundamental going on here, this is just a mathematical convenience—for example, we'll soon find some use in time having the same units as space.

Suppose an event $E$ occurs at the origin. The bottom cone, called the [absolute past] of $E$, contains all events which could possible have influenced $E$ in any way. (Events outside the cone would need their influence to travel faster than the speed of light, which is impossible.) Similarly, the top cone, called the *absolute future* of $E$, contains all events which $E$ could possibly influence.

The above is true for any reference frame. If an event happens in the absolute past in one frame, it also happens in the past for all others. Outside the double-cone, however, all bets are off—here a past event in one frame could be a future event in another. This region is often simply called *elsewhere*.

To make this a little more mathematical, we can draw a rough analogy to something more familiar. In

Euclidean space, rotations leave distances unchanged. If we take two points and rotate them about the origin, the "before" distance is the same as the "after" distance. Therefore, we say that the quantity

$$(\Delta r)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2$$

is invariant under rotations. (Notice how this looks like the equation for a sphere or, in 2D, a circle!)

It's useful to think of spacetime has having a geometry that has werid rules for rotations involving both space and time. Just as rotations in Euclidean space leave circles unaffected, "rotations" in spacetime leave hyperbolas unaffected.[1] As a result, an event occurs on the same hyperbola in all reference frames. This means the quantity $s^2 = -(ct)^2 + x^2 + y^2 + z^2$ is unvariant under the Lorentz transformation!

We can use this fact to define the spacetime interval, a sort of absolute "distance" between events that is independent of reference frames:

$$(\Delta s)^2 \equiv -(c\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2.$$

An interesting characteristic of $(\Delta s)^2$ is that it can be positive, zero, or negative! Each corresponds to a different type of separation between events.

- If $(\Delta s)^2 > 0$ then we have a spacelike interval—there is more separation in space than in time. Here we can always find a reference frame in which the events are simultaneous; the spatial separation in this frame is known as the *proper length* between the events.
- If $(\Delta s)^2 < 0$ then we have a timelike interval—there is more separation in time than in space. Here we can always find a reference frame in which the events occur in the same place; the temporal separation in this frame is known as the *proper time* between the events.
- If $(\Delta s)^2 = 0$ then we have a null interval. This corresponds to events that occur along a light ray.

Events that are separated by a timelike or null interval are in each others' absolute past and futures. For spacelike intervals, the order of events is relative.

**Summary**

Lorentz transformations can be thought of as rotations in a space built upon hyperbolas rather than circles. This hyperbolic geometry gives us the spacetime interval

$$(\Delta s)^2 \equiv -(c\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2,$$

which can be interpreted as a sort of "distance" between events in spacetime. The value of a spacetime interval is not affected by Lorentz transformations, and it gives us information about whether the separation between the events is mostly spatial or temporal (or both).

## 2.2 Momentum

Armed with a better understanding of spacetime, we can now re-examine some fundamental concepts from Newtonian mechanics. To do this, we must define the *four-vector*—the spacetime analog to Euclidean vectors.

It may not be surprising that four-vectors are described by four numbers:

$$A_\mu = (A_0, A_x, A_y, A_z).$$

Four-vectors must transform according to the Lorentz transformation:

$$\begin{aligned}
A_0' &= \gamma\left(A_0 - \frac{VA_x}{c}\right) \\
A_x' &= \gamma\left(A_x - \frac{VA_0}{c}\right) \\
A_y' &= A_y \\
A_z' &= A_z
\end{aligned}$$

[1] A good Desmos visual can be found here: https://www.desmos.com/calculator/r75hsbvmbj

So it follows that, for any four-vector, the quantity

$$-A_0^2 + A_x^2 + A_y^2 + A_z^2$$

is invariant under the Lorentz transformation. (This makes it a *four-scalar*, a number whose value does not change after a Lorentz transformation.)

Now, consider a particle floating around in spacetime. The simplest four-vector describing the particle is the four-position

$$x_\mu = (ct, x, y, z).$$

We can define the four-momentum as

$$p_\mu \equiv m\frac{dx_\mu}{d\tau};$$

where $m$ is the mass of the particle and $\tau$ is its proper time—that is, the time read by a clock moving alongside the particle. The quantity $dx_\mu/d\tau$ is called the four-velocity, and we differentiate it with respect to $\tau$ because it is a four-scalar. (Otherwise, it would give us something different in each reference frame!)

We can expand our expression a bit:

$$p_\mu = m\left(c\frac{dt}{d\tau}, \frac{dx}{d\tau}, \frac{dy}{d\tau}, \frac{dz}{d\tau}\right).$$

Due to time dilation,

$$\frac{dt}{d\tau} = \frac{1}{\sqrt{1-(v/c)^2}},$$

where $v$ is the speed at which the particle is observed to move. We'll call this quantity $\gamma$. (Notice how this $\gamma$ is subtly different from the one we defined previously, in which $V$ was a relative speed between reference frames; we can reconcile this by interpreting $v$ as the speed between our frame and the particle's rest frame.)

Now, by the chain rule,

$$\frac{dx}{d\tau} = \frac{dx}{dt}\frac{dt}{d\tau} = \gamma v_x,$$

where $v_x$ is the $x$-component of the particle's observed velocity. Therefore,

$$p_\mu = (\gamma mc, \gamma mv_x, \gamma mv_y, \gamma mv_z) = (\gamma mc, \gamma m\mathbf{v}).$$

The spatial component of this four-vector is the relativistic momentum! That is,

$$\mathbf{p} = \gamma m\mathbf{v}.$$

Notice that at low speeds $\gamma \approx 1$, and we recover the classical $\mathbf{p} = m\mathbf{v}$.

It has been shown, through, experiment that this momentum is conserved in collisions. It is also easy to show, using the Lorentz transformation, that if momentum is conserved in one frame then it is conserved in all other frames. So this quantity has the universal importance we seek!

**Summary**

Four-vectors are objects that transform according to the Lorentz transformation. If a four-vector is specified by the components $A_\mu = (A_0, A_x, A_y, A_z)$, then the quantity

$$-A_0^2 + A_x^2 + A_y^2 + A_z^2$$

is invariant under the Lorentz transformation.

The most basic four-vector is the four-position $x_\mu = (ct, x, y, z)$. Another is the four-momentum, defined by

$$p_\mu \equiv \frac{dx_\mu}{d\tau} = (\gamma mc, \gamma m\mathbf{v}).$$

The spatial $\mathbf{p} = \gamma m\mathbf{v}$ is the relativistic momentum, and it is conserved in collisions.

## 2.3 Energy

The relativistic momentum $\mathbf{p}$ that we found in the previous section only encompasses the spatial components of the four-momentum. What about the temporal component $p_0$? What does it represent?

Notice that $p_0$ is a scalar that is conserved. The obvious candidate is energy. The unit discrepancy is no problem—we can just multiply by $c$ to get units of energy, just like wedid when formulating spacetime.

To see if this makes sense, consider the series expansion of $p_0 c$:

$$\begin{aligned} p_0 c &= \frac{mc^2}{\sqrt{1-(v/c)^2}} \\ &= mc^2\left(1+\frac{1}{2}\left(\frac{v}{c}\right)^2+\cdots\right) \\ &= mc^2+\frac{1}{2}mv^2+\cdots \end{aligned}$$

We can immediately see that the object's energy increases as it moves faster, which checks out. This is just the idea of kinetic energy. Specifically, the kinetic energy of the object is given by all the velocity-dependent terms in $p_0 c$; that is,

$$K = \gamma mc^2 - mc^2 = (\gamma - 1)mc^2.$$

Also notice that, in the nonrelativistic limit $v \ll c$, this kinetic energy is simply $\frac{1}{2}mv^2$—which is exactly what we'd expect from Newtonian mechanics!

There's this other term, though, that's completely independent of the object's speed, or even its position in space. This quantity,

$$E_0 = mc^2,$$

is called the *mass energy* (or *rest energy*) of the object. It is the energy that the object has just by existing!

So $E = p_0 c$ is the energy of an object. We can now write the four-momentum as

$$p_\mu = (\gamma mc^2, \gamma m\mathbf{v}) = \left(\frac{E}{c}, \mathbf{p}\right).$$

The invariant associated with this four-vector is

$$-(\gamma mc^2)^2 + (\gamma mv)^2 = \left(\frac{E}{c}\right)^2 + p^2.$$

Manipulating the left-hand side gives

$$-(mc)^2 = \left(\frac{E}{c}\right)^2 + p^2$$

or, alternatively,

$$E^2 = (pc)^2 + (mc^2)^2.$$

This is another very important Lorentz invariant, this time showing that energy and momentum are fundamentally related! It's also true for all objects, including massless ones. Thus the energy and momentum of a photon are related by

$$E = pc.$$

As one final tidbit, we'll give a particularly interesting consequence of these conservation laws.

Consider the decay of a particle of mass $M$ into two identical particles, each of mass $m$. By conservation of momentum, the resulting particles fly off in opposite directions with the same speed $v$; by conservation of energy, though,

$$Mc^2 = 2\gamma mc^2 \implies M = \gamma \cdot 2m.$$

Since $\gamma \neq 1$, mass has not been conserved! Specifically, some of the original mass energy has been converted into kinetic energy, so $M > 2m$. Yet another wildly unintuitive result from special relativity, that conservation of mass does not exist.

**Summary**

The temporal component of an objet's four-momentum represents its total energy,

$$E = \gamma mc^2.$$

This energy has two components: the mass energy $E_0 = mc^2$ and the kinetic energy $K = (\gamma - 1)mc^2$. This leads to a new Lorentz invariant,

$$E^2 = (pc)^2 + (mc^2)^2,$$

which applies to all particles, including massless ones.

Together, conservation of energy and momentum leads to the surprising fact that mass is not necessarily conserved in collisions.