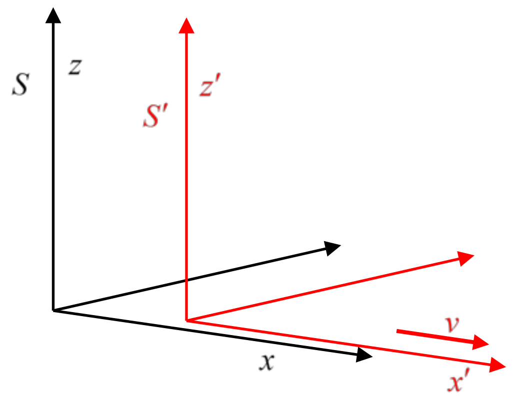
3/28/18

15. Special Relativity

Introduction

In the Modern Physics lectures just reviewed, we found the Lorentz transformations between two parallel frames S, S' with S' moving at constant speed v along the common x axis relative to S, both taking the zero of time to be when the origins coincide.



The result was that the coordinates (x, y, z, t) of an event in S in terms of the coordinates (x', y', z', t') of the same event in frame S' are:

$$x = \frac{x' + vt'}{\sqrt{1 - v^2/c^2}},$$
$$y = y',$$
$$z = z',$$
$$t = \frac{t' + vx'/c^2}{\sqrt{1 - v^2/c^2}}.$$

It's standard notation to write $1/\sqrt{1-(v/c)^2} = \gamma$, very common to write $(\vec{v}/c) = \vec{\beta}$, and often to take $c = 1$.

Standard Relativistic Notation

An "event" has four coordinates: position in three-dimensional space, plus time. That is, it's just a point in *four*-dimensional space, a.k.a. space time. The standard notation is

$$(ct, x, y, z) \equiv (x^0, x^1, x^2, x^3).$$

Matrix Form of Lorentz Transformation

The Lorentz transformation equations from frame S', moving at v in the x-direction relative to S, can be written in matrix form:

$$\begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} \gamma & v\gamma & 0 & 0 \\ v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^{0'} \\ x^{1'} \\ x^{2'} \\ x^{3'} \end{pmatrix},$$

The standard notation for this equation is: $x^\alpha = \Lambda^\alpha{}_{\beta'} x^{\beta'}$.

(*Hint*: to check the sign, take the nonrelativistic limit, $x^1 = v\gamma x^{0'} + \gamma x^{1'} \cong vt + x'$.)
Points to note: we have some indices up, some down. (This is really important in *general* relativity, less so here. There are various notations: contravariant and covariant, vectors and dual vectors, vectors and forms, etc. We'll just call them up and down indices.)

The $\Lambda^{\alpha'}{}_{\beta}$ are just the elements of the matrix written above. "Lambda" is used as *L* for Lorentz.

Four-Vectors: the Metric Tensor, Magnitude of a Vector

Definition of a four-vector: a set of four numbers in any inertial frame, $\vec{A} \xrightarrow[S]{} \{A^\alpha\}$, that transform from one frame to another like the coordinates of an event $\{x^\alpha\}$: that is, $A^\alpha = \Lambda^\alpha{}_{\beta'} A^{\beta'}$. Also called a *contravariant* vector.

Exercise: check that under a Lorentz transformation, $\vec{x}^2 - c^2t^2$ is invariant. In fact, this quantity is called the *magnitude* of the four-vector. Unlike most magnitudes, this one can be *negative*, or zero for a nonzero vector. To write it in terms of the new notation, we have to "square" the vector x^μ, but also include the information that the time and space contributions have opposite signs.

The way this is done is to introduce a *metric tensor*,

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

(Some authors, including Jackson, have an overall minus sign! Watch out if you solve a problem using multiple sources…)
With this, the position vector (x^0, x^1, x^2, x^3) can be converted to one with *down* indices by:

$$x_\mu = g_{\mu\nu} x^\nu,$$

and we see this gives $(x_0, x_1, x_2, x_3) = (-x^0, x^1, x^2, x^3)$.
(The index can be lowered with $g^{\mu\nu}$, which is the inverse of $g_{\mu\nu}$, except that in our special relativity case, they're the same.)

Finally, the *magnitude* of the vector is written

$$x^\mu x_\mu = g_{\mu\nu} x^\mu x^\nu,$$

giving the result $\vec{x}^2 - c^2t^2$.

The Interval

One more piece of jargon: the *interval*. Since the Lorentz transformation is *linear*, and true for arbitrary space time points, the four-vector difference Δx^μ between two space time points clearly also transforms as a four vector, its magnitude is

$$ds^2 = \Delta x^\mu \Delta x_\mu,$$

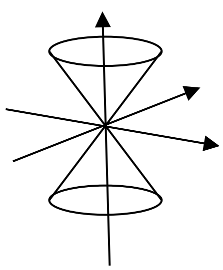
and this "square", the so-called magnitude, rather than the vector itself, is called the *interval*.
It can be positive, negative, or zero.

Spacelike, Timelike, Lightlike

Two events (ct_1, x_1, y_1, z_1), (ct_2, x_2, y_2, z_2) are said to be spacelike separated if the interval between them $-c^2(t_2 - t_1)^2 + (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 = \Delta s^2 > 0$. It is important to note that spacelike separation in one inertial frame of reference means spacelike separation in *all* inertial frames, since the magnitude is invariant under Lorentz transformation.

Similarly, timelike separation is $\Delta s^2 < 0$, lightlike separation $\Delta s^2 = 0$.

Points lightlike separated from the origin are said to be on the *light cone*, which is really two cones having vertices at the origin, the forward (in time) light cone, and the backward light cone. A light signal sent from the origin (meaning $ct = x = 0$) could trigger an event (a bomb?) anywhere on the forward light cone, a light signal from anywhere on the backward light cone could trigger an event at the origin.



An event at the origin *cannot* be the cause of another event which is outside the forward light cone.

Worldlines

As a particle moves through space time, the path traced is termed the *worldline*. Since particles travel at less than the speed of light, the world line lies within the forward light cone. A particle at rest has a worldline along the axis of the cone: in other words, the time axis. A photon has a world line on the surface of the light cone.

Relativistic Addition of Velocities

As stated above, $(c\Delta t, \Delta x, \Delta y, \Delta z)$ transforms just as (ct, x, y, z) does, we'll write the transformation

$$\Delta t = \gamma\left(\Delta t' + v\Delta x'/c^2\right)$$
$$\Delta x = \gamma\left(v\Delta t' + \Delta x'\right)$$
$$\Delta y = \Delta y', \quad \Delta z = \Delta z'.$$

From these equations in the limit of small displacements, $\Delta x/\Delta t$ gives the *addition of velocities* formulas

$$\frac{\Delta x}{\Delta t} = u_x = \frac{\Delta x' + v\Delta t'}{\Delta t' + v\Delta x'/c^2} = \frac{u'_x + v}{1 + u'_x v/c^2} \quad \text{and} \quad u_y = \frac{u'_y}{\gamma\left(1 + u'_x v/c^2\right)}.$$

(Recall the primed frame is moving at v in the positive x -direction relative to the unprimed frame.)
A standard exercise is to consider a space station moving at v in the x -direction relative to an observer sending a rocket ship forward at u relative to the ship. What is the velocity of the rocket ship relative to the "stationary" observer? The answer is:

$$"u+v" = \frac{u+v}{1+uv/c^2}$$

Rotations and Boosts, Rapidity

Notice now that the 4 x 4 Lorentz matrix can also represent ordinary rotations in the three-dimensional space:

$$\begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^{0'} \\ x^{1'} \\ x^{2'} \\ x^{3'} \end{pmatrix},$$

and manifestly $x^\mu x_\mu$ is invariant. Any three-dimensional rotation can be represented by the lower-right 3 X 3 minor.

In fact, the Lorentz transformation to a moving frame—called a "*boost*"—can be formulated in a strikingly similar way, in terms of a variable much favored by high energy physicists, the *rapidity* ψ, defined by

$$\frac{v}{c} = \beta = \tanh\psi, \quad \gamma = 1/\sqrt{1-\beta^2} = \cosh\psi.$$

Rapidity proves to be a very useful parameter, because for one thing

$$\tanh\left(\psi + \psi'\right) = \frac{\tanh\psi + \tanh\psi'}{1 + \tanh\psi\,\tanh\psi'}$$

which is *exactly* the Lorentz addition formula for velocities! (Recall $"u+v"=\frac{u+v}{1+uv/c^2}$.) This means that in successive boosts *you just add the rapidities*.

The Lorentz transformation for boosting from rest to a rapidity ψ along the x-axis is:

$$\begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} \cosh\psi & \sinh\psi & 0 & 0 \\ \sinh\psi & \cosh\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^{0'} \\ x^{1'} \\ x^{2'} \\ x^{3'} \end{pmatrix}.$$

That is, a particle at rest in the moving (boosted) frame is moving with rapidity ψ in the original frame.

Notice the similarity to the three-dimensional rotation! The sign difference ensures the transformations are unitary. Some authors (for instance, Zangwill) take time to be an *imaginary* variable, so the rotation and boost transformations look identical, but we'll stick with the more common practice. (There are in fact deep mathematical differences between rotations and boosts, as we'll see.)

*Lorentz Transformation for Arbitrary Direction

For a boost of *v* in the *x*-direction the coordinates *in the boosted frame* are:

$$\begin{pmatrix} x^{0'} \\ x^{1'} \\ x^{2'} \\ x^{3'} \end{pmatrix} = \begin{pmatrix} \gamma & -v\gamma & 0 & 0 \\ -v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$$

For a boost of *v* in the direction $\begin{pmatrix} \hat{v}_x & \hat{v}_y & \hat{v}_z \end{pmatrix}$ (the hats meaning components of a unit vector $\hat{\mathbf{v}}$) the corresponding matrix is:

$$\begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & -v_x\gamma & -v_y\gamma & -v_z\gamma \\ -v_x\gamma & 1+(\gamma-1)\frac{v_x^2}{v^2} & (\gamma-1)\frac{v_x v_y}{v^2} & (\gamma-1)\frac{v_x v_z}{v^2} \\ -v_y\gamma & (\gamma-1)\frac{v_y v_x}{v^2} & 1+(\gamma-1)\frac{v_y^2}{v^2} & (\gamma-1)\frac{v_y v_z}{v^2} \\ -v_z\gamma & (\gamma-1)\frac{v_z v_x}{v^2} & (\gamma-1)\frac{v_z v_y}{v^2} & 1+(\gamma-1)\frac{v_z^2}{v^2} \end{bmatrix} \begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix},$$

Notice first that this does give the right answer for the boost along the *x*-axis.

Our strategy for boosting in an arbitrary direction is to reorient the system so that that direction becomes the *x*-axis, apply our known boost, then rotate it back.
To see how this works, we write the above matrix in terms of blocks, as follows:

$$\begin{pmatrix} \gamma & -\gamma \mathbf{v}^T \\ -\gamma \mathbf{v} & I + (\gamma - 1)\hat{\mathbf{v}}\hat{\mathbf{v}}^T \end{pmatrix}.$$

In this same block notation, a three-dimensional rotation has the form

$$\begin{pmatrix} 1 & 0 \\ 0 & \mathbf{R} \end{pmatrix}$$

and its inverse is $\begin{pmatrix} 1 & 0 \\ 0 & \mathbf{R}^T \end{pmatrix}$. If we choose $\mathbf{R}$ such that $\mathbf{Rv}$ points along the x-axis, then

$$\begin{pmatrix} 1 & 0 \\ 0 & \mathbf{R} \end{pmatrix}\begin{pmatrix} \gamma & -\gamma \mathbf{v}^T \\ -\gamma \mathbf{v} & I + (\gamma - 1)\hat{\mathbf{v}}\hat{\mathbf{v}}^T \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \mathbf{R}^T \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma \mathbf{v}^T\mathbf{R}^T \\ -\gamma \mathbf{Rv} & I + (\gamma - 1)\mathbf{R}\hat{\mathbf{v}}\hat{\mathbf{v}}^T\mathbf{R}^T \end{pmatrix}$$

where $\mathbf{Rv} = \begin{pmatrix} v \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{v}^T\mathbf{R}^T = \begin{pmatrix} v & 0 & 0 \end{pmatrix}$. The reader can easily check that this is the form we found for a boost along the *x*-axis— so reversing the process gives us the general case. (And, we don't have to find the explicit form of $\mathbf{R}$!)

A Bit of Group Theory

The Lorentz boosts along the *x*-axis formed an Abelian (commutative) group, just as the set of rotations in a plane do. The rotations in a plane are a subgroup of the group of three-dimensional rotations, which is of course non-abelian. What about the set of all Lorentz boosts? It turns out that this is *not* a group. A product of two Lorentz boosts in different directions is not just a Lorentz boost in some combined direction, it also has some rotation. The Lorentz group is the group of boosts *plus* rotations.

Proper Time and Four-Velocity

Consider a spaceship going from one planet to another, the planets might have quite different velocities, so the distance covered by the ship will be different in the two planet rest frames. One thing that won't be different is the time elapsed as measured by the crew of the spaceship. This is called the *proper time* of the spaceship, the clock is always with the ship.

An increment of proper time is denoted by $d\tau$.

If the spaceship moves Δx^μ in time $d\tau$, this incremental displacement transforms as a Lorentz four-vector. Therefore, so does

$$U^\mu = \frac{dx^\mu}{d\tau}.$$

The four-vector U^μ is called the *four-velocity*. In the nonrelativistic limit it becomes (c, v^i), the spatial part just the ordinary velocity, and $\tau \to t,\ x^0 = ct$.

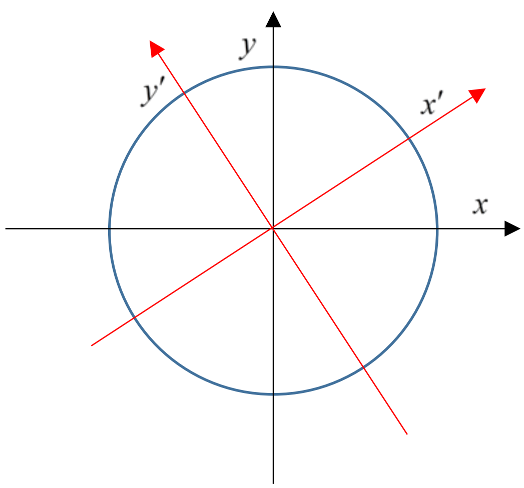
Now $U^\mu U_\mu = \frac{dx^\mu dx_\mu}{(d\tau)^2}$, but $dx^\mu dx_\mu$ is just the interval, which has the same value in all frames, including the frame in the ship, where it is $-(d\tau)^2$, so it follows that

$$U^\mu U_\mu = -c^2.$$

So in the rest frame, where the incremental movement along the world line dx^μ is purely in the time direction, and is just $d\tau$, the four-velocity is $(c,0,0,0)$.
In general, it's $(\gamma c, \gamma v^1, \gamma v^2, \gamma v^3)$. (In relativity papers and books, these formulas usually appear with $c = 1$.)

New Frame Axes and Scales in the Old Frame

For ordinary rotations in a plane, circles centered at the origin are invariant, and in particular the unit circle cuts the axes, old and new, at the point one, so it connects—rather trivially—the two *scales*, in the new frame and the old frame.



In order to discuss length contraction and time dilation, it is essential to have a way to connect axes and scales in different frames.

For the Lorentz transformation, instead of the simple invariant circles $x^2 + y^2 = R^2$, we evidently have invariant *hyperbolae*, $-c^2t^2 + x^2 = a^2$, or $-c^2t^2 + x^2 = -b^2$ (a, b real.)

The circles came from a rotation matrix

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

and $\cos^2\theta + \sin^2\theta = 1$.

The invariant hyperbolae come analogously from

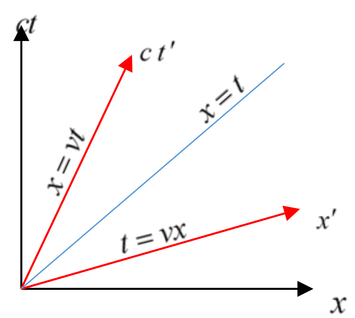
$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \cosh\psi & -\sinh\psi \\ -\sinh\psi & \cosh\psi \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} = \mathbf{\Lambda}(\psi) \begin{pmatrix} ct \\ x \end{pmatrix}.$$

and $\cosh^2\psi - \sinh^2\psi = 1$.

Recall that $\tanh\psi = 0$ for $\psi = 0$, and $\tanh\psi \to \pm 1$ as $\psi \to \pm\infty$.

First, the lines $ct = \pm x$ must go to $ct' = \pm x'$, they constitute the two-dimensional version of the light cone.

This light cone invariance only works because there is one sign change in $\Lambda(\psi)$ compared with $R(\theta)$, and that sign change means that the ct', x' axes turn in *opposite* directions from the t, x axes, in contrast to the ordinary rotation, so going to larger and larger boosts, the axes close like scissors around the line $x = t$, never reaching it, of course.



This is easy to see from the equations: the t-axis is the line $x = t$ $x = 0$, the t' axis is the line $x' = 0$, or $x = vt$.

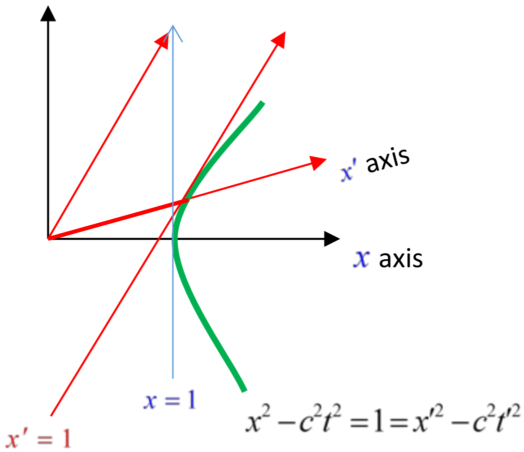
Put another way: the t-axis is the "world line", meaning the path in space time, of an object at rest at the origin in the original frame, the t'-axis is the world line of an object at rest at the origin of the primed frame.

The x' axis is the line $t' = 0$, so $t = vx$ in the original frame.

So the primed frame axes are the original axes turned through opposite angles $\pm\theta$, $\tan\theta = v = \tanh\psi$. This means that for small speeds, θ, v, ψ are close, but as ψ goes to infinity, θ just approaches 45°.

These diagrams are often called *Minkowski diagrams*—first drawn by Minkowski a few years after Einstein published his special relativity paper.

Finding Length and Time Scales in a New Frame: Invariant Hyperbolae



We've now seen how the axes move, but we haven't tracked what happens to the *calibration*—the scale on the axes. The way to do that is to use an invariant hyperbola, for example $-t^2 + x^2 = 1 = -t'^2 + x'^2$.
This hyperbola cuts the x-axis at $x = 1$, and the x' axis at $x' = 1$. Note that $x' = 1$ is the tangent line to the unit hyperbola $x'^2 - t'^2 = 1$, it's the minimum possible value of x' on that hyperbola.

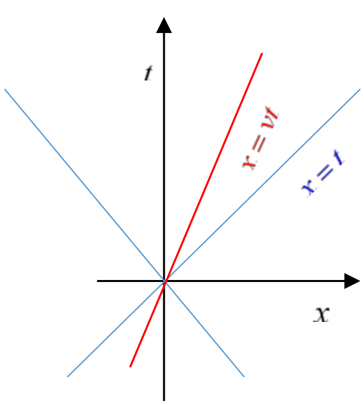
Lorentz Contraction

Notice that from the diagram, in the (x, t) plane the point $x' = 1$ is further from the origin than the point $x = 1$. Does this mean that a rod of unit length at rest in the primed frame (say, stretching from $x' = 0$ to $x' = 1$ will appear longer than unity in the (x, t) frame?

Presumably not—that would be the *opposite* of Lorentz contraction.
So what's going on? The essential point is that we're looking at the x -positions of the ends of the rod at *different times* t. To measure the length of a moving rod, we obviously need to find the x-values of the end points *at the same time* t.

World Lines

As we mentioned earlier, the world line of a particle (or of a small part of a solid object) is its *path in four-dimensional spacetime*.



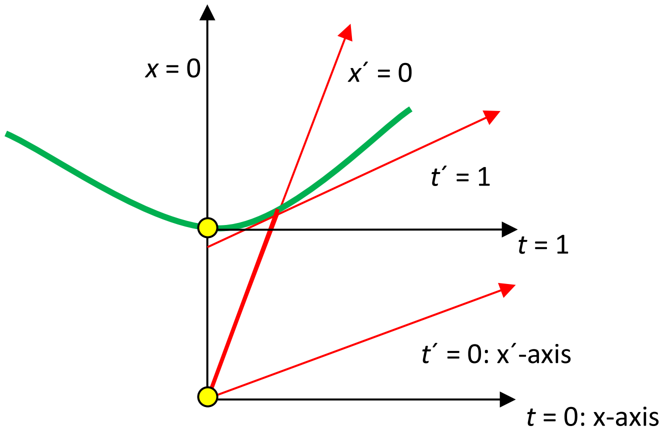
Here are some sample world lines in a two dimensional subspace. First, the light cone sections are world lines of photons, traveling at $c = 1$. A particle moving at constant velocity $x = vt$ is shown. The world line must be *steeper* than the light cone, since $v < c$. A particle *at rest* has a world line parallel to the t axis, so the t axis itself is the world line of a particle at rest at the origin.
Now, back to measuring the moving rod. We need to plot the world lines of the two ends, and find how far apart they are at, say, $t = 0$. Look back at the original diagram. The world line of the left end is just that of the primed origin, that is, it's the t' axis, $x' = 0$. The other end is moving at the same velocity, so its world line has the same slope: it's $x' = 1$, the tangent line to the unit hyperbola $x'^2 - t'^2 = 1$,as mentioned earlier.
Plotting the two world lines, we can see that their simultaneous intercepts on $t = 0$ are at *points less than one unit apart*.

Exercise: check your understanding by using similar arguments to show that a rod of unit length at rest in the original unprimed frame will have length measured as less than one in the primed frame.

Time Dilation

We've just seen how an invariant space-like hyperbola can explain how each observer can see the other as Lorentz contracted. A *time-like* invariant hyperbola can show us that each sees the other's clock as running slow.

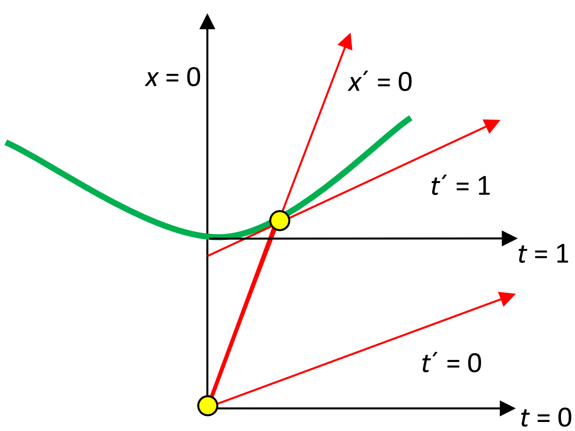
Here is the invariant hyperbola is $-t^2 + x^2 = -1 = -t'^2 + x'^2$:



Two yellow circles represent flashes one second apart at origin in *unprimed* frame

The red parallel lines here are the lines $t' = 0$ and $t' = 1$, both lines of simultaneity in the primed frame. Suppose first that a clock in the unprimed frame flashes once a second. The initial flash is seen by both observers to be at their common origin, $t = t' = 0$. The second flash, at $t = 1$, is clearly at $t' > 1$--the clock is running slow in the primed frame.

Now suppose a clock in the *primed* frame is flashing once a second. As before, the initial flash is at the common origin. The next flash, at $t' = 1, x' = 0$ is clearly at $t > 1$:



Two yellow circles represent flashes one second apart at origin in *primed* frame

Look at the invariant hyperbolae *and scale markings* on [this animation](#)!

Four-Acceleration

The *acceleration four vector* is defined as $\vec{a} = d\vec{U} / d\tau$. Notice that since the four velocity has constant magnitude $\vec{U} \cdot \vec{U} = -1$, *the four acceleration is always orthogonal to the four velocity*: $\vec{U} \cdot d\vec{U} / d\tau = 0$, and so *in the frame of the moving object the four acceleration has only spatial components*.

The acceleration four vector can also be written $a^{\alpha} = \dfrac{dU^{\alpha}}{d\tau} = \dfrac{dU^{\alpha}}{dx^{\beta}} \dfrac{dx^{\beta}}{d\tau} = U^{\beta} \partial_{\beta} U^{\alpha}$.