PHY141 Lectures 24,25 notes

Alice Quillen

December 1, 2022

Contents

1 Special Relativity!

1.1 Space time

The development of special relativity involved redefining what is meant by **space time**.

There are two ways to think about the argument that leads to Lorenz transformations. One, that by Einstein, is based on the invariance of the speed of light. Alternatively one can think about a theoretical maximal speed of information transmission which must be invariant, and with speed coinciding with the speed of light in vacuum.

Einstein based his theory of special relativity on two fundamental postulates. First, all physical laws are the same for all inertial frames of reference, regardless of their relative state of motion; and second, the speed of light in free space is the same in all inertial frames of reference.

Let's consider two points in space time defined by position and time x_1, y_1, z_1, t_1 and x_2, y_2, z_2, t_2. A photon leaves the first position and travels to the second one. The distance between the two positions is

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

This must be equal to $d = c(t_2 - t_1)$ if the photon travels at c the speed of light. This gives

$$(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 - c^2(t_2 - t_1) = 0$$

If the speed of light is the same in all reference frames then this condition must hold in all reference frames.

For a small travel time or distance we can compute the interval

$$ds^2 = c^2 dt^2 - (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$$

We can consider a transfer of reference frame to a new coordinate system giving ds'^2. Suppose that $ds = ads'$. What can a depend on? We would like space to be isotropic. So a cannot depend on the actual positions of the events as that would make space-time inhomogeneous. The function a could depend on the relative velocity between the frames. If it depended on the velocity direction, then space-time would not be isotropic. What if it depends on the velocity amplitude? Then we wind up with a contradiction if we do three transformations in a row. We would like transformation and inverse transformations to be similar. This pushes $a = 1$.

1.2 The Lorenz transformation

To simplify a search for a transformation we restrict our transformation to x, t only. We can shift the origins for both coordinate systems so that x_1, t_1 is the origin for the first coordinate system and x'_1, t'_1 is the origin for the second one. We set $x_2 = x$, $t_2 = t$ and $x'_2 = x'$, $t'_2 = t'$. We search for a transformation that preserves

$$c^2 t^2 - x^2 = c^2 t'^2 - x'^2$$

Assume a linear transformation

$$\begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x' \\ t' \end{pmatrix} = \begin{pmatrix} Ax' + Bt' \\ Cx' + Dt' \end{pmatrix}$$

$$c^2 t'^2 - x'^2 = c^2 (Cx' + Dt')^2 - (Ax' + Bt')^2$$

We find that

$$\begin{aligned} A^2 - c^2 C^2 &= 1 \\ 2(c^2 CD - AB) &= 0 \\ c^2 D^2 - B^2 &= 1 \end{aligned}$$

It helps to define $C' = C/c$ and $D' = D/c$ giving

$$\begin{aligned} A^2 - C'^2 &= 1 \\ C'D' &= AB \\ D'^2 - B^2 &= 1 \end{aligned}$$

We have 4 unknowns and 3 constraints. Regular sines and cosines obey

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

but hyperbolic sines and cosines obey

$$\cosh^2\theta - \sinh^2\theta = 1$$

A general solution is

$$\begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} \cosh\Psi & c\sinh\Psi \\ \sinh\Psi & c\cosh\Psi \end{pmatrix} \begin{pmatrix} x' \\ t' \end{pmatrix}$$

Giving

$$x = \cosh\Psi x' + \sinh\Psi ct'$$
$$t = \sinh\Psi x' + \cosh\Psi ct'$$

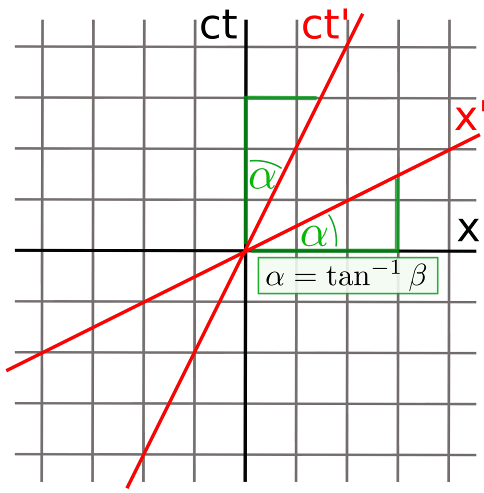

Figure 1: The x' and t' axes in a coordinate system with axes x, t are related by the Lorenz transformation with $\beta = v/c$. On this 2d plot, we let $x = 1$ and $ct = 1$ span the same distance on the plot.

Consider the origin of the prime frame at $x' = 0$. This transforms to $x = \sinh\Psi ct'$ and $t = \cosh\Psi ct'$. We take the ratio

$$\frac{x}{ct} = \tanh\Psi$$

We associate $v = x/t$ as the relative motion of the two frames.

$$\beta = \frac{v}{c} = \tanh\Psi$$

The transformation becomes

$$\begin{aligned} x &= \gamma x' + \gamma\beta c t' \\ t &= \gamma\beta c^{-1} x' + \gamma t' \end{aligned} \tag{1}$$

with

$$\gamma \equiv (1-\beta^2)^{-\frac{1}{2}}.$$

The inverse transform

$$\begin{aligned} x' &= \gamma x - \gamma\beta c t \\ t' &= -\gamma\beta c^{-1} x + \gamma t \end{aligned} \tag{2}$$

We can also write the transformation as

$$\begin{pmatrix} x \\ ct \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\beta \\ \gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} x' \\ ct' \end{pmatrix} \tag{3}$$

and the inverse transform

$$\begin{pmatrix} x' \\ ct' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} x \\ ct \end{pmatrix} \tag{4}$$

We verify that the transform applied to the inverse transform gives the identity transformation

$$\begin{aligned} \gamma \begin{pmatrix} 1 & \beta \\ \beta & 1 \end{pmatrix} \gamma \begin{pmatrix} 1 & -\beta \\ -\beta & 1 \end{pmatrix} &= \gamma^2 \begin{pmatrix} 1+\beta^2 & \beta-\beta \\ \beta-\beta & 1+\beta^2 \end{pmatrix} \\ &= \gamma^2(1+\beta^2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

The x' axis is all points with $t' = 0$. Looking at equation 1 this is all points with $\beta c^{-1} x - t = 0$. This has $ct = \beta x$ and which has a slope of β in Figure 1.

The t' axis is all points with $x' = 0$. Looking at equation 1 this is a line with $x - \beta ct = 0$. Putting ct as a y axis $ct = x/\beta$ and the line's slope in Figure 1 is $1/\beta$.

What's the inverse transform look like? Similar to Figure 1 except the lines have negative slope and so go through quadrant's 2 and 4 instead of 1 and 3. See Figure 2.

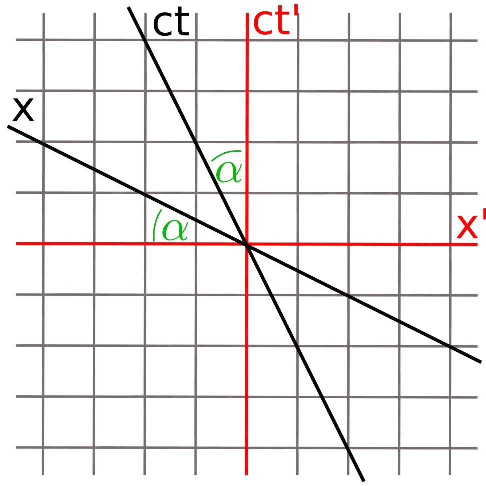

Figure 2: The inverse transform of that in Figure 1. .

1.3 Minkowski space

Space time consists of points $(t, \mathbf{x}) = (t, x, y, z)$. We can define a 4-vector

$$\mathbf{A} = (A_t, A_x, A_y, A_z).$$

A notion of length can be defined with a modification to the dot product.

$$\mathbf{A} \cdot \mathbf{A} = c^2 A_t^2 - A_x^2 - A_y^2 - A_z^2. \tag{5}$$

Lorenz transformations preserve this dot product. In other words if we transform $\mathbf{A}$ using a Lorentz transformation to $\mathbf{A}'$, then $\mathbf{A}' \cdot \mathbf{A}' = \mathbf{A} \cdot \mathbf{A}$ using the dot product in equation 5.

Space time is points $t, \mathbf{x}$ and distances between points are measured via the dot product. For for small distances in space time $dt, d\mathbf{x}$ we can define a length ds

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2.$$

This is often called the Minkowski metric.

Lorenz transformations transfer between reference frames.

Lorenz transformation form a continuous group.

Lorenz transformation preserves the dot product of a 4-vector with itself in equation 5. That means the dot product of a 4-vector with itself is a **relativistic invariant**. The dot product of a 4-vector is independent of reference frame.

In many settings, instead of defining the dot product as in equation 5 the signs are reversed, giving

$$\mathbf{A} \cdot \mathbf{A} = -c^2 A_t^2 + A_x^2 + A_y^2 + A_z^2.. \tag{6}$$

1.4 Causality

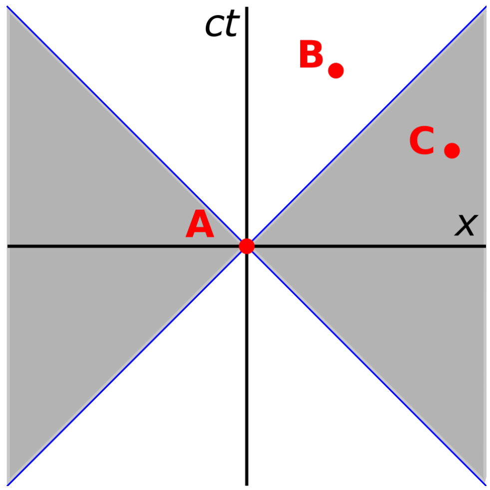

Figure 3: The slope of a line on this plot gives its speed. Events A and B can be causally connected but not events A and C. Information passing from A to C would be moving faster than the speed of light.

Consider two points in space time $\mathbf{x}_1, t_1$ and $\mathbf{x}_2, t_2$. We take $dt = t_2 - t_1$, $dx = x_2 - x_1$, $dy = y_2 - y_1$ and $dz = z_2 - z_1$. We compute

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2$$

Points in space time are separated by ds^2. This can be negative!

Since light travels at the speed of light, two points in space time along the path of a photon have $ds = 0$.

If $ds^2 > 0$ then a path moving at a speed below the speed of light can connect the two points.

If $ds^2 < 0$ then only paths moving above the speed of light can connect the two points.

Information cannot travel faster than the speed of light. Points with $ds^2 > 0$ are **causally connected** and points with $ds^2 < 0$ are not.

Sometimes the causally connected region is referred to as 'within the light cone' or where there are time-like, rather than space-like trajectories.

1.5 Lorenz contraction

We consider two objects that are separated by L and are observed in a frame where both are stationary. Object a is at $x_a = 0$ at all times t. The another object b is at $x_b = L$ at

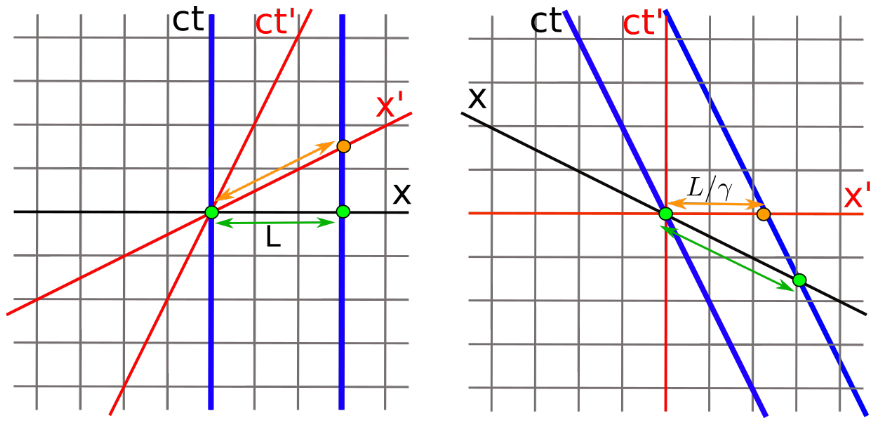

Figure 4: The Lorenz contraction shown in both reference frames. The thick black lines show the world lines of two objects. They are stationary in the x, ct coordinate system on the left and separated by distance L. Three events are shown in both coordinate systems. On the right the distance between the two objects at the same time is L/γ due to the Lorenz contraction.

all times.

The Lorenz transformation to a frame moving with speed β is (repeating equation 2),

$$
\begin{aligned}
x' &= \gamma x - \gamma\beta ct \\
t &= -\gamma\beta c^{-1} x + \gamma t.
\end{aligned}
\tag{7}
$$

Using this Lorenz transformation for object a

$$
\begin{aligned}
x'_a &= -\gamma\beta ct \\
t'_a &= \gamma t
\end{aligned}
$$

we can write x'_a in terms of t'_a with

$$x'_a = -\beta ct'_a$$

The object looks like it is moving, as expected.

We now look at the transformation of object b,

$$
\begin{aligned}
x'_b &= \gamma L - \gamma\beta ct \\
t'_b &= -\gamma\beta c^{-1} L + \gamma t
\end{aligned}
$$

We solve for

$$\gamma t = -t'_b + \gamma\beta c^{-1} L$$

Insert this into the expression for x'_b

$$x'_b = \gamma L - \beta c(t'_b + \gamma\beta c^{-1} L)$$
$$= \gamma L(1-\beta^2) - \beta c t'_b$$

recalling that $\gamma = (1-\beta^2)^{-\frac{1}{2}}$

$$x'_b = L\sqrt{1-\beta^2} - \beta c t'_b.$$

If we measure the distance between object a and object b at the same time in the new frame (so $t'_a = t'_b$) then we measure a distance between the two $(x'_b - x'_a)$ of

$$L' = L\sqrt{1-\beta^2} = L\gamma^{-1}$$

If object a and object b are really the ends of a train, then the lenght between them in the x, t frame is L. In the x', t' frame we measure the distance between the ends of the moving train at the same time in this frame. The distance we measure would be $L' < L$. This means the train appears to be shorter! This is known as the Lorenz contraction. Note that the length is the distance between two space time positions with the same time in a particular reference frame. The space time coordinates where we measured L are not the same as the space time coordinates where we measured L', as shown in Figure 4.

Notice that the expression for contraction was independent of the sign of β (setting the direction of one frame moving w.r.t to the other).

Figure 4, left side, illustrates that distances along the x' axis are shorter than they seem. Likewise, distances along the x axis on the right, are shorter than they seem. The length of the orange arrow on the left, is L/γ measured along the x' axis in the x', t' reference frame, and it looks like it is longer than L which is the distance of the green arrow which is measured in the x, t coordinate frame.

A **standard ruler** gives the distance between two points at the same time, measured in a particular reference frame. Figure 4 shows that each reference frame carries its own notion of a standard ruler.

1.6 Time dilation

We consider two events, the first at $x_a = 0, t_a = 0$ and the second at $x_b = 0, t_b = T$. In this coordinate frame the time between the two events is T and both events are at the same location, as shown in Figure 5. We do a Lorenz transformation to another frame using equation 7

$$x'_a = 0$$
$$t'_a = 0$$
$$x'_b = -\gamma\beta c T$$
$$t'_b = \gamma T$$

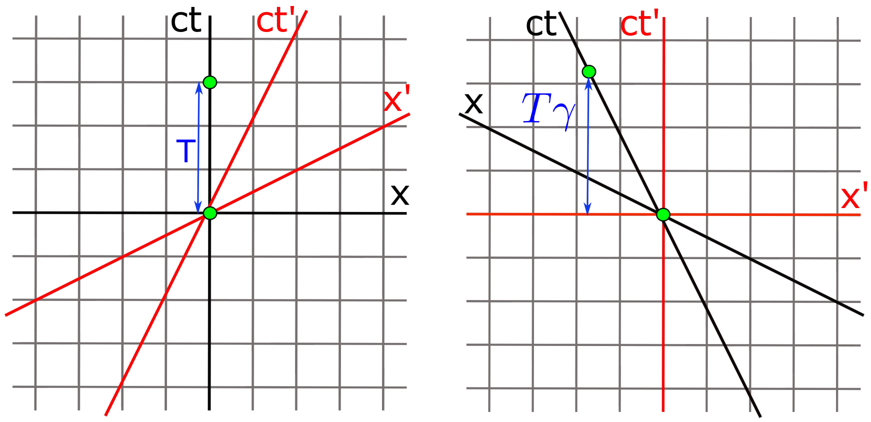

Figure 5: Time dilation shown in both frames. On the left two events are separated in space but not time in the x, ct frame.

We notice that the two events appear to be separated in time by γT. As $\gamma > 1$ the interval of time in the new frame is larger than that in the original frame. This is known as time dilation. Notice that the position of object b has changed!

Again notice that the size of the dilation does not depend on the direction of motion.

Time dilation is not the same as the Doppler shift. Time dilation is for events that can be at two different locations rather than the time between events (peaks of light and related to the frequency of light) seen at a specific location.

Time dilation implies that two clocks that are moving relativistically with respect to one another tick at different rates, as viewed from a third reference frame.

A clock tick of 1 second is measured in the frame of the clock, so the clock remains at the same position in that reference frame. Now consider another reference frame. The clock is moving in that reference frame. An observer (with his or her own clock for reference) would measure a longer length of time between the moving clock's ticks. The size of the time dilation does not depend on whether the clock is moving away from or toward the observer.

A clock can move away and then come back and the passage of time this clock experiences would have been less than for the stationary clock.

1.7 Four velocity

Consider a particle with mass m. We first consider its reference frame, where it is not moving. We place the particle at the origin and give the particle a clock. Time advances as usual. The trajectory of the particle is shown in Figure 6 on the left.

Let's define a vector that characterizes the arrow of time as seen in the particle's

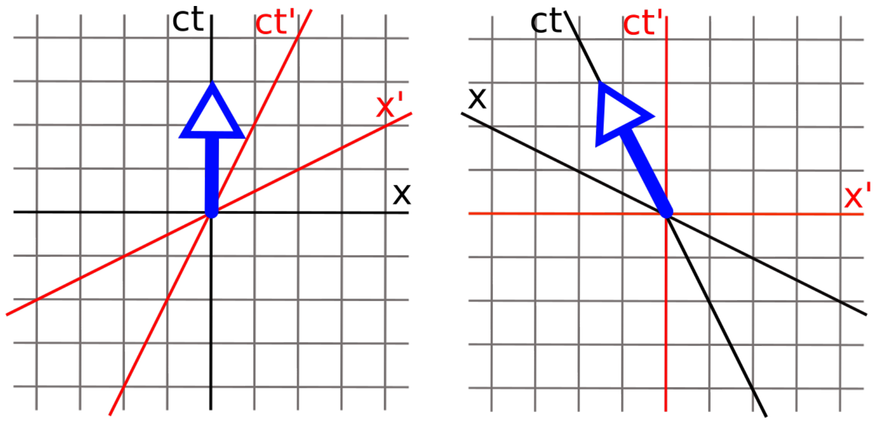

Figure 6: On the left we show the trajectory of a particle in its reference frame.

reference frame

$$\mathbf{u} = (1, 0, 0, 0)$$

in the particle's coordinate frame with coordinates (ct, x, y, z). This vector can also written as

$$\mathbf{u} = (u_t, u_x, u_y, u_z).$$

We now transform this vector $\mathbf{u}$ using a Minkowski transformation to a frame that moves with velocity v (with $\beta = v/c$ and $\gamma = 1/\sqrt{1-\beta^2}$ as usual). With v in the x direction, and $u_t = c$

$$\begin{aligned} u'_t &= \gamma u_t = \gamma c \\ u'_x &= \gamma\beta u_t = \gamma\beta c \\ u'_y &= u'_z = 0 \end{aligned}$$

In the new reference frame the vector becomes

$$\mathbf{u}' = (\gamma c, \gamma\beta c, 0, 0). \tag{8}$$

Since in our original reference, the particle was not moving, in the new reference frame the particle moves with velocity v.

Because $\mathbf{u}$ and $\mathbf{u}'$ are related via Lorenz transformation, they have the same length, where length is computed via the dot product of equation 5.

$$\mathbf{u} \cdot \mathbf{u} = c^2 = \mathbf{u}' \cdot \mathbf{u}'.$$

Notice that u'_x in equation 8 has units of velocity!

The **4-velocity** is a 4-vector defined with the Lorenz transformation that relates the rest frame of a particle to that of the observer in a frame where the particle has velocity $\mathbf{v}$;

$$\mathbf{u} \equiv (\gamma c, \gamma \mathbf{v}). \tag{9}$$

Here $\mathbf{v}$ is a three-vector.

1.8 Four momentum

From the definition of 4-velocity (equation 9) we can define a **4-momentum**

$$\mathbf{P} = m\mathbf{u} = (m\gamma c, m\gamma \mathbf{v}) \tag{10}$$

where $\mathbf{v}$ is the velocity of the particle. Notice that the x, y, z coordinates of this is $\mathbf{p} = m\gamma\mathbf{v}$ which we recognize as a three-vector describing the relativistic generalization of momentum. What is does the p_t component correspond to? If we take $p_t c = m\gamma c^2$ we recognize the relativistic generalization of energy.

The **4-momentum** can be defined as

$$\mathbf{P} = (E/c, \mathbf{p}) = (m\gamma c, m\gamma \mathbf{v}). \tag{11}$$

Here the 3-vector for relativistic momentum

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

and relativistic energy

$$E = \gamma m c^2.$$

With dot product as defined in equation 5, the relativistic invariant

$$\begin{aligned}
\mathbf{P}\cdot\mathbf{P} &= (E/c)^2 - p^2 \\
&= \gamma^2 m^2 c^2 - \gamma^2 m^2 v^2 \\
&= \gamma^2 m^2 c^2 (1 - v^2/c^2) \\
&= m^2 c^2. \\
E^2 - p^2 c^2 &= m^2 c^4.
\end{aligned}$$

We recognize the relativistic invariant that can be computed from the energy and momentum of a relativistic particle.

Notice that if we define the 4-momentum as in equation 10 then it makes sense to associate the energy with the t-component of the 4-momentum.

Note you will see variants of these definitions that have different factors of the speed of light c in them. Sometimes c is set to 1. Sometimes c is removed from the dot product and time is multiplied by c in the coordinates describing space-time.