Lorentz Transformations

J. Franklin

April 20, 2020

A common strategy in physics is to identify a physically relevant invariant quantity and generate transformations that preserve it. As an example, we expect a meter stick to have the same length no matter what you do to it.[1] For a physical theory to manifest this length invariance, its equations must be unchanged, in form, under the transformation that leaves lengths unchanged (a notion known as "covariance").

We'll work out the example of a general transformation that preserves length as a warmup, and then describe another, less familiar invariant quantity that comes directly from observation. The transformation that preserves this new "length" is a fundamental building block in any physical theory that deals with high energy particles. Both the length and its associated transformation provide examples of hyperbolic geometry in action, hence this motivating note.

Length Preservation

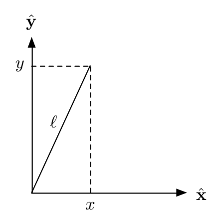


Figure 1: A point in two dimensions can be labelled by its x and y coordinates.

For the two-dimensional plane, we have orthogonal "x" and "y" directions, each with a unit vector pointing in the direction of increasing coordinate value, $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ (bold face for vectors, hats to indicate unit length). Any point can be identified with its x and y locations along these axes as shown in Figure 1, and then the distance to the origin is given by Pythagoras:

$$\ell = \sqrt{x^2 + y^2}.$$

[1]For the most part.

We want a new coordinate system in two dimensions, with new basis vectors $\hat{\bar{\mathbf{x}}}$ and $\hat{\bar{\mathbf{y}}}$, and new point descriptions $\bar{x}$ and $\bar{y}$ measured relative to these basis vectors. The most general linear transformation that leaves the origin unchanged (our choice, for simplicity) is

$$\bar{x} = Ax + By \qquad \bar{y} = Fx + Gy, \tag{1}$$

for constants $\{A, B, F, G\}$. The target invariance is $\bar{\ell} = \sqrt{\bar{x}^2 + \bar{y}^2} = \ell$. It's easier to work with the squares, so we'll enforce $\bar{\ell}^2 = \ell^2$:

$$\bar{x}^2 + \bar{y}^2 = x^2 \left(A^2 + F^2\right) + 2xy\,(AB + FG) + y^2 \left(B^2 + G^2\right) = x^2 + y^2. \tag{2}$$

This equation must hold for all points in the plane, so that it implies the three independent constraints

$$A^2 + F^2 = 1 \qquad AB + FG = 0 \qquad B^2 + G^2 = 1, \tag{3}$$

leaving us with a one parameter family of transformations. Let A be that parameter, then we have

$$F = \pm\sqrt{1 - A^2} \qquad B = \mp\sqrt{1 - A^2} \qquad G = A \tag{4}$$

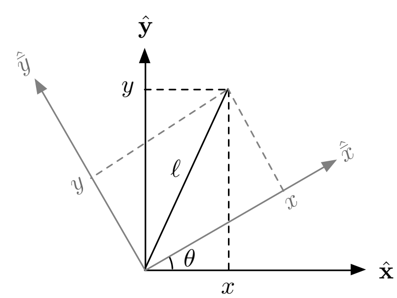
(there's also a set of solutions with $G = -A$, but we'll leave that for now). Picking the lower signs,

$$\bar{x} = Ax + \sqrt{1 - A^2}y \qquad \bar{y} = -\sqrt{1 - A^2}x + Ay. \tag{5}$$

It's a strange-looking way to express what we know must be rotations (either clockwise or counterclockwise depending on our choice of upper or lower signs above), so let $A = \cos\theta$ (appropriately constraining $A \in [-1, 1]$ which makes the transformation real), then we have, in matrix form, the familiar relation between the coordinate systems

$$\begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{6}$$

and the two sets of axes and point labels are shown below.



To produce a rotation in the other direction, just take $\theta \to -\theta$ in (6), a procedure that also generates the inverse of the matrix in (6) with little computation.

Frame Setup

Our next job is to find a new invariant, something other than length, and develop the transformation that preserves it. That task will require both some setup and some physical inputs. We'll do the setup first.

You have a horizontal and vertical axis, and make measurements with respect to those — call that coordinate system L. Moving at constant speed v to the right (tradition) through L is another pair of axes, call those $\bar{L}$. At time $t = 0$, the origins of L and $\bar{L}$ coincide. A picture of the temporal evolution (a "stillie") is shown in Figure 2.

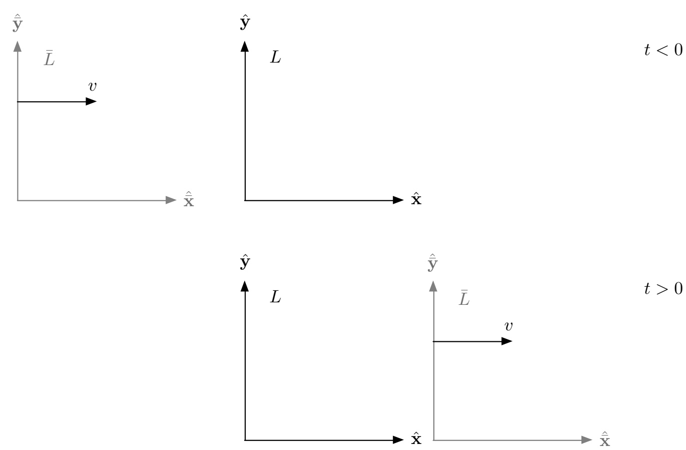


Figure 2: A coordinate system L is at rest, with $\bar{L}$ moving at constant speed to the right with speed v relative to L.

The typical way to represent these two "inertial frames" (coordinate systems moving with constant relative velocity) is to produce a snapshot at $t > 0$, with the $\bar{L}$ axes shifted up a bit, and drawn "in" the L system, as in Figure 3. The separation is solely for visual clarity, the horizontal axes in L and $\bar{L}$ are shared.

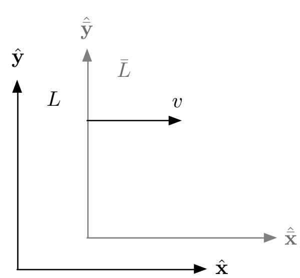


Figure 3: It is typical to draw the $\bar{L}$ axes shifted up a little so that you can see them.

Any point in L can be represented in $\bar{L}$ and vice versa. As an example, the origin of

$\bar{L}$ has coordinate values $\bar{x} = 0$ and $\bar{y} = 0$ in $\bar{L}$. That same point has coordinates $x = vt$, $y = 0$ when described in L.

Velocity Addition

Suppose you have a "ball" (arbitrary massive body) moving with velocity $v_{B\bar{L}}$ relative to $\bar{L}$. What is the velocity of the ball relative to L, v_{BL}? The situation is shown graphically in Figure 4.

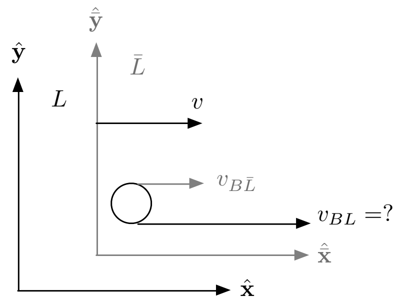


Figure 4: A ball moves with velocity $v_{B\bar{L}}$ relative to $\bar{L}$. What is the ball's speed relative to L?

We can use our everyday experience with relative velocity to generate the correct expression.[2] If $v_{B\bar{L}} = 0$, we must have $v_{BL} = v$, and if $v_{B\bar{L}} = v$, the velocity with respect to L should be $v_{BL} = 2v$. So in general, we have $v_{BL} = v + v_{B\bar{L}}$. Turning it around, if you know the ball's velocity with respect to L, v_{BL}, you can find the velocity of the ball relative to $\bar{L}$: $v_{B\bar{L}} = v_{BL} - v$. The point is that the ball's speed is different in the two frames. We can track the ball's location in each. If we throw the ball at $t = 0$ (when the origins of L and $\bar{L}$ coincide), then the location of the ball in L is $x = v_{BL}t$, and in $\bar{L}$, $\bar{x} = v_{B\bar{L}}t = (v_{BL} - v)t$.

Now for the physics — if the ball is instead the front of a light beam that was switched on at time $t = 0$, the *experimental fact* is that the beam is measured traveling at speed $v_{BL} = c$ in L, where c is the "speed of light", $c \approx 3 \times 10^8$ m/s, and in $\bar{L}$ its speed is *also* c, $v_{B\bar{L}} = c$. The speed of light is the same in both L and $\bar{L}$. This measurement has been done many times, and always holds. It forces us to change our notion of space, both to include time, so that we have "space-time," but also to extend the idea of "length" to include time.

New Invariant

The immediate shift that makes any of this possible is to promote time to a "coordinate." In physics, we are obsessed with units, and time is measured in seconds while traditional

[2]Although as will become clear in a moment, this form of velocity addition cannot hold for all speeds.

lengths are measured in meters. Those aren't the same, we need a factor that can turn meters into seconds, and that everyone can agree upon. The speed of light, c, serves nicely — everyone measures it to have the same speed, and since it is a speed, its m/s units can turn t in seconds into ct in meters. Going back to L and $\bar{L}$, in addition to spatial axes, we give each one an orthogonal (?) "temporal" axis. Now L gives point locations by reporting x, y and ct, while $\bar{L}$ reports $\bar{x}$, $\bar{y}$ and $c\bar{t}$, it has its own, different, clock.

Thinking of the front of the light beam, its horizontal location in $\bar{L}$ is $\bar{x} = c\bar{t}$, and in L: $x = ct$. Squaring both of these, we get

$$\bar{x}^2 - (c\bar{t})^2 = 0 = x^2 - (ct)^2 . \tag{7}$$

We have a quadratic relation on either side of the zero here that is similar to our previous $\bar{x}^2 + \bar{y}^2 = x^2 + y^2$. If we promote the equality to a general rule (i.e. for *all* points now, not just those with[3] $x = ct$), then we have a new invariant "Minkowski length" in the two coordinate systems

$$\bar{x}^2 - (c\bar{t})^2 = x^2 - (ct)^2 . \tag{8}$$

What is the coordinate transformation, now involving time, that preserves this new target invariant?

Lorentz Boost

To develop the transformation that supports (8), start with the linear, origin-preserving,

$$c\bar{t} = Act + Bx \qquad \bar{x} = Fct + Gx, \tag{9}$$

where we don't have to worry about putting bars on c (that's the point). Inserting this form in (8) gives

$$\bar{x}^2 - (c\bar{t})^2 = x^2\left(-B^2 + G^2\right) + 2(ct)x\left(FG - AB\right) + (ct)^2\left(-A^2 + F^2\right) = x^2 - (ct)^2 \tag{10}$$

and we learn that (compare with (3))

$$-A^2 + F^2 = -1 \qquad AB - FG = 0 \qquad -B^2 + G^2 = 1. \tag{11}$$

We again have a set of three equations for four "unknowns," leaving us with a one-parameter transformation, pick A as that parameter,

$$F = \pm\sqrt{-1 + A^2} \qquad B = \pm\sqrt{-1 + A^2} \qquad G = A, \tag{12}$$

where, again, there is a solution pair with $G = -A$ that we omit. Taking the lower signs, the transformation is

$$c\bar{t} = Act - \sqrt{-1 + A^2}x \qquad \bar{x} = -\sqrt{-1 + A^2}ct + Ax. \tag{13}$$

[3] This loophole can be closed by considering the beam of a flashlight pointed in the $\hat{\mathbf{y}}$ direction at time $t = 0$.

It would be natural, especially in the setting of this class, to take $A = \cosh(\eta)$ and write, in analogy with (6),

$$\begin{pmatrix} c\bar{t} \\ \bar{x} \end{pmatrix} = \begin{pmatrix} \cosh(\eta) & -\sinh(\eta) \\ -\sinh(\eta) & \cosh(\eta) \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}, \tag{14}$$

but it's easier to connect the form in (13) to the physical configuration shown in Figure 3.

Take the horizontal location of the origin in $\bar{L}$, $\bar{x} = 0$, expressed in L, $x = vt$. We can use this concrete pair in $\bar{x} = -\sqrt{-1 + A^2}ct + Ax$ to isolate A:

$$0 = -\sqrt{-1 + A^2}ct + Avt \longrightarrow A = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{15}$$

and then we can write the transformation in matrix form making explicit reference to the relative speed v,

$$\begin{pmatrix} c\bar{t} \\ \bar{x} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & -\frac{\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}} \\ -\frac{\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}} & \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}, \tag{16}$$

with the inverse transformation given by taking $v \to -v$ (after all, from the point of view of $\bar{L}$, L is moving to the *left* with constant speed v). This transformation, relating time and space in L and $\bar{L}$ while preserving the Minkowski length between points is called a "Lorentz boost."

Time Dilation

One immediate consequence of the transformation (16) is the phenomenon of "time dilation." Suppose you have a clock at rest at the origin in $\bar{L}$. "At rest" means that the clock is not moving relative to $\bar{L}$, so that at time $\bar{t} = 0$, the clock is at $\bar{x} = 0$, and at a later time $\bar{t}$ in $\bar{L}$, we still have $\bar{x} = 0$, the clock is still at the origin. Meanwhile, in L, at time $t = 0$ (we'll synchronize the times so that at $\bar{t} = 0$, we have $t = 0$), the clock is at $x = 0$, and at time t, it is at $x = vt$. Question: For the time interval $0 \to \bar{t}$ that elapses in $\bar{L}$, how much time has elapsed in L, $0 \to t$?

We'll use the Lorentz boost to relate t and $\bar{t}$, noting that $x = vt$,

$$c\bar{t} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \left(ct - \frac{v}{c} x \right) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \left(ct - \frac{v}{c} vt \right) = ct \frac{1 - \frac{v^2}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{17}$$

or, solving for t,

$$t = \frac{\bar{t}}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{18}$$

The factor multiplying $\bar{t}$ is greater than one, so that $t > \bar{t}$, more time has elapsed in L than in $\bar{L}$.

As an example, take $v = (3/5)c$, then

$$t = \frac{5}{4}\bar{t}, \tag{19}$$

so that if one year elapses on the clock in $\bar{L}$ (i.e. $\bar{t} = 1$ year), $t = 1.25$ year. This effect is real, and has been measured using a pair of clocks — you put one clock on a jet (that clock is at rest in "$\bar{L}$"), leave one on the ground (the clock at rest in "L"), send the jet off at "high speed" (still slow compared to c), and compare the clocks when the jet returns. The one that was on the jet shows less elapsed time than the one that remained on the ground.

Proper Time

A final note, related to the above time dilation section. Imagine *you* are $\bar{L}$, moving around through L. Your wristwatch (whatever that is!) remains at rest, relative to you, on your wrist, and ticks off a time $\bar{t}$. That time is different from the time registered on a clock hanging on the wall (which is not at rest relative to you) in L. As you move an infinitesimal distance dx through L in time dt, your wristwatch does not move relative to you, $d\bar{x} = 0$, and ticks off a time $d\bar{t}$ that is different from dt. But the invariant Minkowski length from (8), applied to these infinitesimals, must be the same, that's what started us off,

$$-c^2 d\bar{t}^2 = -c^2 dt^2 + dx^2. \tag{20}$$

This special time, the time registered on your wristwatch, at rest on your arm, is called the "proper time" and is typically denoted τ. Its defining property, from (20), is

$$d\tau^2 = dt^2 - \frac{dx^2}{c^2}. \tag{21}$$

Suppose you now move along a trajectory in L given by $x(t)$. The infinitesimal form of the Minkowski length relation (21) holds for each interval, and $dx = \frac{dx(t)}{dt} dt$, so we could write

$$d\tau^2 = \left(1 - \frac{\dot{x}(t)^2}{c^2}\right) dt^2, \tag{22}$$

really just an infinitesimal generalization of time dilation (solve (18) for $\bar{t}$ and square) with "instantaneous" speed $v = \dot{x}(t)$.[4] But there's another option, we could parametrize the motion through L using the proper time τ instead of the L clock's time t. We have to

[4]A dot refers to the derivative of a function with respect to its argument, so that $\dot{x}(t) \equiv \frac{dx(t)}{dt}$, while $\dot{x}(\tau) \equiv \frac{dx(\tau)}{d\tau}$.

parametrize both position and time in L using τ: $x(\tau)$ and $t(\tau)$ both change with τ. Now (20), with $dx = \frac{dx(\tau)}{d\tau} d\tau$ and $dt = \frac{dt(\tau)}{d\tau} d\tau$ becomes

$$-c^2 = -\left(c\dot{t}(\tau)\right)^2 + \left(\dot{x}(\tau)\right)^2 \tag{23}$$

which gives a version of arc-length parametrization for the "motion." Think of defining the velocity vector, tangent to the motion in time and space,

$$\dot{x}^\mu(\tau) \doteq \begin{pmatrix} c\dot{t}(\tau) \\ \dot{x}(\tau) \end{pmatrix}, \tag{24}$$

and the "metric"

$$\eta_{\mu\nu} \doteq \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \tag{25}$$

then we can express (23) as a generalized length requirement for the tangent vector

$$\sum_{\mu,\nu} \dot{x}^\mu \eta_{\mu\nu} \dot{x}^\nu = -c^2. \tag{26}$$