Physics 200a Relativity notes Shankar (2006)

Let us go over how the Lorentz transformation was derived and what it represents.

An event is something that happens at a definite time and place, like a firecracker going off. Let us say I assign to it coordinates (x, t) and you, moving to the right at velocity u, assign coordinates (x', t'). It is assumed that when you, sitting at $x' = 0$ passed me (sitting at $x = 0$), we set our clocks to zero: $t = t' = 0$. Thus our origins in space-time coincide.

Let us see how the coordinates would have been related **in pre-Einstein days**.

First, once we synchronize clocks at $t = t' = 0$, they remain synchronized for all future times $t = t'$. This is the notion of absolute time we all believe in our daily life.

You will assign to the event the spatial coordinates

$$x' = x - ut. \tag{1}$$

This relationship is easy to understand: your origin is ut meters to the right of mine (Figure (2)) because you have been moving to the right at velocity u for time t so the x-coordinate you assign to an event will be less by ut compared to what I assign. You can invert this relation to say

$$x = x' + ut. \tag{2}$$

How are these two equations modified post-Einstein? If the velocity of light is to be same for both you and me, it is clear we do not agree on lengths or times or both. Thus if I predict you will say the event is at $x' = x - ut$, you will say that my lengths need to be modified by a factor γ so that the correct answer is
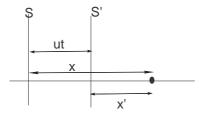
$$x' = \gamma(x - ut). \tag{3}$$

FIG. 1. The same event (solid dot) is assigned coordinates (x, t) by me (S) and (x', t') by you (S'). In the pre-Einstein days $t = t'$ not only initially (when our origins crossed) but always.

Likewise when you predict I will say $x = x' + ut'$ I will say, "No, your lengths are off, so the correct result is

$$x = \gamma(x' + ut')."\tag{4}$$

Note two things. First, I leave open the option that the time elapsed between when we synchronized clocks and when the firecracker went off is t' for you and t for me, with two times being possibly different. Next, the "fudge factor" for converting your lengths to mine and mine to yours are the same γ. This comes from the postulate that both observers are equivalent.

So let us look at the equations we have:

$$x = \gamma(x' + ut')\tag{5}$$
$$x' = \gamma(x - ut).\tag{6}$$

We proceed to nail down γ as follows. The event in question was a fire cracker going off. Suppose when our origins coincided we sent off a light pulse that this pulse set off the firecracker. Since the light pulse took t seconds to travel x meters according to me and took t' seconds to go x' meters according to you and we both agree on the value of c, it must be true for *this particular event* that

$$x = ct \quad \text{and} \quad x' = ct'.\tag{7}$$

Let us multiply the LHS of Eqn 5 by the LHS of 6 and equate the result to the product of the RHS's to get

$$xx' = \gamma^2(xx' + xut' - x'ut - u^2tt'), \text{and upon setting } x = ct,\ x' = ct' \text{ we get}\tag{8}$$
$$c^2tt' = \gamma^2(c^2tt' + uctt' - uct't - u^2tt') \quad \text{and now upon cancelling } tt'\tag{9}$$
$$\gamma^2 = \frac{1}{1 - \frac{u^2}{c^2}}\tag{10}$$
$$\gamma = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}}.\tag{11}$$

Note that once we have found γ it does not matter that we got it form this specific event involving a light pulse. It can applied to Eqns. (5,6) valid for a generic event. Putting γ back into Eqn. (6) we obtain

$$x' = \frac{x - ut}{\sqrt{1 - \frac{u^2}{c^2}}}.\tag{12}$$

Open Yale courses

If we now go to Eqn 5 and isolate t' we find

$$t' = \frac{t - \frac{ux}{c^2}}{\sqrt{1 - \frac{u^2}{c^2}}} \tag{13}$$

upon remembering that

$$1 - \frac{1}{\gamma^2} = \frac{u^2}{c^2}. \tag{14}$$

To summarize, if I am frame S and you are S', and you are moving to the right (increasing x direction) at speed u and my coordinates for an event are (x, t) and yours are (x', t') the Lorentz transformation tells us that

$$x' = \frac{x - ut}{\sqrt{1 - u^2/c^2}} \tag{15}$$

$$t' = \frac{t - \frac{u}{c^2}x}{\sqrt{1 - u^2/c^2}} \tag{16}$$

If we consider two events labelled 1 and 2, then the coordinate differences obey

$$\Delta x' = \frac{\Delta x - u\Delta t}{\sqrt{1 - u^2/c^2}} \tag{17}$$

$$\Delta t' = \frac{\Delta t - \frac{u}{c^2}\Delta x}{\sqrt{1 - u^2/c^2}} \tag{18}$$

where $\Delta x = x_2 - x_1$ etc., and differences are not necessarily small.

If you want to get my coordinates in terms of yours, you need to invert Eqns. (15) and 16. The answer is we get the same equations but with $u \to -u$. The same goes for the differences. As a result the differences will be related as follows;

$$\Delta x = \frac{\Delta x' + u\Delta t'}{\sqrt{1 - u^2/c^2}} \tag{19}$$

$$\Delta t = \frac{\Delta t' + \frac{u}{c^2}\Delta x'}{\sqrt{1 - u^2/c^2}} \tag{20}$$

Now for the velocity transformation law .

Let us follow a particle as it moves by an amount Δx in time Δt according to me and $\Delta x'$ in time $\Delta t'$ according to you. Let us agree that velocities are defined as follows once and for all:

$$v = \frac{\Delta x}{\Delta t} \qquad \text{according to me} \tag{21}$$

$$w = \frac{\Delta x'}{\Delta t'} \qquad \text{according to you} \tag{22}$$

$$u = \text{ your velocity relative to me} \tag{23}$$

In the above all $\Delta' s$ better be infinitesimals going to zero, as we are talking about instantaneous velocities and the derivative needs to be taken.

Suppose I see a particle with velocity v. What is the velocity w according to you? To get this we take the ratio of the equations (37 -38) that give the primed coordinate differences in terms of unprimed ones:

$$w = \frac{\Delta x'}{\Delta t'} = \frac{\Delta x - u\Delta t}{\Delta t - \frac{u}{c^2}\Delta x} = \frac{v - u}{1 - \frac{uv}{c^2}} \tag{24}$$

where in the last step we have divided top an bottom by Δt. It is good to check that for small velocities (dropping the terms that go as $1/c^2$) we get results agreeing with common sense.

Let us get used to going from your description to mine. Suppose you think a particle has velocity w. What will I think its velocity is? Now we use Eqns.(19-20). Taking the ratios as before and recalling the definition of w we get

$$v = \frac{w + u}{1 + \frac{vw}{c^2}}, \tag{25}$$

which just has the sign of u reversed in Eqn. (24) as expected and $w \leftrightarrow v$.

Suppose you see an object moving at $w = 3c/4$ and you yourself are moving relative to me at $u = 3c/4$. In the old days I would expect that object to be moving at $1.5c$. However the correct answer is

$$v = \frac{\frac{3c}{4} + \frac{3c}{4}}{1 + \frac{9}{16}} = \frac{24}{25}c. \tag{26}$$

It is interesting to see that if you choose to apply this to a light pulse seen by you ($w = c$) the speed I will find (for any relative velocity u between us)

$$v = \frac{c + u}{1 + u/c} = c. \tag{27}$$

Time dilatation. Let me carry a clock with ticks $\Delta t = \tau_0$ seconds apart. If the two events are two successive ticks, the spatial separation is $\Delta x = 0$ since the clock is at rest for me. You think the time difference between ticks is

$$\Delta t' = \frac{\tau_0 - 0\frac{u^2}{c^2}}{\sqrt{1 - u^2/c^2}} = \frac{\tau_0}{\sqrt{1 - u^2/c^2}} \tag{28}$$

Thus you say my clock is slow. This is what I will hear from anyone moving relative to me.

Let us rederive this using Eqns (19-20). Thus set

$$\Delta t = \tau_0 = \frac{\Delta t' + \frac{u}{c^2}\Delta x'}{\sqrt{1 - u^2/c^2}} \tag{29}$$

and note that the other equation says

$$\Delta x = 0 = \frac{\Delta x' + ut'}{\sqrt{1 - u^2/c^2}} \tag{30}$$

which means $\Delta x' = -u\Delta t'$. (This just means you think I and my clock are moving to the left at speed u) Feeding this into Eqn.29 to get

$$\tau_0 = \frac{\Delta t'(1 - u^2/c^2)}{\sqrt{1 - u^2/c^2}} \tag{31}$$

which agrees with Eqn. 28.

Length contraction Suppose you are carrying a rod of length L_0. This is its rest length. Suppose you are moving relative to me at velocity u. How long will I say it is? To find its length I have two of my co-moving assistants measure the location of its front and back ends *at the same time.* These two measurements, which can be seen as two events, have $\Delta x = L$, the length according to me, and $\Delta t = 0$ since the measurements of a moving rod's ends have to be simultaneous. (What happens if I measure one end, take a break and them measure the other?)

The separation in space between the two events is the rest length L_0 of the rod according to you. Thus

$$L_0 = \frac{L - 0}{\sqrt{1 - u^2/c^2}} \text{or} \ \ L = L_0\sqrt{1 - u^2/c^2}. \tag{32}$$

that is I will say your meter stick is actually $\sqrt{1 - u^2/c^2}$ meters long, i.e., shortened.

Open Yale courses

Likewise you will see my meter sticks are contracted. Suppose the contraction factor $\sqrt{1 - u^2/c^2} = .5$. I say your meter sticks are $.5m$ long. You say yours are fine and mine are $.5m$ long. How do you understand a guy with a shortened meter stick accusing you of having a shortened meter stick? Should I not say your are $2m$ long? The answer is this : You will say that I measured the leading end of your meter stick first and then after a delay , during which time the stick slid by, I measured the other end. That is, you will say I did not measure the ends simultaneously. I will insist I did, and both are right since simultaneity is not absolute.

Order of events and causality If event 1 causes event 2, no one should see the cause come after the effect for this leaves time for some one to prevent the cause from itself occurring. On the other hand if the two events are not causally related, a reversed order, while strange from our daily experience, will not lead to logical contradictions.

According to the LT

$$\Delta t' = \frac{\Delta t - \frac{u}{c^2}\Delta x}{\sqrt{1 - u^2/c^2}} \tag{33}$$
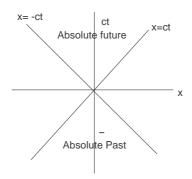
Let $\Delta t = t_2 - t_1 > 0$, so that I think 2 occurred after 1. Then we can find an observer for whom $\Delta t' < 0$ provided

$$\frac{u}{c^2}\Delta x > \Delta t \tag{34}$$

$$\frac{u}{c} > \frac{c\Delta t}{\Delta x} \tag{35}$$

If $c\Delta t > \Delta x$ the order can be reversed only if $u/c > 1$ which is impossible. In other words if a light signal can go a distance greater than the spatial separation between events, their order cannot be reversed in this theory. This is because if a light signal has enough time to interpolate between the events, event 1 *could have been* the cause of event 2 and we do not want the time- order of such (potentially) causally related pairs reversed. On the other hand if $c\Delta t < \Delta x$, even a light pulse could not have made it from one event to the other and the allows us to find frames with $u/c < 1$ in which the order is reversed. This is harmless since if a light signal cannot connect the events, neither can any other means, and the events could not have been causally connected.

In summary:

The light cone is defined by x=± ct

FIG. 2. The nature of space-time around the point $(0,0)$: events in the forward and backward light-cones lie in the absolute future and past (with the time-order of events relative to $(0,0)$ not negotiable), while those outside have non-absolute time-ordering relative to $(0,0)$: for example those that occur just above the $x-axis$ occur later than the event $(0,0)$ according to me but could be earlier than that event according to another observer.

If there is time for a light signal to go from the first event to the second, the theory assumes they could have been causally connected and does not allow us to find a frame with the order reversed, while if there is not enough time for a light signal to go from the first event to the second, the theory assumes they could not have been causally connected and allows us to find frames with the order reversed.

Thus if I am at $(0,0)$, space-time is divided into three regions: absolute future (which has $ct > |x|$ which I can affect from the origin using light or something slower)) absolute past ($ct < |x|$, which could have affected me here and now using light or something slower) and the "elsewhere" region where events I say happen later (earlier) could have happened earlier (later) for another observer moving at a speed $u < c$. See Figure 2.

Four vectors Let us define a four-vector

$$X = (x_0, x_1, x_2, x_3) \equiv (ct, x, y, z) \equiv (x_0, \mathbf{r}) \tag{36}$$

where all four components now have units of length. I urge you to check that under a LT describing motion along the x-axis, the components of X transform as follows:

$$x_1' = \frac{x_1 - \beta x_0}{\sqrt{1 - \beta^2}} \tag{37}$$

$$x_0' = \frac{x_0 - \beta x_1}{\sqrt{1 - \beta^2}}, \tag{38}$$

$$x_2' = x_2 \tag{39}$$

$$x_3' = x_3 \tag{40}$$

Open Yale courses

where $\beta = \frac{u}{c}$.

If you rewrote this as

$$x_1^{'} = x_1 \frac{1}{\sqrt{1-\beta^2}} - x_0 \frac{\beta}{\sqrt{1-\beta^2}} \tag{41}$$

$$x_0^{'} = x_0 \frac{1}{\sqrt{1-\beta^2}} - x_1 \frac{\beta}{\sqrt{1-\beta^2}} \tag{42}$$

you can see how similar this is to the equations relating (x, y) to (x', y') under rotation of axes:

$$x' = x \cos\theta + y \sin\theta \tag{43}$$

$$y' = y \cos\theta - x \sin\theta \tag{44}$$

No one is surprised that two people cannot agree on what the x-component of a point is or whether or not two points lie on the x-axis (since my x-axis is not yours) but it is surprising to see that this applies to $(x_0, x_1) \equiv (ct, x)$. Before Einstein we still needed t to label an event (besides (x, y, z), but no transformation of frame (including going to a moving frame) could change the time coordinate t. The reason time is now the fourth dimension is that it mixes freely with the other three under LT's.

Now, we saw that in the case of rotated axes, two observers who could not agree on the coordinates of a point, still agreed on one how far the point was from the origin:

$$x^2 + y^2 = x^{'2} + y^{'2} \tag{45}$$

More generally they agreed on the dot product of two vectors

$$\mathbf{A} \cdot \mathbf{B} = A_x B_x + A_y B_y = AB \cos\theta = A_x' B_x' + A_y' B_y' \tag{46}$$

with $\mathbf{A}$ being the position vector $\mathbf{r} = \mathbf{i}x + \mathbf{j}y$ in the above example.

In the LT, we cannot hope that

$$\frac{1}{\sqrt{1-\beta^2}} = \cos\theta \qquad \frac{\beta}{\sqrt{1-\beta^2}} = \sin\theta \tag{47}$$

since this does not respect $\cos^2\theta + \sin^2\theta = 1$. Thus the LT is not a rotation in the (ct, x) plane. However I strongly urge you to check that the there is a dot product $X \cdot X$ that does the trick:

$$X \cdot X = x_0^2 - x_1^2 = x_0^{'2} - x_1^{'2} \tag{48}$$

Open Yale courses

We call this the space-time interval s^2 between the origin and the point (x_0, x_1). More generally if there are two events separated in space by Δx_1 and time by Δx_0, then the square of the space-time interval between the events is

$$(\Delta s)^2 = (\Delta x_0)^2 - (\Delta x_1)^2 \tag{49}$$

and is the same for all observers. Note that s^2 and $(\Delta s)^2$ are not positive definite. We use the following terminology:

- space-like separation or space-like interval $\rightarrow (\Delta s)^2 < 0$, that is $\Delta x_1 > \Delta x_0$

- time-like separation $\rightarrow (\Delta s)^2 > 0$, that is $\Delta x_0 > \Delta x_1$

- light-like separation $\rightarrow (\Delta s)^2 = 0$, that is $\Delta x_0 = \Delta x_1$, events could lie on path of a light pulse.

In Figure 2, the points in the absolute past and future are separated from the origin by time-like separations, points in the other two regions are space-like separated and points on the light cone have a light-like separation.

Four momentum In the old days we manufactured new vectors from the one by taking derivatives:

$$\frac{d\mathbf{r}}{dt} = \mathbf{v} \tag{50}$$

is a vector since $d\mathbf{r}$ is a vector and dividing a scalar (something that does not transform when you rotate axes) like dt does not change this fact. We then made up a new vector called momentum by multiplying $\mathbf{v}$ by another scalar called m:

$$\mathbf{p} = m\mathbf{v}. \tag{51}$$

Let us try that here by analogy. Clearly $\Delta X = (\Delta x_0, \Delta x_1, \Delta x_2, \Delta x_3) \equiv (c\Delta t, \Delta \mathbf{r})$ is the change in the coordinates of a point in space-time. We cannot divide these four by Δt since t is just another component. This will be like dividing δx by Δy and that does not give us a velocity. The only scalar, something same for al observers, (or an invariant) is the space-time interval:

$$ds = \sqrt{(dx_0)^2 - (dx_1)^2} = cdt\sqrt{1 - \left(\frac{dx}{cdt}\right)^2} = cdt\sqrt{1 - v^2/c^2} \tag{52}$$

Open Yale courses

where v is the velocity of the particle as seen by the observer who has assigned the coordinates of X to that particle. It is better to drop the c which does not change the scalar nature, and use a quantity with dimensions of time

$$d\tau = dt\sqrt{1 - v^2/c^2} \qquad (53)$$

where $d\tau$ is called the proper-time between the two events. What is this quantity with units of time and yet same for all? It is the time measured by a clock moving with the particle. To see this let us go back to the definition

$$d\tau = \sqrt{dt^2 - dx^2/c^2}. \qquad (54)$$

This is the same in all frames. What is it in the frame moving with the particle? In this frame the two events on its trajectory occur at the same point – $dx = 0$, and $d\tau$ reduces to the time interval in the on-board or co-moving clock.

Let us now define a four-momentum as follows:

$$P = m\left(\frac{dx_0}{d\tau}, \frac{dx_1}{d\tau}, \frac{dx_2}{d\tau}, \frac{dx_3}{d\tau}\right) \qquad (55)$$

$$\equiv \left(c\frac{dt}{d\tau}, \frac{d\mathbf{r}}{d\tau}\right) \qquad (56)$$

$$\equiv (P_0, P_1, P_2, P_3) \qquad (57)$$

$$\equiv (P_0, \mathbf{P}) \qquad (58)$$

where m is called the mass of the particle (postulated to be the same in al frames) and I have written the expression for the components of P in many possible notations found in the literature.

Note that we have cleverly dealt with the question of taking the time-derivative: instead of taking my time-derivative or yours, we take it in the particle's frame. You and I do not agree on how many seconds it took the particle to go the distance Δx separating two nearby events in the particle's trajectory, but we all agree if we ask how long the particle says it took according to its own clock.

For now on let us ignore the components in the y and z directions, assuming all motion along x. (I will still call P the four-momentum.)

Given that P is a four-vector means the following: if in one frame it has components (P_0, P_1) then in another moving at speed u, it will have components P_0', P_1' which are related to each other just like the components of X were as in Eqns. (37,38).

We will exploit this in a moment but let us look at all the τ derivatives, in Eqn 55, which are unfamiliar quantities. The particle can have its own clock, but I want to see it all in terms of my clock and my time t. So let us trade τ derivatives of any function f for t=derivatives as follows:

$$d\tau = dt\sqrt{1 - v^2/c^2} \tag{59}$$

$$\frac{df}{d\tau} = \frac{df}{dt} \cdot \frac{dt}{d\tau} = \frac{df}{dt} \cdot \frac{1}{\sqrt{1 - v^2/c^2}}. \tag{60}$$

We can now write the four-momentum in Eq. 56 as

$$P = (P_0, P_1) = \left(\frac{mc}{\sqrt{1 - v^2/c^2}}, \frac{mv}{\sqrt{1 - v^2/c^2}} \right) \tag{61}$$

where we now have familiar velocity v everywhere. So take your pick: a familiar looking expression for four-momentum in terms of an unfamiliar quantity, the proper time τ as Eqn. (57), or the above expressions with new functions of a familiar velocity. The second version allows us to understand what the components of P mean. Consider first

$$P_1 = \frac{mv}{\sqrt{1 - v^2/c^2}} \tag{62}$$

If we consider a particle moving slowly, i.e., $v/c << 1$, we find (upon setting the denominator to unity)

$$P_1 = mv \tag{63}$$

Thus we conclude that P_1 stands for the momentum of the particle in the relativistic theory. However if the particle picks up speed we need to take into account the denominator. As $v \to c$, P_1 grows without limit: that is, in this theory, there is a limit to the particle velocity, but not its momentum.

Some people like to write

$$P_1 = \left(\frac{m}{\sqrt{1 - v^2/c^2}} \right) v \equiv m(v)v \tag{64}$$

where $m(v) = m/\sqrt{1 - v^2/c^2}$ is a new velocity dependent mass. They also refer to $m(0) = m$ as the rest mass m_0. Their point is that if you introduce a velocity dependent mass, then momentum can still be mass times velocity as in the old days. We will will not do that: for us m is always the rest mass and momentum is now a more complicated function of this mass and velocity.

Suppose v/c is small but not utterly negligible. Then we can use a slightly better formula for momentum by using

$$(1+x)^n = 1 + nx + \ldots$$

for $x << 1$, to write

$$\frac{1}{\sqrt{1 - v^2/c^2}} = (1 - v^2/c^2)^{-1/2} = 1 + v^2/2c^2 + .. \tag{65}$$

and

$$P_1 = mv + mv^3/2c^2 + .. \tag{66}$$

where the ellipses (dots) stand for smaller corrections we are ignoring. We can calculate more such correction terms or simply use the exact expression with the $\sqrt{1 - v^2/c^2}$ in it.

It is also clear that if we bring in P_2 and P_3 we just get the vector

$$\mathbf{P} = \frac{m\mathbf{v}}{\sqrt{1 - v^2/c^2}} \tag{67}$$

What does

$$P_0 = \frac{mc}{\sqrt{1 - v^2/c^2}}$$

stand for? If we set $v = 0$ we get the mass of the particle. Let us go to the next level of approximation and write as before,

$$P_0 = mc + \frac{1}{2c}mv^2 + \ldots \tag{68}$$

We see that if we multiply both sides by c something familiar emerges:

$$cP_0 = mc^2 + \frac{1}{2}mv^2 + \ldots \tag{69}$$

We see that the second term is just the kinetic energy. So it must be that all higher powers v/c shown by the ellipsis stand for corrections to kinetic energy as we consider faster particles.

But it must then be that the first term mc^2 also stands for energy, but of a particle at rest. This is called rest energy. Einstein did not tell us how to extract this energy (in contrast to kinetic energy of motion which can be extracted, say in hydroelectric power generators using turbines that slow

down the water). However when later on when people discovered fusion or fission, they found that some amount of mass was missing at the end of the reaction and that this missing mass (upon multiplying by c^2) exactly equalled the additional kinetic energy of the final particles. You should read up on this if you want. I will limit myself to the discussion of four-vectors.

To summarize what we have seen two four-vectors:

$$P = (P_0, P_1) = (\frac{E}{c}, p) \tag{70}$$

and

$$X = (x_0, x_1) = (ct, x). \tag{71}$$

In the relativistic picture, there emerge four vectors, three of whose components are familiar vectors from the old days and the fourth is what used to be a scalar (unaffected by rotations.) Now all four components can shuffled with each other, under Lorentz Transformations. (Note that under just ordinary rotations, the fourth components P_0 and x_0 are indeed unaffected, it takes a Lorentz transformation, or going to a moving and not just rotated frame to mix up all four components.) Note also that the forth components are familiar things times c as in $x_0 = ct$ or divided by c, as in $P_O = E/c$. It is for this reason that people often use unit in which $c = 1$. For example you can choose to measure distance in light seconds (distance light travels in 1 second, that is $3 \cdot 10^8 m/s$) and time in seconds. In these units the speed of light is one light second per second, i.e, $c = 1$. In cosmology you may want to use light years for distance and years for time. Later I will give you some practice in setting $c = 1$.

Consequences of P being a four vector.

- *The components of P transform as a four vector* as we go from one frame to another:

$$P_1' = \frac{P_1 - \beta P_0}{\sqrt{1 - \beta^2}} \tag{72}$$

$$P_0' = \frac{P_0 - \beta P_1}{\sqrt{1 - \beta^2}}, \tag{73}$$

$$P_2' = P_2 \tag{74}$$

$$P_3' = P_3 \tag{75}$$

where $\beta = \frac{u}{c}$. We will ignore P_2 and P_3 from now on.

Open Yale courses

It will be instructive for you verify that this transformation law is in accord with another way of finding P_0' and P_1'. To this end, consider a particle moving at speed v as seen by unprimed observer. Write explicit formulas for P_0 and P_1 in terms of m and v and keep them handy. Now ask what velocity w this particle will have as seen by the primed observer moving at velocity u. Write down her expressions for P_0' and P_1' in terms of w and now write w in terms of u and v. Check that the result agrees with what the LT above gives. This was assigned as a homework problem. To save paper you may set $c = 1$ in this exercise.

- $P_A \cdot P_B$ *is invariant, where P_A and P_B are are any two four-momenta, they could refer to two different particles A and B or the same particle, in which case $A = B$.*

First consider just one particle and the value of $P \cdot P$. Since it can be found in any frame, find it in its own co-moving frame. In this case $P_1 = 0$ and $P_0 = mc$. Thus

$$P \cdot P = P_0^2 - P_1^2 = m^2 c^2 \tag{76}$$

You may verify that if you went to a generic frame and wrote down expressions for P_0 and P_1 in terms of v, you will get the same result.

This is a very important result: $P^2 = m^2 c^2$.
(I will start referring to the dot product of a vector with itself as its square: $P^2 \equiv P \cdot P$).
If there are two particles we may assert that

$$P_A \cdot P_B = P_{A0} P_{B0} - P_{A1} P_{B1} = \frac{E_A E_B}{c^2} - P_{A1} P_{B1} \tag{77}$$

will have the same value for all observers. For example I can go the rest frame of particle B, I get $P_A \cdot P_B = E_A m_B$ where by E_A I mean the energy of A as measured in the rest-frame of B.

A photon has no rest mass. This means

$$K \cdot K = 0 \tag{78}$$

where K is a common name for the momentum as applied to a photon. It has no rest frame, in any frame it moves at c. However its components

K_0, K_1 will undergo LT as we change frames of reference. The components of K again stand for E/c and momentum of the photon. Zero rest mass means that

$$K_0 = K_1 \qquad \text{or } E = cp. \tag{79}$$

We will denote the components of K as follows

$$K = (\omega/c, k), \tag{80}$$

that is, the photon energy is denoted by ω and its momentum by k with $\omega = kc$ because of the condition $K \cdot K = 0$.

If a photon of energy ω hits me in the face, it might get absorbed and disappear in which case I will recoil with its momentum k and my mass will go up due to the absorbed energy.

What is my new mass m'? Your guess may be $m' = m + \omega/c^2$. Let us work it out.

Let $P = (mc, 0)$ denote my initial energy -momentum and $K = (\omega/c, k)$ that of the photon. In the end there is just me, but moving at some speed v to conserve momentum and a new rest mass m'. So we write

$$P + K = P' \tag{81}$$

where

$$P' = (\frac{m'c}{\sqrt{1 - v^2/c^2}}, \frac{m'v}{\sqrt{1 - v^2/c^2}}) \tag{82}$$

is my final energy-momentum.

Note that Eq. (81) is really four equations, one for each component of energy-momentum, just as $\mathbf{F} = m\mathbf{a}$ is really three equations. In our problem two of these, in the y and z directions vanish on both sides assuming the photon is along the x-axis. Thus in relativity conservation of energy and momentum become the single conservation law of energy-momentum.

Let me write down the two nontrivial equations:

$$mc + \omega/c = \frac{m'c}{\sqrt{1 - v^2/c^2}} \qquad \text{conservation of energy, more precisely, of } E/c. \tag{83}$$

$$0 + k = = \frac{m'v}{\sqrt{1 - v^2/c^2}} \qquad \text{momentum conservation} \tag{84}$$

Open Yale courses

Now you can juggle these equations and solve for m'. But here is a quicker way using four-vectors. Since we just want m' we need only calculate $P' \cdot P' = m'^2 c^2$. So we proceed as follows

$$m'^2 c^2 = P' \cdot P' \tag{85}$$
$$= (P + K) \cdot (P + K) \tag{86}$$
$$= P \cdot P + K \cdot K + 2P \cdot K \tag{87}$$
$$= m^2 c^2 + 0 + 2(mc\frac{\omega}{c} - 0 \cdot k) \tag{88}$$
$$m' = \sqrt{m^2 + 2m\frac{\omega}{c^2}} \tag{89}$$

If $\omega/mc^2 << 1$, i..e, the photon energy is much less than my rest energy, we can approximate as follows:

$$m' = m\sqrt{1 + 2\frac{\omega}{mc^2}} = m(1 + \frac{\omega}{mc^2} + ..) = m + \frac{\omega}{c^2} + .. \tag{90}$$

in accord with the naive expectation which ignores recoil. In other words not all the photon energy can go into boosting my rest mass since I also need to move to conserve the initial photon momentum. Thus the increased rest energy plus kinetic energy has to equal the photon energy.

Remember the following trick: in these four-vector equations., square that four-momentum about which you know the least since the answer for the square is always $m^2 c^2$. Another hint: sometimes the momentum you need to square may not be standing alone in one side of the equation. If this happens, isolate it (by moving other terms to the other side) and square it.

- Four momentum if conserved in one frame is conserved in any frame

  The virtue of momentum as we derived it is that if it is conserved in one frame, it will be conserved in any other. For example if particles A and B turn into C, D, E, and in one frame we have

$$P_{initial} = P_A + P_B = P_C + P_D + P_E = P_{final} \tag{91}$$

then in any other frame we will have

$$P'_{initial} = P'_{final} \tag{92}$$

because if two vectors $P_{initial}$ and P_{final} are equal in one frame they are equal in any other. This must be clear from analogy with usual vectors: if $\mathbf{A} + \mathbf{B} = \mathbf{C}$, that is the three form a triangle in one frame, then they will also form a triangle in a rotated frame. Or one can say that since $\mathbf{A} + \mathbf{B} - \mathbf{C} = 0$, the null vector, the LHS will be the null (zero) vector in any frame since the rotated version of the null vector is the null vector. In the case of the LT, if $P_{final} - P_{initial} = 0$, the difference vector will vanish in any frame if it vanishes in one.

It also follows that if you did not like my definition of four-momentum and made up your own, yours may not have the property of being conserved in all frames if conserved in one. Conservation is what makes an momentum important quantity and it has to be a four-vector if it is to be conserved in all frames.

Another worked example: Compton Effect

Imagine a photon of momentum K moving along the axis hitting a sitting electron of momentum P. If the photon comes back along the x-axis what is its new energy? I will show you how to do this problem using units in which $c = 1$. This makes the notation easier. I will also show you how in the end we can put back the c's guided by dimensions. With these units the initial and final photon momenta look as follows:

$$K = (\omega, \omega), \quad K' = (\omega', -\omega') \tag{93}$$

Why? First since $K \cdot K = K_0^2 - K_1^2 = 0$ implies $K_0 = \pm K_1$, that is $\omega/c = \pm k$ which (on setting $c = 1$) simplifies to $\omega = \pm k$ with the plus (minus) sign for the incoming (reflected) photon. The electron's four-momenta are, before and after,

$$P = (m, 0) \quad P' = (E, p) \text{ with } E^2 - p^2 = m^2 \tag{94}$$

Begin with the conservation law

$$K + P = K' + P' \tag{95}$$

Since no one cares about the details of P' we isolate and square it (take dot product with itself) because we know that for any particle, no matter how it is moving, the square of the four momentum is $m^2 c^2 = m^2$.

Here are the details.

$$P' \cdot P' \equiv (P')^2 = m^2 = (P + K - K')^2 \tag{96}$$
$$= P^2 + K^2 + K'^2 + 2(P \cdot K - 2P \cdot K' - 2K \cdot K') \tag{97}$$
$$= m^2 + 0 + 0 + 2\left(m\omega - m\omega' - (\omega\omega' - kk')\right) \tag{98}$$
$$0 = m(\omega - \omega') - 2\omega\omega' \quad (\text{ used } kk' = (\omega)(-\omega')) \tag{99}$$
$$\frac{1}{\omega'} = \frac{1}{\omega} + \frac{2}{m} \tag{100}$$
$$\frac{1}{\omega'} = \frac{1}{\omega} + \frac{2}{mc^2} \tag{101}$$

where in the last equation I restored the c^2 since ω is an energy and so is mc^2.

Yet another example Question: What is the minimum energy E of the incident proton that strikes the proton at rest so that in the end we have a p,p, p, and $\bar{p}$? (Antiparticles have the same mass as particles.)

Answer: The energy-momentum of the incident proton is $P_1 = (E, p)$, that of the target is $P_2 = (m, 0)$. (Remember $c = 1$). In the lab frame the total momentum is

$$P_{Tot}^{Lab} = P_1 + P_2 = (E + m, p)$$

. See Fig. 3.

In the minimal reaction the final four particles will have the minimum energy, but they can't all be at rest due to momentum conservation. So we go to the CM frame in which the two initial protons approach each other with opposing spatial momenta, that is their four-momenta are $(\bar{E}, \bar{p})$ and $(\bar{E}, -\bar{p})$. Thus the initial total momentum is $P_{Tot}^{CM} = (2\bar{E}, 0)$. The final four-momentum of the four particles, now allowed to be created at rest, is

$$P_{Tot}^{CM} = (4m, 0).$$

Now recall that for any four-vector $V^\mu = (V^0, V^1)$,

$$V \cdot V = (V^0)^2 - (V^1)^2$$

is invariant, same in all frames. Apply this to the total momentum and remember that for any one particle $P \cdot P = E^2 - p^2 = m^2$. We find

$$\text{CM frame} \qquad P_{Tot} \cdot P_{Tot} = 16m^2$$

$$\text{Lab frame} \qquad P_{Tot} \cdot P_{Tot} = E^2 + m^2 + 2mE - p^2 = m^2 + m^2 + 2Em$$

Open Yale courses

which gives the desired result
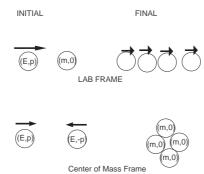
$$E = 7m$$

or $7mc^2$ if you put back c.

FIG. 3. The reaction seen in lab and CM frames. (E=Energy, p=momentum) given in brackets. Arrows indicate the spatial motion.