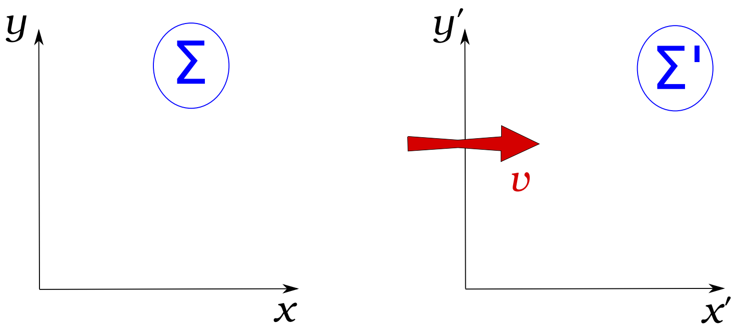
1 This writeup

In this writeup, the Lorentz transformations are derived for a standard boost, in three different ways.

The 'standard' configuration: Frames Σ' and Σ are coincident at time $t = t' = 0$ and their relative motion is in the common x, x' direction.



We consider the relativistic transformation connecting the same event observed from the two frames. This is an example of a Lorentz transformation. Lorentz transformations in general can include relative rotations between the two frames in addition to a boost (relative motion). The standard configuration contains two frames with a relative boost only. We refer to this transformation as the standard Lorentz boost.

For the standard configuration (representing a pure boost in the x direction), the Lorentz transformation between time and space coordinates is

$$ct' = \gamma_v \left(ct - \frac{v}{c} x \right) \tag{1}$$

$$x' = \gamma_v \left(x - \frac{v}{c} ct \right) \tag{2}$$

$$y' = y \tag{3}$$

$$z' = z \tag{4}$$

with

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

2 Deriving the LT: Preliminaries

We want to now derive the transformation equations for the standard configuration, i.e., the equations relating the coordinates (ct, x, y, z) and (ct', x', y', z') for the same event as measured from frame Σ and from frame Σ'.

It is reasonable to suppose that the transverse directions are unaffected, i.e., $y' = y$ and $z' = z$. This can be argued by considering a thought experiment in which observers in Σ and Σ' use identically constructed sticks to create marks on a wall at a certain height as they pass by it; one can then argue a contradiction if the marks are not at equal height. We thus only have to relate (ct', x') to (ct, x). The transverse spatial coordinates are only along for the ride.

A second observation is that the transformation from (ct, x) to (ct', x') must be *linear*. If not, a body observed from Σ to have constant velocity will be observed from Σ' to have acceleration, i.e., there would be fictitious forces felt in either Σ, or Σ', or both. This violates the tenet that physics is identical in all inertial frames. Thus, we can expect the transformation equations to look like

$$ct' = Act + Bx\,, \quad x' = Dct + Ex\,. \tag{5}$$

We now have to simply find the four constants A, B, D and E. The constants are expected to be functions of v and c, but (to preserve linearity).

We know what these transformations should look like in the limit $v \to 0$: they should reproduce the Galiliean transformations, $ct' = ct$, $x' = x - vt = \left(-\frac{v}{c}\right) ct + x$. This limit, together with dimensional considerations, could be used to restrict the transformations further. However, we will use something more direct: the definition of the standard configuration is that frame Σ' moves with speed v in the x-direction with respect to Σ, and that the origins coincided at $t = t' = 0$. Thus the origin of Σ' must satisfy $x = vt$ (seen from the Σ frame); since this point is defined as $x' = 0$, we get

$$0 = Dct + E(vt) \quad \implies \quad E = -\frac{c}{v}D$$

so that the transformations are

$$ct' = Act + Bx\,, \quad x' = Dct - \frac{c}{v}Dx\,. \tag{6}$$

Note that the second equation now has the form $x' = -D\frac{c}{v}(x - vt)$, so it's easy to imagine this reducing to the Galilean form for small v, provided that $D \to -\frac{v}{c}$ as $v \to 0$. We now have only three constants to determine.

To obtain these constants, we can impose the fact that the invariant interval is invariant: $c^2dt'^2 - dx'^2 = c^2dt^2 - dx^2$. For the standard configuration, this also means

$$c^2t'^2 - x'^2 = c^2t^2 - x^2 \tag{7}$$

because the origins of the two frames coincide at $t = t' = 0$. Using this requirement straightforwardly gives the constants we are after. This is done in three different ways below. They achieve the same purpose but it is instructive to go through each; in particular the 3rd derivation introduces infinitesimal transformations, an important idea.

3 Derivation 1: Brute force

Our first derivation is a bit "brute-force"-ish and will win no prizes for elegance. Simply take the transformation equations (6) with unknown coefficients, and plug them into the equation (7) for the invariant interval.

$$(Act + Bx)^2 - (Dct - Dcx/v)^2 \;=\; c^2t^2 - x^2$$

Equating coefficients gives

$$A^2 - D^2 = 1, \quad D^2(c/v)^2 - B^2 = 1, \quad AB + D^2(c/v) = 0.$$

We have three equations for three coefficients, so obtaining the coefficients is now strakghtforward, although admittedly a bit tedious. When solving, I strongly suggest use $\beta = v/c$ to simplify the clutter while solving.

The equations involve squares of D and A, so the signs are not uniquely determined from this set of equations, but from the requirement that $A \to 1$, $D \to -\frac{v}{c}$ at small v (Galilean transformations), we can infer that A and D should respectively be positive and negative. The solutions are

$$A = \frac{1}{\sqrt{1 - v^2/c^2}}, \quad B = \frac{-v/c}{\sqrt{1 - v^2/c^2}}, \quad D = \frac{-v/c}{\sqrt{(1 - v^2/c^2)}}$$

So the transformation equations are

$$ct' = \frac{ct - vx/c}{\sqrt{1 - v^2/c^2}}, \quad x' = \frac{x - vt}{\sqrt{1 - v^2/c^2}}$$

We now have clear reason to define

$$\gamma_v \;=\; \frac{1}{\sqrt{1 - v^2/c^2}}$$

so that the derived Lorentz transformations are

$$ct' = \gamma_v(ct - vx/c), \quad x' = \gamma_v(x - (v/c)ct),$$

and of course $y' = y$, $z' = z$.

4 Derivation 2: Light cone coordinates

This derivation is a bit more elegant.
Let's define the coordinates

$$x^{\pm} = ct \pm x$$

which are known as 'light cone' coordinates. The invariant interval is then

$$c^t dt^2 - dx^2 = dx^+ dx^-$$

This quantity is invariant under the transformation. Thus if the transformation decreases x^+ by some factor it needs to increase x^- by the same factor. Call this factor e^{ϕ}:

$$x'^+ = e^{-\phi} x^+, \quad x'^- = e^{\phi} x^-.$$

Then

$$\begin{aligned} ct' = \tfrac{1}{2}(x'^+ + x'^-) &= \tfrac{1}{2}(e^{-\phi} x^+ + e^{\phi} x^-) \\ &= \tfrac{1}{2}(e^{-\phi}(ct + x) + e^{\phi}(ct - x)) = \cosh\phi ct - \sinh\phi x \end{aligned}$$

and similarly $x' = -\sinh\phi ct + \cosh\phi x$. Here $\sinh\phi = \frac{1}{2}(e^{\phi} - e^{-\phi})$ and $\cosh\phi = \frac{1}{2}(e^{\phi} + e^{-\phi})$ are the hyperbolic sine and hyperbolic cosine functions.
We have obtained the transformation equations

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \cosh\phi & -\sinh\phi \\ -\sinh\phi & \cosh\phi \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}$$

in terms of the variable ϕ, which we still have to relate to v/c.
Comparing with Eq. (6), $x' = Dct - (c/v)Dx$, we get

$$\tanh\phi = \frac{v}{c}$$

which leads to

$$\cosh^2\phi = \frac{1}{1 - \tanh^2\phi} = \frac{1}{1 - v^2/c^2} = \gamma_v^2$$

and

$$\sinh^2\phi = \cosh^2\phi - 1 = \frac{1}{1 - v^2/c^2} - 1 = \frac{v^2/c^2}{1 - v^2/c^2} = \gamma_v^2 \frac{v^2}{c^2}$$

The requirement of reducing to the Galilean transformations for small v implies that $\cosh\phi$ and $\sinh\phi$ are both positive; thus $\cosh\phi = \gamma_v$ and $\sinh\phi = \gamma_v(v/c)$.
We have thus re-derived the transformation equations for the Lorentz boost in standard configuration:

$$ct' = \gamma_v(ct - vx/c), \quad x' = \gamma_v(x - (v/c)ct),$$

The variable ϕ is known as the rapidity.

<table>
<tr>
<td>
The derivation of LT using light cone coordinates mirrors the following derivation of the transformation matrix for rotations:
Rotation preserves $x^2 + y^2$. Defining $z = x + iy$, this means $|z'|^2 = |z|^2$ or $z'z'^* = zz^*$. So z' can differ from z only by a phase. Try $z' = ze^{-i\zeta}$, $z'^* = z^*e^{i\zeta}$; then

$$\begin{aligned} x' &= \frac{1}{2}\left(z' + z'^*\right) = \frac{1}{2}\left(ze^{-i\zeta} + z'^*e^{i\zeta}\right) \\ &= \frac{1}{2}\left(x\left[e^{-i\zeta} + e^{i\zeta}\right] + iy\left[e^{-i\zeta} - e^{i\zeta}\right]\right) \\ &= x\cos\zeta + y\sin\zeta \end{aligned}$$

and similarly $y' = -x\cos\zeta + y\sin\zeta$.
</td>
<td>
This is the imaginary version of our derivation 2 using light cone coordinates. 1+1 dimensional Lorentz boosts are very much like 2-dimensional rotations, except for the occational minus sign or factor of i.
Can we interpret ζ as the angle of rotation? That would require additional input. (Of course, we know from other derivations that ζ is the angle of rotation)
</td>
</tr>
</table>

5 Derivation 3: Infinitesimal boosts

This derivation will use matrix language. This will be elegant in a different way, and seemingly a bit roundabout, but the detour will be instructive. As a bonus we will learn a technique — focusing on infinitesimal transformations — that is at the heart of the study of *Lie groups*.

5.1 Formulate as a matrix

Since we are after the transformation that gives (ct', x') in terms of (ct, x), we need a 2×2 matrix, which we call Λ:

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \Lambda \begin{pmatrix} ct \\ x \end{pmatrix} \tag{8}$$

Of course, a full Lorentz transformation is a 4×4 matrix, but in this case, the perpendicular spatial directions are untouched by the transformation. This is the simplicity gained by focusing on the standard configuration. A sophisticated-sounding way of saying this: we are restricting to 1+1 dimensional spacetime.

We now try to express the invariance of our familiar 'interval' in matrix language, in particular, as a condition on Λ. Noting that

$$c^2t^2 - x^2 = \begin{pmatrix} ct & x \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}, \tag{9}$$

it seems useful to give a name to the matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Its extension to 4 dimensions is called the **metric tensor**. For now, we will refer to this 2×2

matrix as the metric tensor

$$g = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(If you have had some quantum mechanics, you would also recognize this matrix as the third Pauli matrix σ_z or σ_3.)
Eq. (8) implies $\begin{pmatrix} ct' & x' \end{pmatrix} = \begin{pmatrix} ct & x \end{pmatrix} \Lambda^{\mathrm{T}}$, where Λ^{T} is the transpose of Λ. The invariance of $c^2t^2 - x^2$ thus means

$$\begin{pmatrix} ct & x \end{pmatrix} \Lambda^{\mathrm{T}} g \Lambda \begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} ct & x \end{pmatrix} g \begin{pmatrix} ct \\ x \end{pmatrix}$$

Since this is true for any event (ct, x), it implies

$$\Lambda^{\mathrm{T}} g \Lambda = g$$

This is the condition that the transformation matrix Λ must satisfy.

5.2 Focus on an infinitesimal piece

Having formulated the problem as a matrix condition, we first try to find the transformation matrix for an infinitesimal boost. Noting that Λ is the unit matrix for the case of zero velocity (zero boost), it makes sense to try

$$\Lambda(\epsilon) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \epsilon K = I + \epsilon K$$

The trick of considering infinitesimal transformation is a common trick in the study of *Lie groups*. This trick is due to Sophus Lie, after whom *Lie algebra* and *Lie groups* are named.

Here K is a 2×2 matrix which we will try to determine. Here ϵ is an infinitesimal quantity, proportional to the velocity provided both ϵ and v are small. The matrix K (sans a conventional factor of i) is called the *generator* of the set of 1-dimensional Lorentz boosts.
Inserting $\Lambda = I + i\epsilon K$ into the invariance condition and retaining only leading terms in ϵ, we get

$$K^{\mathrm{T}} g + gK = 0$$

This is the condition for our infiniteslimal boost to obey the invariance of the interval $\Delta\tau$. This equation is satisfied by matrices of the form (see panel) $K = \begin{pmatrix} 0 & k_2 \\ k_2 & 0 \end{pmatrix}$. We can choose $k_2 = 1$ without loss of generality, since any constant factor can be abosrbed in the definition of ϵ. Then an infinitesimal boost would be $\Lambda = \begin{pmatrix} 1 & \epsilon \\ \epsilon & 1 \end{pmatrix}$.

You could define $K = \begin{pmatrix} k_1 & k_2 \\ k_3 & k_4 \end{pmatrix}$ and use it in the condition $K^{\mathrm{T}} g + gK = 0$. This leads to $k_1 = k_4 = 0$ and $k_2 = k_3$.

However, we expect the off-diagonal terms in the transformation matrix to have negative sign, in order to connect to the Galilean transformations for in the $c \to \infty$ limit. So it makes more sense to choose

$$K = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}; \qquad \Lambda(\epsilon) = \begin{pmatrix} 1 & -\epsilon \\ -\epsilon & 1 \end{pmatrix}$$

This choice might feel like looking ahead unfairly, but you don't really have to cheat — the derivation will work out correctly with any choice of k_2.

Let's celebrate: we have obtained the generator and the form of the infinitesimal Lorentz boost!

5.3 From the infinitesimal to the finite boost

How do we get a finite boost from an infinitesimal one? Easy, we just apply the tiny boost successively as many times as we need it. Imagine that we need to apply it N times to obtain a boost of finite strength $\zeta = N\epsilon$. Thus

$$\Lambda(\zeta) = \Lambda(\epsilon)\lambda(\epsilon)\lambda(\epsilon)\ldots\lambda(\epsilon) = [\lambda(\epsilon)]^N .$$

I am being a bit cavalier here, since ϵ is infinitesimal and ζ is finite, we need the $N \to \infty$ limit, but this will not be a problem. Now note that $I + \epsilon K = e^{\epsilon K}$ because ϵ is infinitesimal, so that $\Lambda(\zeta) = \left[e^{\epsilon K}\right]^N = e^{\zeta K}$. Thus we have found the boost transformation

$$\Lambda(\zeta) = \exp[\zeta K] = \exp\begin{pmatrix} 0 & -\zeta \\ -\zeta & 0 \end{pmatrix}$$

We don't yet know how to interpret ζ; is it maybe proportional to the velocity v? To find out, we expand out the exponential and obtain series

Here's a more picturesque path from the infinitesimal boost K to the finite boost $e^{\zeta K}$:

$$\Lambda(\zeta) \sim \lambda(\epsilon)^N \sim \left(I + \frac{\zeta}{N}K\right)^N .$$

Since ϵ is infinitesimal, this is exact in the limit $N \to \infty$:

$$\Lambda(\zeta) = \lim_{N\to\infty}\left(I + \frac{\zeta}{N}K\right)^N$$

This is a famous representation of the exponential function (right panel). Thus $\Lambda(\zeta) = e^{\zeta K}$.

To physics students, $y = \lim_{N\to\infty}\left(1 + \frac{x}{N}\right)^N$ is not the best known represetntation of the exponential function $y = e^x$. To derive, expand in a binomial series. The N-dependence in each term cancels for $N \to \infty$, leaving

$$y = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = e^x$$

Each of these steps also works if x is a *matrix*; we are using this result for the matrix ζK.
What is the exponential of a matrix? It is generally defined as the series $\sum \frac{x^m}{m!}$ used above.
The same definition is used for the exonential of operators in quantum mechanics.

in ζ for each element; the result is

$$\Lambda(\zeta) = \begin{pmatrix} \cosh\zeta & -\sinh\zeta \\ -\sinh\zeta & \cosh\zeta \end{pmatrix}.$$

We can now relate ζ to v/c by comparing with Eqs. (5). However, comparing with the previous derivation, we recognize that ζ is exactly the rapidity $\phi = \tanh^{-1}(v/c)$. Obtaining the boost transformation in terms of the speed is now the same steps as in the previous derivation; we won't repeat.

If we had chosen K to be $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, we would have had $\zeta = -\phi$, but would end up with the same Lorentz boost equations. With our choice, ϵ has the interpretation of an infinitesimal rapidity. (This is the same as an infinitesimal (v/c), as $\lim_{x\to 0}\tanh x = x$). However, the perhaps artificial-looking choice of $k_2 = -1$ is not necessary for the derivation; please feel free to derive the boost equations using another choice of k_2. In that case, ϵ would be the infinitesimal boost divided by whatever constant you defined k_2 to be.