Derivation of the Lorentz Transformation

Lecture note for course Phys171H "Introductory Physics: Mechanics and Relativity"

written by Victor Yakovenko

Department of Physics, University of Maryland, College Park

15 November 2004

In most textbooks, the Lorentz transformation is derived from the two postulates: the equivalence of all inertial reference frames and the invariance of the speed of light. However, the most general transformation of space and time coordinates can be derived using only the equivalence of all inertial reference frames and the symmetries of space and time. The general transformation depends on one free parameter with the dimensionality of speed, which can be then identified with the speed of light c. This derivation uses the group property of the Lorentz transformations, which means that a combination of two Lorentz transformations also belongs to the class Lorentz transformations.

The derivation can be compactly written in matrix form. However, for those not familiar with matrix notation, I also write it without matrices.

1) Let us consider two inertial reference frames O and O'. The reference frame O' moves relative to O with velocity v in along the x axis. We know that the coordinates y and z perpendicular to the velocity are the same in both reference frames: $y = y'$ and $z = z'$. So, it is sufficient to consider only transformation of the coordinates x and t from the reference frame O to $x' = f_x(x,t)$ and $t' = f_t(x,t)$ in the reference frame O'.

From translational symmetry of space and time, we conclude that the functions $f_x(x,t)$ and $f_t(x,t)$ must be linear functions. Indeed, the relative distances between two events in one reference frame must depend only on the relative distances in another frame:

$$x_1' - x_2' = f_x(x_1 - x_2, t_1 - t_2), \qquad t_1' - t_2' = f_t(x_1 - x_2, t_1 - t_2). \tag{1}$$

Because Eq. (1) must be valid for any two events, the functions $f_x(x,t)$ and $f_t(x,t)$ must be linear functions. Thus

$$x' = Ax + Bt, \tag{2}$$
$$t' = Cx + Dt, \tag{3}$$

where A, B, C, and D are some coefficients that depend on v. In matrix form, Eqs. (2) and (3) are written as

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} \tag{4}$$

with four unknown functions A, B, C, and D of v.

2) The origin of the reference frame O' has the coordinate $x' = 0$ and moves with velocity v relative to the reference frame O, so that $x = vt$. Substituting these values into Eq. (2), we find $B = -vA$. Thus, Eq. (2) has the form

$$x' = A(x - vt), \tag{5}$$

so we need to find only three unknown functions A, C, and D of v.

3) The origin of the reference frame O has the coordinate $x = 0$ and moves with velocity $-v$ relative to the reference frame O', so that $x' = -vt'$. Substituting these values in Eqs. (2) and (3), we find $D = A$. Thus, Eq. (3) has the form

$$t' = Cx + At = A(Ex + t), \tag{6}$$

where we introduced the new variable $E = C/A$.

Let us change to the more common notation $A = \gamma$. Then Eqs. (5) and (6) have the form

$$x' = \gamma\,(x - vt), \tag{7}$$
$$t' = \gamma\,(Ex + t), \tag{8}$$

or in the matrix form

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \gamma \begin{pmatrix} 1 & -v \\ E & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix}, \tag{9}$$

Now we need to find only two unknown functions γ_v and E_v of v.

4) A combination of two Lorentz transformations also must be a Lorentz transformation. Let us consider a reference frame O' moving relative to O with velocity v_1 and a reference frame O'' moving relative to O' with velocity v_2. Then

$$\begin{aligned} x'' &= \gamma_{v_2}\,(x' - v_2 t'), \qquad & x' &= \gamma_{v_1}\,(x - v_1 t), \\ t'' &= \gamma_{v_2}\,(E_{v_2} x' + t'), \qquad & t' &= \gamma_{v_1}\,(E_{v_1} x + t), \end{aligned} \tag{10}$$

or in the matrix form

$$\begin{pmatrix} x'' \\ t'' \end{pmatrix} = \gamma_{v_2} \begin{pmatrix} 1 & -v_{v_2} \\ E_{v_2} & 1 \end{pmatrix} \begin{pmatrix} x' \\ t' \end{pmatrix}, \qquad \begin{pmatrix} x' \\ t' \end{pmatrix} = \gamma_{v_1} \begin{pmatrix} 1 & -v_{v_1} \\ E_{v_1} & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix}. \tag{11}$$

Substituting x' and t' from the second Eq. (35) into the first Eq. (35), we find

$$\begin{aligned} x'' &= \gamma_{v_2}\,\gamma_{v_1}\,[(1 - E_{v_1} v_2)x - (v_1 + v_2)t)], \\ t'' &= \gamma_{v_2}\,\gamma_{v_1}\,[(E_{v_1} + E_{v_2})x + (1 - E_{v_2} v_1)t], \end{aligned} \tag{12}$$

or in the matrix form

$$\begin{pmatrix} x'' \\ t'' \end{pmatrix} = \gamma_{v_2}\,\gamma_{v_1} \begin{pmatrix} 1 & -v_{v_2} \\ E_{v_2} & 1 \end{pmatrix} \begin{pmatrix} 1 & -v_{v_1} \\ E_{v_1} & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} = \gamma_{v_2}\,\gamma_{v_1} \begin{pmatrix} 1 - E_{v_1} v_2 & -v_1 - v_2 \\ E_{v_1} + E_{v_2} & 1 - E_{v_2} v_1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix}. \tag{13}$$

For a general Lorentz transformation, the coefficients in front of x in Eq. (7) and in front of t in Eq. (8) are equal, i.e. the diagonal matrix elements in Eq. (9) are equal. Eqs. (12) and (13) must also satisfy this requirement:

$$1 - E_{v_1} v_2 = 1 - E_{v_2} v_1 \quad \Rightarrow \quad \frac{v_2}{E_{v_2}} = \frac{v_1}{E_{v_1}} \, . \tag{14}$$

In the second Eq. (14), the left-land side depends only on v_2, and the right-hand side only on v_1. This equation can be satisfied only if the ratio v/E_v is a constant a independent of velocity v, i.e.

$$E_v = v/a, \tag{15}$$

Substituting Eq. (15) into Eqs. (7) and (8), as well as (9), we find

$$x' = \gamma_v \, (x - vt), \qquad t' = \gamma_v \, (xv/a + t), \tag{16}$$

or in the matrix form

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \gamma_v \begin{pmatrix} 1 & -v \\ v/a & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} . \tag{17}$$

Now we need to find only one unknown function γ_v, whereas the coefficient a is a fundamental constant independent on v.

5) Let us make the Lorentz transformation from the reference frame O to O' and then from O' back to O. The first transformation is performed with velocity v, and the second transformation with velocity $-v$. The equations are similar to Eqs. (35) and (11):

$$\begin{aligned} x &= \gamma_{-v} \, (x' + vt'), & x' &= \gamma_v \, (x - vt), \\ t &= \gamma_{-v} \, (-x'v/a + t'), & t' &= \gamma_v \, (xv/a + t), \end{aligned} \tag{18}$$

or in the matrix form

$$\begin{pmatrix} x \\ t \end{pmatrix} = \gamma_{-v} \begin{pmatrix} 1 & v \\ -v/a & 1 \end{pmatrix} \begin{pmatrix} x' \\ t' \end{pmatrix}, \qquad \begin{pmatrix} x' \\ t' \end{pmatrix} = \gamma_v \begin{pmatrix} 1 & -v \\ v/a & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} . \tag{19}$$

Substituting x' and t' from the first equation (18) into the second one, we find

$$x = \gamma_{-v} \, \gamma_v \, (1 + v^2/a) \, x, \qquad t = \gamma_{-v} \, \gamma_v \, (1 + v^2/a) \, t, \tag{20}$$

or in the matrix form

$$\begin{pmatrix} x \\ t \end{pmatrix} = \gamma_{-v} \, \gamma_v \begin{pmatrix} 1 & v \\ -v/a & 1 \end{pmatrix} \begin{pmatrix} 1 & -v \\ v/a & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} = \gamma_{-v} \, \gamma_v \begin{pmatrix} 1 + v^2/a & 0 \\ 0 & 1 + v^2/a \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} . \tag{21}$$

Eqs. (20) and (21) must be valid for any x and t, so

$$\gamma_{-v} \, \gamma_v = \frac{1}{1 + v^2/a} \, . \tag{22}$$

Because of the space symmetry, the function γ_v must depend only on the absolute value of velocity v, but not on its direction, so $\gamma_{-v} = \gamma_v$. Thus we find

$$\gamma_v = \frac{1}{\sqrt{1 + v^2/a}} \,. \tag{23}$$

6) Substituting Eq. (23) into Eqs. (16) and (17), we find the final expressions for the transformation

$$x' = \frac{x - vt}{\sqrt{1 + v^2/a}}, \qquad t' = \frac{xv/a + t}{\sqrt{1 + v^2/a}}, \tag{24}$$

or in the matrix form

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \frac{1}{\sqrt{1 + v^2/a}} \begin{pmatrix} 1 & -v \\ v/a & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix}. \tag{25}$$

Eqs. (24) and (25) have one fundamental parameter a, which has the dimensionality of velocity squared.

If $a < 0$, we can write it as

$$a = -c^2. \tag{26}$$

Then Eqs. (24) and (25) become the standard Lorentz transformation:

$$x' = \frac{x - vt}{\sqrt{1 - v^2/c^2}}, \qquad t' = \frac{-xv/c^2 + t}{\sqrt{1 - v^2/c^2}}, \tag{27}$$

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \frac{1}{\sqrt{1 - v^2/c^2}} \begin{pmatrix} 1 & -v \\ -v/c^2 & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix}. \tag{28}$$

It is easy to check from Eq. (27) that, if a particle moves with velocity c in one reference frame, it also moves with velocity c in any other reference frame, i.e. if $x = ct$ then $x' = ct'$. Thus the parameter c is the invariant speed. Knowing about the Maxwell equations and electromagnetic waves, we can identify this parameter with the speed of light. It is straightforward to check that the Lorentz transformation (27) and (28) preserves the space-time interval

$$(ct')^2 - (x')^2 = (ct)^2 - x^2, \tag{29}$$

so it has the Minkowski metric.

If $a = \infty$, then Eqs. (24) and (25) produce the non-relativistic Galileo transformation:

$$x' = x - vt, \qquad t' = t, \qquad \begin{pmatrix} x' \\ t' \end{pmatrix} = \begin{pmatrix} 1 & -v \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix}. \tag{30}$$

If $a > 0$, we can write it as $a = \sigma^2$. Then Eqs. (24) and (25) describe a Euclidean space-time and preserve the space-time distance: $(x')^2 + (\sigma t')^2 = x^2 + (\sigma t)^2$.

Additional note for course Phys374
"Intermediate Theoretical Methods"
17 October 2006

The Lorentz transformation (28) can be written more symmetrically as

$$\begin{pmatrix} x' \\ ct' \end{pmatrix} = \frac{1}{\sqrt{1 - v^2/c^2}} \begin{pmatrix} 1 & -v/c \\ -v/c & 1 \end{pmatrix} \begin{pmatrix} x \\ ct \end{pmatrix}. \tag{31}$$

Instead of velocity v, let us introduce a dimensionless variable β, called *rapidity* and defined as

$$\tanh\beta = v/c, \tag{32}$$

where tanh is the hyperbolic tangent. Then Eq. (31) acquires the following form:

$$\begin{pmatrix} x' \\ ct' \end{pmatrix} = \begin{pmatrix} \cosh\beta & -\sinh\beta \\ -\sinh\beta & \cosh\beta \end{pmatrix} \begin{pmatrix} x \\ ct \end{pmatrix}. \tag{33}$$

Let us consider a combination of two consequtive Lorentz transformations (boosts) with velocities v_1 and v_2, as described in the first part. The rapidity β of the combined boost has a simple relation to the rapidities β_1 and β_2 of each boost:

$$\beta = \beta_1 + \beta_2. \tag{34}$$

Indeed, Eq. (34) represents the relativistic law of adding velocities

$$\tanh\beta = \frac{\tanh\beta_1 + \tanh\beta_2}{1 + \tanh\beta_1 \tanh\beta_2} \quad \Rightarrow \quad v = \frac{v_1 + v_2}{1 + v_1 v_2/c^2}. \tag{35}$$

Let us denote the 2×2 matrix in Eq. (33) as $\mathbf{M}(\beta)$. Then, the combination of two boosts has the simple matrix form

$$\mathbf{M}(\beta_1 + \beta_2) = \mathbf{M}(\beta_2)\,\mathbf{M}(\beta_1). \tag{36}$$

We see that the Lorentz transformations form a *group*, similar to the group of rotations, with the rapidity β being the (imaginary) rotation angle.