Einstein's 1905 Paper rides on a Kinematical Mistake to get the Lorentz Transformation

Karunesh M. Tripathi,

Retd. (vol.) Chief Engineer, Western Railway, Mumbai, India.

Author Email: karuneshtripathi@hotmail.com

M. Tech. Received June 1984, from I.I.T, Delhi, India.

Abstract:

Einstein's derivation of the Lorentz transformation in section 3 of his famous 1905 paper is based on kinematics of light rays in X, Y and Z directions. However, the velocity of light worked out in Y and Z directions, with respect to the frame moving in X direction, and as observed from the stationary frame, is found to be incorrect. This renders the entire derivation invalid.

It is further pointed out that even if the mistake was ignored, the relations are achieved by taking a time that is for travel of light from origin to a specific location (mirror). This limits the applicability of the relations only to events of light. Despite the restriction, however, Einstein extensively applied the relations to rigid bodies and moving clocks, even in the same paper.

It is also discussed as to how, in the derivation of his 1916 book, he used the relations derived for a moving light signal, on rigid bodies.

Further, a new objection-free method has been presented based on kinematics, using the principle of reciprocity of relative velocity between the two frames.

Key Words: Special Relativity, Lorentz Transformation,

PACS Nos, 03.30.+p, 11.30.Cp, 42.25.Fx, 42.25.Bs,

Introduction:

Einstein, in his 1905 paper, presented the theory of Special Relativity based on the principle that the time of travel of light between two locations was not the same in both the directions (of travel), except when the locations were in the observer's stationary frame. The corresponding exercise, appearing in section 3 of his paper (discussed below), culminated into the Lorentz transformation.

The derivation is all about equating, in a moving frame, the times of travel of light between two points, in the two directions , after casting them as functions of their own corresponding values, along with the distance between the two locations, as observed from the stationary frame. However, a kinematical mistake is found in calculation of the relative velocity of light in Y and Z directions, as observed from the stationary frame, and it renders the derivation invalid.

Further, even if the mistake is ignored for a moment, the Lorentz transformation relations have been achieved by taking a time that is for travel of light to a specific location. This limits the applicability of the relations only to events of light.

Working towards the solution, a new objection-free method has been presented towards the end, based on the reciprocity of relative velocity between the two frames, by kinematics.

The Section 3 of the paper is reproduced below in italics. My observations are intermittently placed, in normal font, between two dotted lines, as the matter progresses.

$3. Theory of the Transformation of Co-ordinates and Times from a Stationary System to another System in Uniform Motion of Translation Relatively to the Former

Let us in "stationary" space take two systems of co-ordinates, i.e. two sys- tems, each of three rigid material lines, perpendicular to one another, and issuing from a point. Let the axes of X of the two systems coincide, and their axes of Y and Z respectively be parallel. Let each system be provided with a rigid measuring-rod and a number of clocks, and let the two measuring-rods, and likewise all the clocks of the two systems, be in all respects alike.

Now to the origin of one of the two systems (k) let a constant velocity v be imparted in the direction of the increasing x of the other stationary system (K), and let this velocity be communicated to the axes of the co-ordinates, the relevant measuring-rod, and the clocks. To any time of the stationary system K there then will correspond a definite position of the axes of the moving system, and from reasons of symmetry we are entitled to assume that the motion of k may be such that the axes of the moving system are at the time t (this "t" always denotes a time of the stationary system) parallel to the axes of the stationary system.

We now imagine space to be measured from the stationary system K by means of the stationary measuring-rod, and also from the moving system k by means of the measuring-rod moving with it; and that we thus obtain the co-ordinates x, y, z, and ξ, η, ζ respectively. Further, let the time t of the stationary system be determined for all points thereof at which there are clocks by means of light signals in the manner indicated in $1; similarly let the time τ of the moving system be determined for all points of the moving system at which there are clocks at rest relatively to that system by applying the method, given in $1, of light signals between the points at which the latter clocks are located.

To any system of values x, y, z, t, which completely defines the place and time of an event in the stationary system, there belongs a system of values ξ, η, ζ, τ, determining that event relatively to the system k, and our task is now to find the system of equations connecting these quantities.

In the first place it is clear that the equations must be linear on account of the properties of homogeneity which we attribute to space and time.

If we place $x^J = x - vt$ it is clear that a point at rest in the system k must have a system of values x^J, y, z, independent of time. We first define τ as a function of x^J, y, z, and t. To do this we have to express in equations that τ is nothing else than the summary of the data of clocks at rest in system k, which have been synchronized according to the rule given in $1.

From the origin of system k let a ray be emitted at the time τ_0 along the X-axis to x^J, and at the time τ_1 be reflected thence to the origin of the co- ordinates, arriving there at the time τ_2; we then must have $\frac{1}{2}(\tau_0 + \tau_2) = \tau_1$, or, by inserting the arguments of the function τ and applying the principle of the constancy of the velocity of light in the stationary system:—

$$\frac{1}{2}\left[\tau(0,0,0,t)+\tau\left(0,0,0,t+\frac{x'}{c-v}+\frac{x'}{c+v}\right)\right]=\tau\left(x',0,0,t+\frac{x'}{c-v}\right)$$

Hence, if x' be chosen infinitesimally small,

$$\frac{1}{2}\left(\frac{1}{c-v}+\frac{1}{c+v}\right)\frac{\partial\tau}{\partial t}=\frac{\partial\tau}{\partial x'}+\frac{1}{c-v}\frac{\partial\tau}{\partial t}$$

Or,

$$\frac{\partial\tau}{\partial x'}+\frac{v}{c^2-v^2}\frac{\partial\tau}{\partial t}=0$$

It is to be noted that instead of the origin of the co-ordinates we might have chosen any other point for the point of origin of the ray, and the equation just obtained is therefore valid for all values of x', y, z.

An analogous consideration—applied to the axes of Y and Z—it being borne in mind that light is always propagated along these axes, when viewed from the stationary system, with the velocity $\sqrt{c^2-v^2}$ gives us

$$\frac{\partial\tau}{\partial y}=0\ ,\frac{\partial\tau}{\partial z}=0$$

Since τ is a *linear* function, it follows from these equations that

$$\tau=a\left(t-\frac{v}{c^2-v^2}x'\right)$$

where a is a function φ(v) at present unknown, and where for brevity it is assumed that at the origin of k, τ = 0, when t = 0.

Observations 1:

i) The relation obtained just above is a general relation for any value of t and x', provided that at the origin of k, τ = 0, when t = 0.
To check whether the requirement mentioned at the start is met with, let us work out the values of τ_0, τ_1 and τ_2. On substituting the parameters of τ, at the three instants, from the starting equation, we get the values as follows.

$$\tau_0=a\left(t-\frac{v}{c^2-v^2}\times 0\right)=at$$

$$\tau_1=a\left(\left(t+\frac{x'}{c-v}\right)-\frac{v}{c^2-v^2}x'\right)=a\left(t+\frac{c}{c^2-v^2}x'\right)$$

$$\tau_2=a\left(\left(t+\frac{x'}{c-v}+\frac{x'}{c+v}\right)-\frac{v}{c^2-v^2}\times 0\right)=a\left(t+\frac{2c}{c^2-v^2}x'\right)$$

Therefore,

$$\tau_0 + \tau_2 = at + a\left(t + \frac{2c}{c^2 - v^2}x'\right) - 2a\left(t + \frac{c}{c^2 - v^2}x'\right) = 2\tau_1$$

Thus the relation meets the stipulated requirement.

It is pointed out here that though the relation meets the assumptions it is derived from, it fails to get the Lorentz transformation, when correct values are inputted to it, as shown in my Observations 3 and 4 ahead. The reason is a mistake which is discussed at (ii) below.

ii) The statement that the velocity of light in Y and Z directions, when viewed from the stationary system, would always be $\sqrt{c^2 - v^2}$ is incorrect, as explained below.

First of all, it is clarified that at this stage, the value of relative velocity of light in Y or Z direction is to be taken **by kinematics and not relativity,** which is also obvious from the terms of time traversed by light moving in X direction, in the starting equation itself, such as $(t + \frac{x\prime}{c-v} + \frac{x\prime}{c+v})$ and $(t + \frac{x\prime}{c-v})$. Thus the relative velocity of light with respect to the moving frame can be both, less as well as more than c by kinematics.

Following the same principles, the relative velocity of light with respect to the origin of frame moving in X direction, as observed from the stationary frame, would always be $\sqrt{c^2 + v^2}$ (by kinematics) in the diagonal direction in XY plane or XZ plane respectively, as the two velocities, i.e. c and v, are mutually perpendicular to each other. However, the same velocity $\sqrt{c^2 + v^2}$ (in the diagonal direction) would appear as velocity c in the Y and Z directions (as the case may be), and as velocity v in the X direction, as its components.

To make it clearer, if we consider c and v as vectors, the relative velocity of light with respect to that of the moving frame, as seen from the stationary frame, is $(\vec{c} - \vec{v})$. When the two are collinear, as along X direction, these are simply added $(c + v)$ or subtracted $(c - v)$, as also done by the author at start itself. On the other hand, when the two velocities are perpendicular to each other, with the light ray moving either in the Y direction or in the Z direction, the magnitude of the aforementioned relative velocity becomes $\sqrt{c^2 + v^2}$, irrespective of whether $\vec{v}$ is in (+)ve direction or in (-)ve direction. If, however, the relative velocity was sought only in the Y direction or only in the Z direction, it would always be c, as there is no component of velocity v in these directions.

Even if one argues that Einstein took the velocity of light in the diagonal direction as c relativistically (though it is not to be done at this stage), the velocity of light in in Y or Z direction would still continue to be c by the same principle of relativity. Thus the figure of $\sqrt{c^2 - v^2}$ worked out by Einstein finds no place, either relativistically or non-relativistically.

Although the correction does not alter the relations mentioned ahead i.e. $\frac{\partial \tau}{\partial y} = 0$, $\frac{\partial \tau}{\partial z} = 0$ yet the derivation ahead fails to get Lorentz transformation, as will be seen shortly.

With the help of this result we easily determine the quantities ξ, η, ζ by expressing in equations that light (as required by the principle of the constancy of the velocity of light, in combination with the principle of relativity) is also propagated with velocity c when measured in the moving system. For a ray of light emitted at the time τ = 0 in the direction of the increasing ξ

$$\xi = c\tau \text{ or } \xi = ac\left(t - \frac{v}{c^2-v^2}x'\right)$$

But the ray moves relatively to the initial point of k, when measured in the stationary system, with the velocity $c - v$, so that

$$\frac{x'}{c-v} = t$$

If we insert this value of t in the equation for ξ, we obtain

$$\xi = a\frac{c^2}{c^2-v^2}x'$$

Observations 2:

As already mentioned in Observations 1 (i) above, the expression of τ, as a function of t and x', is of a general nature, given the theory of transformation proposed. However, selecting t as the time for light to travel a distance of x' curtails the applicability of the resultant relation only to events of light.

Despite such a limiting condition, Einstein used the relations extensively on all space–time sets, the first being on a rigid sphere in the next section 4 titled "Physical Meaning of the Equations Obtained in Respect to Moving Rigid Bodies and Moving Clocks". The deviation is also practiced universally.

The above issue (of deviation from assumptions of derivation), however, gets relegated behind, when one discovers a mistake, which is already explained in my Observations 1 (ii) above.

In an analogous manner we find, by considering rays moving along the two other axes, that

$$\eta = c\tau = ac\left(t - \frac{v}{c^2-v^2}x'\right)$$

when

$$\frac{y}{\sqrt{c^2-v^2}} = t, x' = 0$$

Thus

$$\eta = a\frac{c}{\sqrt{c^2-v^2}}y \text{ and } \zeta = a\frac{c}{\sqrt{c^2-v^2}}z$$

Substituting for x' its value, we obtain

$$\tau = \emptyset(v)\beta(t - vx/c^2),$$

$$\xi = \emptyset(v)\beta(x - vt),$$

$$\eta = \emptyset(v)y,$$

$$\zeta = \emptyset(v)z,$$

where

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

Observations 3:

It has been explained in my Observations 1 (ii) above that the velocity of light in Y and Z directions is incorrectly taken as $\sqrt{c^2 - v^2}$, and it should correctly be c from simple kinematics. On applying this correction by replacing $\sqrt{c^2 - v^2}$ with c, the expressions of η and ζ become as follows.

$$\eta = a\frac{c}{c}y = ay \text{ and } \zeta = a\frac{c}{c}z = az$$

Further, in the above relations, the function $\emptyset(v)$ has been taken as follows.

$$\emptyset(v) = a\frac{c}{\sqrt{c^2 - v^2}} = a\beta$$

When the above correction is applied, one gets $\emptyset(v) = a\frac{c}{c} = a$.

Therefore, to correct the above-mentioned 4 relations, one has to multiply the RHS by β. On doing so, one gets,

$$\tau = \emptyset(v)\beta^2(t - vx/c^2),$$

$$\xi = \emptyset(v)\beta^2(x - vt),$$

$$\eta = \emptyset(v)\beta y,$$

$$\zeta = \emptyset(v)\beta z,$$

Notes:

The derivation goes on further to find out the value of $\emptyset(v)$ which turns out to be 1.

Observations 4:

When the calculated value of $\emptyset(v)$ as 1 is substituted, the relations become as follows.

$$\tau = \beta^2(t - vx/c^2),$$

$$\xi = \beta^2(x - vt),$$

$$\eta = \beta y,$$

$$\zeta = \beta z,$$

These are different from the Lorentz transformation relations and also cannot be the targeted relations, as these would not conform to constancy of light speed c in the moving frame.

Thus the assumptions made for transformation of coordinates and times fail to get the Lorentz transformation.

Does it mean that the Lorentz Transformation are Incorrect?

No. The above exercise is only a failed attempt to derive the Lorentz transformation for light by kinematics, which has otherwise been established by electrodynamics. However, these relations for events other than those of light are not yet derived, though used universally. The same Lorentz transformation relations can be derived for such events too, but with a different postulate/assumption.

Alternative Methods:

1. Einstein's 1916 Book:

Annexure I of the book also presents a derivation by kinematics. However, the derivation starts with building up equations (1) to (5) for a moving light signal, but soon thereafter digresses to apply the set of equations (5) on other-than-light objects such as origin of the moving frame, meter-rods and clocks placed in the stationary and moving frames etc.
Thus, this derivation too is not correct, and therefore, is unsuitable for adoption.

2. A New Derivation Based on Reciprocity of Velocity:

Two of Einstein's derivations of the Lorentz transformation relations by kinematics are found to be incorrect, as explained above. On the other hand, by application of the principle of velocity reciprocity between the two frames, it is possible to achieve the relations by kinematics.

The same is presented below [2].

The following fig.1 may be referred.

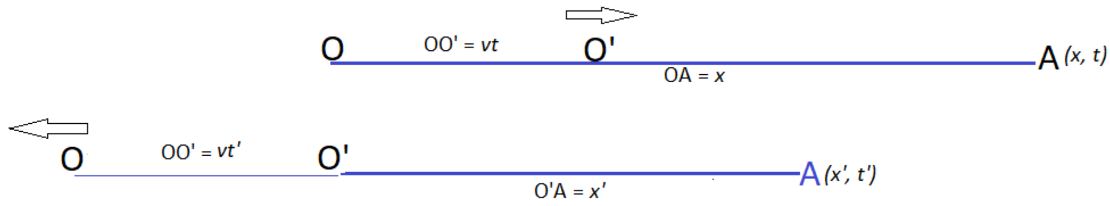

Fig.1

The point A is the event which may be of a moving light signal or any arbitrary distance-time set. So, it is shown merely as (x, t) in the non-primed frame and as (x', t') in the primed frame. In case of light, the two parameters of the event are related as $(x = ct)$ and $(x' = ct')$ respectively.

The points O and O' are the locations of origins of the non-primed and the primed frames respectively, which are in relative motion with respect to each other with a uniform relative velocity v in the direction of location of the event A from origin O i.e. v is along the line OO'A. At time $t = t' = 0$, both the origins were coincident, and at the same time, one of the frames starts its motion. Either of O and O' could be taken as moving toward A, along the line OO'A. In the instant case, however, O' is considered to be so.

Both the diagrams show the locations of O, O' and A at a particular time t after the start.

The top one represents O as stationary and O' as moving, with the observer at O watching O' and A.

Similarly, the bottom one represents O' as stationary and O as moving in the opposite direction, with the observer at O' watching O and A.

x is the distance of the event A at time t in the non-primed frame, and similarly, x' is the corresponding distance of the event A at time t' in the primed frame.

Both the representations are equally correct and interchangeable, by the essence of relativity i.e. reciprocity.

Applying classical kinematics, one may write the following relations for the two cases respectively.

$$\left.\begin{aligned} \mathrm{O'A} = x' = x - vt \\ \mathrm{OA} = x = x' + vt' \end{aligned}\right\}$$

However, according to the Special Relativity, the distance and time in the moving frame are so modified that these conform to the postulate of constancy of light speed (in vacuum) in all inertial frames, and/or to the Lorentz Transformation Condition.

Therefore, a bridging parameter, in the form of a constant, need to be applied to both the relations to strike conformity.

Let a be such a constant to be applied to both, as follows.

$$\left.\begin{aligned} x' = a(x - vt) \\ x = a(x' + vt') \end{aligned}\right\} \ldots\ldots (1)$$

On separating t' and t, one-by-one, from the above two relations (1), one gets the following.

$$\left.\begin{aligned} t' &= a\left[t - \left(1 - \frac{1}{a^2}\right)\frac{x}{v}\right] \\ t &= a\left[t' + \left(1 - \frac{1}{a^2}\right)\frac{x'}{v}\right] \end{aligned}\right\} \ldots\ldots (2)$$

Since both the frames are inertial, the postulate stipulates that $x = ct$ as well as $x' = ct'$.

Using the above relations (1) and (2), in conjunction with the postulate, separately in the two frames i.e. $x = ct$ and $x' = ct'$, let us proceed as follows, on two separate threads (columns of the following table).

$x' = ct'$	$x = ct$
Substitute the expressions of x' and t' from relations (1) and (2) above to get $$a(x - vt) = c\left(a\left[t - \left(1 - \frac{1}{a^2}\right)\frac{x}{v}\right]\right)$$	Substitute the expressions of x and t from relations (1) and (2) above to get $$a(x' + vt') = c\left(a\left[t' + \left(1 - \frac{1}{a^2}\right)\frac{x'}{v}\right]\right)$$
Substitute x with ct to get $$a(ct - vt) = c\left(a\left[t - \left(1 - \frac{1}{a^2}\right)\frac{ct}{v}\right]\right)$$	Substitute x' with ct' to get $$a(ct' + vt') = c\left(a\left[t' + \left(1 - \frac{1}{a^2}\right)\frac{ct'}{v}\right]\right)$$
Divide both the sides of eq. by ct to get $$a\left(1 - \frac{v}{c}\right) = a\left[1 - \left(1 - \frac{1}{a^2}\right)\frac{c}{v}\right]$$	Divide both the sides of eq. by ct' to get $$a\left(1 + \frac{v}{c}\right) = a\left[1 + \left(1 - \frac{1}{a^2}\right)\frac{c}{v}\right]$$
Or, $$1 - \frac{v}{c} = 1 - \left(1 - \frac{1}{a^2}\right)\frac{c}{v}$$	Or, $$1 + \frac{v}{c} = 1 + \left(1 - \frac{1}{a^2}\right)\frac{c}{v}$$
Or, $$a = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \gamma$$	$$a = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \gamma$$
Substitute the above value of a in relations (1) and (2), and replace a with γ to get $$\left.\begin{aligned} x' &= \gamma(x - vt) \\ t' &= \gamma\left(t - \frac{vx}{c^2}\right) \end{aligned}\right\}$$ $$\left.\begin{aligned} x &= \gamma(x' + vt') \\ t &= \gamma\left(t' + \frac{vx'}{c^2}\right) \end{aligned}\right\}$$	Substitute the above value of a in relations (1) and (2), and replace a with γ to get $$\left.\begin{aligned} x' &= \gamma(x - vt) \\ t' &= \gamma\left(t - \frac{vx}{c^2}\right) \end{aligned}\right\}$$ $$\left.\begin{aligned} x &= \gamma(x' + vt') \\ t &= \gamma\left(t' + \frac{vx'}{c^2}\right) \end{aligned}\right\}$$

Thus the Lorentz transformation is achieved for both the frames separately, by the new method.

Conclusion:

The article shows as to how Einstein, in his 1905 paper, derived the Lorentz transformation by kinematics, with a mistake which has remained unnoticed by the Physics community.

His next attempt in his 1916 book is also riddled with mistakes, which have been pointed out briefly in the article.
The mistakes, however, do not undermine the validity of the Lorentz transformation, which has been established beyond doubt for light, by electrodynamics.
The essence of relativity i.e. reciprocity of the relative velocity between the two frames has been shown to be strong enough to throw out the Lorentz transformation by kinematics, in both the frames individually.

References:

1. A. Einstein, "On the Electrodynamics of Moving Bodies", June 30, 1905, distributed by http://www.fourmilab.ch/
2. Author's Book "Refining Relativity Part 1 (The Special Theory)", 2020.