

The simplest derivation of the Lorentz transformation

J.-M. Levy

► **To cite this version:**

| J.-M. Levy. The simplest derivation of the Lorentz transformation. 2006. in2p3-00079457v4

HAL Id: in2p3-00079457

<http://hal.in2p3.fr/in2p3-00079457v4>

Preprint submitted on 9 Oct 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The simplest derivation of the Lorentz transformation

J.-M. Lévy

Laboratoire de Physique Nucléaire et de Hautes Energies, CNRS - IN2P3 - Universités Paris VI et Paris VII, Paris.

Email: jmlevy@in2p3.fr

Abstract

The Lorentz transformation is derived from the simplest thought experiment by using the simplest vector formula from elementary geometry. The result is further used to obtain general velocity and acceleration transformation equations.

1 Introduction

The light clock is a much used conceptual device to demonstrate time dilation and length contraction in introductory courses to special relativity (see e.g. [1]; [2] gives an extensive bibliography) Less well known is the fact that the Lorentz transformation can be entirely derived once these effects have been established. A paper in which this was shown appeared in the American Journal of Physics a long time ago [3]. However, the author of [3] missed what we think is the easiest way to derive the time transformation formula and was led to obtain it through a rather contrived argument, introducing an artificial extension of the 'time interval'. Also, as in most papers on the subject, the derivation was limited to transformations between two reference frames in the so-called 'standard configuration' [4], viz. parallel axes, OX' sliding along OX with coincident space-time origins.

The purpose of the present paper is to show that the full transformation can be derived from a purely geometrical argument which amounts to writing the basic vector addition formula in the two frames at stake successively, taking into account the length contraction effect.

It is further shown that the same reasoning yields the transformation for an arbitrary velocity between two parallel frames with very little extra effort. In passing, a simple formula is derived for the space part of the transformation.

To make this paper self-contained and also to prevent objections which are often not taken care of in the derivation of the two basic effects using the light-clock, we shall start with a brief review of this derivation in section 2. To go straight to the heart of the argument, we first derive the Lorentz transformation from length contraction between two frames in 'standard configuration' in section 3. Section 4 will then be devoted to the more general case of an arbitrarily oriented relative velocity. Transposing the demonstration to this

more general case forces us to write the transformation in a slightly unusual form which yields the above mentioned formula as a by-product. This allows us to derive very simply general expressions for the velocity and acceleration transformations in Section 5. Section 6 contains our summary and conclusions.

2 Time dilation and length contraction

2.1 The light clock

For those who first discover it, the light clock is a magically simple conceptual device to demonstrate the basic effects of special relativity starting from Einstein's two postulates. Admittedly, the second postulate viz. the constancy of the velocity of light in all inertial frames can be dispensed with through general considerations [5] if the existence of a limiting invariant velocity is established. The only other choice would be a possibly limitless relative velocity and galilean invariance. However, Maxwell's equations are there to settle the matter.

We think nevertheless that for an introductory course, the second postulate should be retained in as much as it allows the students to arrive more quickly at the heart of the matter without leaving the more practically minded ones stranded. They will have plenty of time to assimilate and appreciate the value of general symmetry arguments later in their cursus by getting impregnated with them progressively.

2.1.1 Time dilation

We imagine the following device: a light signal bounces back and forth between two parallel mirrors maintained at a constant separation with the aid of pegs. The signal triggers the registering of a tick each time it hits what we define as the 'lower' mirror (fig.1 a). The question of how this device can be practically constructed

does not concern us. We simply assume that there is a way to sample the signal in order to produce the tick and to compensate for the loss of light incurred thereof. We thus have a kind of perfect clock the period of which is $T_0 = \frac{2L_0}{c}$ with L_0 the distance between the mirrors and c the speed of light.

Let's now look at the clock in a frame wherein it travels at constant speed v in a direction parallel to the planes of the mirrors. To prevent an objection, we might assume that the mirrors are constrained to slide in two parallel straight grooves which have been engraved a constant distance L_0 apart previous to the experiment, so that there can't be any arguing about a variation of the mirror separation when they are moving.

By the first postulate, this moving clock must have the same period in its rest frame than its twin at rest in the laboratory.

On the other hand, it is obvious that the length traveled by the signal in the observer frame is longer than the length in the clock rest frame (see fig.1 b.) If T is the interval between two ticks in the observer frame, then by Einstein's second postulate and Pythagora's theorem we have that $(cT/2)^2 = L_0^2 + (vT/2)^2$ from which $T = \frac{T_0}{\sqrt{1-(\frac{v}{c})^2}}$ follows, which shows that the moving clock runs more slowly in the lab frame.

2.1.2 Length contraction

We now imagine that the moving clock is traveling in a direction perpendicular to the plane of its mirrors relative to us. In this case, no check can be kept of the inter-mirror distance. To make sure that (for the same v) the clock period hasn't changed, we can imagine an observer traveling with the clock and provided with an identical second clock oriented as before, parallel to the velocity with respect to us. Since this second clock has period T , as above, in our frame and since the observer can reassure us that the two clocks tick at the same rate in his frame, we can be sure that the first clock period as measured in our frame hasn't changed either. Anticipating the result which will be forced upon us, we call L the inter-mirror distance as measured in our frame. If we now consider the time taken by the light signal to make its two-way travel in our frame, we see that it needs $\frac{L}{c+v}$ for the lower mirror-upper mirror part and $\frac{L}{c-v}$ for the return part. Since the total must equal T , one is forced to conclude that $L = L_0 \sqrt{1 - (\frac{v}{c})^2}$

This ends our review of the basic light-clock experiment. That the distances in the directions orthogonal to the motion are not changed can be demonstrated by invoking grooves arguments like

the one we used for the time-dilation derivation.

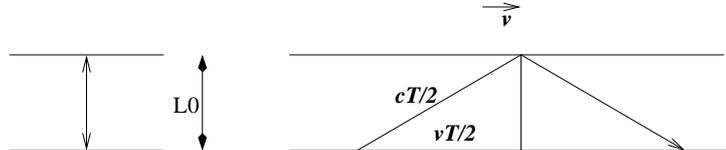


fig 1.a

fig 1.b

The light clock at rest (left) and moving(right)

3 Lorentz transformation along the x axis

Let us now envision two frames in 'standard configuration' with K' having velocity \mathbf{v} with respect to K and let x, t (resp. x', t') be the coordinates of event M in the two frames. Let O and O' be the spatial origins of the frames; O and O' coincide at time $t = t' = 0$

Here comes the pretty argument: all we have to do is to express the relation $\mathbf{OM} = \mathbf{OO'} + \mathbf{O'M}$ between vectors (which here reduce to oriented segments) in both frames.

In K , $\overline{OM} = x$, $\overline{OO'} = vt$ and $\overline{O'M}$ seen from K is $\frac{x'}{\gamma}$ with $\gamma = \frac{1}{\sqrt{1-(\frac{v}{c})^2}}$ since x' is $\overline{O'M}$ as measured in K' Hence the first relation:

$$x = vt + \frac{x'}{\gamma}$$

In K' , $\overline{OM} = \frac{x}{\gamma}$ since x is \overline{OM} as measured in K , $\overline{OO'} = vt'$ and $\overline{O'M} = x'$. Hence a second relation:

$$\frac{x}{\gamma} = vt' + x'$$

The first relation yields immediately

$$x' = \gamma(x - vt)$$

which is the 'space' part of the Lorentz transformation and the second relation yields the inverse

$$x = \gamma(x' + vt')$$

of this 'space part'. Eliminating x' between these two leads quickly to the formula for the transformed time:

$$t' = \gamma(t - vx/c^2)$$

the inverse of which could easily be found by a similar elimination of x .

Needless to say, coordinates on the y and z axes are

unchanged for the already stated reason that distances do not vary in the directions perpendicular to the velocity. The contraction is therefore limited to that part of the coordinate vector which is parallel to the relative velocity.

4 The case of an arbitrary velocity

In the following, \mathbf{v} will denote the velocity vector of K' w.r.t. K and \mathbf{r} (resp. \mathbf{r}') the position vector of the event under consideration as measured in frame K (resp K'). We further define $\mathbf{u} = \frac{\mathbf{v}}{|\mathbf{v}|}$ the unit vector parallel to \mathbf{v} .

From our findings of section 2, we see that only the component of \mathbf{r} parallel to \mathbf{v} is affected when looking at it from the other frame, whilst the normal components are unchanged. We resolve \mathbf{r} into parallel and perpendicular components according to $\mathbf{r} = \mathbf{u}\mathbf{u}\cdot\mathbf{r} + (\mathbf{1} - \mathbf{u} \otimes \mathbf{u})\mathbf{r} = \mathbf{r}_{\parallel} + \mathbf{r}_{\perp}$ where the dot stands for the 3-space scalar product, $\mathbf{1}$ is the identity operator and $\mathbf{u} \otimes \mathbf{u}$ is the dyadic which projects out the component parallel to \mathbf{u} from the vector it operates upon, viz $(\mathbf{u} \otimes \mathbf{u})\mathbf{V} = (\mathbf{u}\cdot\mathbf{V})\mathbf{u}$.

The operator which contracts the projection on \mathbf{u} by γ whilst leaving the orthogonal components unchanged must yield: $\mathbf{u}\frac{\mathbf{u}\cdot\mathbf{r}}{\gamma} + (\mathbf{1} - \mathbf{u} \otimes \mathbf{u})\mathbf{r} = (\mathbf{1} + \frac{1-\gamma}{\gamma}\mathbf{u} \otimes \mathbf{u})\mathbf{r}$. Let us therefore define $\mathbf{Op}(\gamma^{-1}) = \mathbf{1} + \frac{1-\gamma}{\gamma}\mathbf{u} \otimes \mathbf{u}$. The inverse operator must correspond to multiplication of the longitudinal part by γ and is therefore $\mathbf{Op}(\gamma) = \mathbf{Op}(\gamma^{-1})^{-1} = \mathbf{1} + (\gamma - 1)\mathbf{u} \otimes \mathbf{u}$ as can also be checked by multiplication. Note that these operators are even in \mathbf{u} and therefore independent of the orientation of \mathbf{v} .

Mimicking what has been done in section 3, let us now write $\mathbf{OM} = \mathbf{OO}' + \mathbf{O}'\mathbf{M}$ (these are vectors now, no longer oriented segments) taking care of the invariance of the orthogonal parts. We get in frame K :

$$\mathbf{r} = \mathbf{vt} + \mathbf{Op}(\gamma^{-1})\mathbf{r}'$$

and in frame K' :

$$\mathbf{Op}(\gamma^{-1})\mathbf{r} = \mathbf{vt}' + \mathbf{r}'$$

Using the inverse operator the first relation yields immediately:

$$\mathbf{r}' = \mathbf{Op}(\gamma)(\mathbf{r} - \mathbf{vt}) = (\mathbf{1} + (\gamma - 1)\mathbf{u} \otimes \mathbf{u})(\mathbf{r} - \mathbf{vt})$$

which is probably the simplest way to write the space part of the rotation free homogenous Lorentz transformation. The usual γ factor of the one dimensional transformation is simply replaced by the operator

$\mathbf{Op}(\gamma)$

By feeding this result into the second relation above, we find:

$$\mathbf{Op}(\gamma^{-1})\mathbf{r} = \mathbf{vt}' + \mathbf{Op}(\gamma)(\mathbf{r} - \mathbf{vt})$$

or, using $\mathbf{Op}(\gamma)\mathbf{v} = \gamma\mathbf{v}$ and with the explicit form of \mathbf{Op} :

$$(\frac{1-\gamma}{\gamma} - (\gamma - 1))\mathbf{r} + \gamma\mathbf{vt} = \mathbf{vt}'$$

which, using $1 - \gamma^2 = -(\frac{v}{c})^2\gamma^2$ and crossing away \mathbf{v} on both sides yields:

$$t' = \gamma(t - \frac{\mathbf{v}\cdot\mathbf{r}}{c^2})$$

i.e. the time transformation equation.

5 Velocity and acceleration transformations

5.1 Velocity

The two formulas thus obtained for the L.T. are so simple that they can readily be used to yield the velocity transformation equation without the need of complicated thought experiments and algebraic manipulations. Differentiating \mathbf{r}' and t' w.r.t. t and taking the quotient of the equalities thus obtained yields (with $\mathbf{V}' = \frac{d\mathbf{r}'}{dt'}$ and $\mathbf{V} = \frac{d\mathbf{r}}{dt}$)

$$\mathbf{V}' = \frac{1}{\gamma} \frac{(1 + (\gamma - 1)\mathbf{u} \otimes \mathbf{u})(\mathbf{V} - \mathbf{v})}{1 - \frac{\mathbf{v}\cdot\mathbf{V}}{c^2}}$$

which is the general velocity transformation formula.

5.2 Acceleration

Using the compact \mathbf{Op} notation helps keeping things tidy when differentiating \mathbf{V}' ; dividing by the differential of t' one finds

$$\mathbf{A}' = \frac{1}{\gamma^2} \frac{\mathbf{Op}(\gamma)\mathbf{A}(1 - \frac{\mathbf{v}\cdot\mathbf{V}}{c^2}) + \mathbf{Op}(\gamma)(\mathbf{V} - \mathbf{v})\frac{\mathbf{v}\cdot\mathbf{A}}{c^2}}{(1 - \frac{\mathbf{v}\cdot\mathbf{V}}{c^2})^3}$$

Expliciting \mathbf{Op} , simplifying and regrouping terms, one obtains after a page of algebra:

$$\mathbf{A}' = \frac{\mathbf{A} - \frac{\gamma}{\gamma+1}\frac{\mathbf{v}\cdot\mathbf{A}\mathbf{v}}{c^2} + \frac{\mathbf{v}\wedge(\mathbf{V}\wedge\mathbf{A})}{c^2}}{\gamma^2(1 - \frac{\mathbf{v}\cdot\mathbf{V}}{c^2})^3}$$

By making the necessary substitutions: $\mathbf{V} \rightarrow \mathbf{u}'$, $\mathbf{V}' \rightarrow \mathbf{u}$, $\mathbf{v} \rightarrow -\mathbf{V}$ and specializing to \mathbf{V} parallel to Ox ,

one can easily check that the components equations derived from this general formula agree with those published in [2]. They have been, however, derived with much less effort.

As an example of use of this acceleration transformation, by specializing to $\mathbf{V} = \mathbf{v}$ and $\mathbf{v} \cdot \mathbf{A} = 0$, one gets $\mathbf{A}' = \gamma^2 \mathbf{A}$ retrieving the known result that a particle in a circular storage ring undergoes a proper (\mathbf{A}') acceleration that is a factor γ^2 larger than the lab (\mathbf{A}) acceleration. Moreover, the two accelerations are parallel, which is far from obvious a priori. Observe in this respect, that all the terms which can make A' and A different in direction as well as in size vanish in the $c \rightarrow \infty$ limit, consistent with the fact that acceleration is an invariant quantity under a change of inertial frame in newtonian physics.

6 Summary and conclusion

We have shown that the general rotation free homogenous Lorentz transformation can be derived once length contraction has been established by writing the elementary vector relation (sometimes dubbed

'Chasles' relation) $\mathbf{OM} = \mathbf{OO}' + \mathbf{O}'\mathbf{M}$ in the two frames considered [6]. The extension from the special one dimensional case to the 3-dimensional case is completely straightforward. The formula thus obtained allows for a simple derivation of the velocity and acceleration transformations without the need for complicated thought experiments and algebraic manipulations beyond what college students are used to.

References

- [1] Kenneth Krane, *Modern Physics*, John Wiley & sons 1983 p. 23
- [2] W.N. Mathews Jr. Am. J. Phys. **73**, 45-51 (2005)
- [3] David Park, Am. J. Phys. **42**, 909-910, (1974)
- [4] Wolfgang Rindler, *Relativity, Special, General and Cosmological* Oxford U.P. 2001, p. 5
- [5] See e.g. ref. [4] p. 57
- [6] There are much faster derivations than the one presented here. See e.g. Alan Macdonald, Am. J. Phys. **49**, p. 493, (1981). Our purpose here was to be simple, as stated in the title and abstract, and to go beyond the 'standard configuration'