

Getting the Lorentz transformations without requiring an invariant speed

Andrea Pelissetto^{a)} and Massimo Testa^{b)}

Dipartimento di Fisica, Università degli Studi di Roma “La Sapienza” and INFN – Sezione di Roma I Piazzale A. Moro 2, I-00185 Roma, Italy

(Received 15 July 2014; accepted 31 October 2014)

The structure of the Lorentz transformations follows purely from the absence of privileged inertial reference frames and the group structure (closure under composition) of the transformations—two assumptions that are simple and physically necessary. The existence of an invariant speed is *not* a necessary assumption and in fact is a consequence of the principle of relativity (though the finite value of this speed must, of course, be obtained from experiment). Von Ignatowsky derived this result in 1911, but it is still not widely known and is absent from most textbooks. Here, we present a completely elementary proof of the result, suitable for use in an introductory course in special relativity. © 2015 American Association of Physics Teachers.

[<http://dx.doi.org/10.1119/1.4901453>]

I. INTRODUCTION

In standard textbooks, the Lorentz transformation equations, which connect inertial reference frames, are deduced from the invariance of the speed of light, which implies the invariance of the Minkowski interval.¹ A different approach, however, was followed by von Ignatowsky.² Only six years after the formulation of special relativity, he proved that the Lorentz transformations arise under quite general conditions, without assuming *a priori* the existence of an invariant speed. Von Ignatowsky showed that the only admissible transformations consistent with the principle of inertia, the isotropy of space, the absence of preferred inertial frames, and a group structure (i.e., closure under composition), are the Lorentz transformations, in which c can be any velocity scale, or the Galilean transformations. This surprising result shows that the Lorentz transformations are not directly related to the properties of electromagnetic radiation. Electromagnetism is only relevant, if present within the theory, as a way to fix the arbitrary velocity scale, which is then identified with the speed of light. This deep and fascinating result, although well known in the specialized literature,^{3–11} is not commonly found in textbooks,¹² because its usual proof is rather complicated and uninspiring. It is the purpose of the present paper to present a derivation of the von Ignatowsky result using elementary considerations, which, in our opinion, sheds new light on the result.

II. NOTATION FOR TRANSFORMATIONS BETWEEN FRAMES

We wish to characterize the transformations that relate two different inertial frames. Let us consider two inertial observers \mathcal{O} and \mathcal{O}' . Let $\mathbf{r} = (x, y, z)$ and t be space and time coordinates for \mathcal{O} and $\mathbf{r}' = (x', y', z')$ and t' be the corresponding quantities for \mathcal{O}' . In order to simplify the argument, we will restrict our considerations to the subgroup of transformations involving x and t only, setting $y' = y$ and $z' = z$. This is equivalent to choosing coordinates so that \mathcal{O} and \mathcal{O}' are in relative motion along the x and x' direction (see Fig. 1).¹³

We assume the validity of the principle of inertia: in an inertial frame free particles undergo rectilinear motion with constant speed. Therefore, if the trajectory of a particle is a

straight line in the frame of observer \mathcal{O} and its speed is constant, the trajectory is also a straight line in the frame of observer \mathcal{O}' and also the speed in the new frame is constant. This condition implies¹⁴ that the two inertial frames are related by a *linear* transformation. We can therefore write

$$x'^0 = Ax^0 + Bx, \quad (1)$$

$$x' = Cx^0 + Dx, \quad (2)$$

where we have defined the time coordinates x^0 and x'^0 , measured in units of distance, as

$$x^0 = \bar{v}t, \quad (3)$$

where \bar{v} is an *arbitrary* constant with units of speed. The inverse transformation follows immediately:

$$x^0 = \frac{1}{\Delta}(Dx'^0 - Bx'), \quad (4)$$

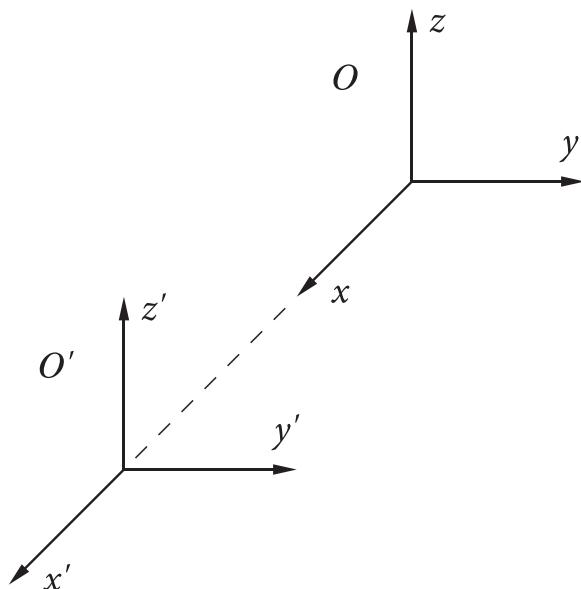


Fig. 1. Two inertial frames in relative motion along the x -axis.

$$x = \frac{1}{\Delta} (-Cx^{0'} + Ax'), \quad (5)$$

with

$$\Delta \equiv AD - BC \neq 0. \quad (6)$$

It is useful to rewrite the transformation in matrix notation, introducing

$$\Lambda \equiv \begin{pmatrix} A & B \\ C & D \end{pmatrix} \quad (7)$$

and its inverse

$$\Lambda^{-1} = \frac{1}{\Delta} \begin{pmatrix} D & -B \\ -C & A \end{pmatrix}. \quad (8)$$

Then, we have

$$\begin{pmatrix} x^{0'} \\ x' \end{pmatrix} = \Lambda \begin{pmatrix} x^0 \\ x \end{pmatrix} \quad (9)$$

and

$$\begin{pmatrix} x^0 \\ x \end{pmatrix} = \Lambda^{-1} \begin{pmatrix} x^{0'} \\ x' \end{pmatrix}. \quad (10)$$

III. EQUIVALENCE OF INERTIAL REFERENCE FRAMES

The main ingredient in the proof is the requirement that there are no privileged frames: *all inertial frames are equivalent*. We wish now to express this hypothesis in a more transparent way that allows us to put constraints on the matrix Λ .

Let us consider two events, E_1 and E_2 , at the same spatial location in frame \mathcal{O} , but separated by a time difference τ . In \mathcal{O}' the two events are separated by a time lapse T' . One natural requirement is that if we consider two different events, E_3 and E_4 , that are at the same spatial location in frame \mathcal{O}' , but separated by the same time difference τ in \mathcal{O}' , then these two events are separated by a time lapse $T = T'$ in \mathcal{O} . Similarly, we require that if observer \mathcal{O} measures the length l (along the x -axis) of a rod that is at rest with respect to \mathcal{O}' and has length l_0 in the \mathcal{O}' frame, the same result $l' = l$ is obtained by \mathcal{O}' for an identical rod at rest with respect to \mathcal{O} . These requirements constrain the coefficients of the transformation Λ .

We first consider the two events E_1 and E_2 at the same spatial location in frame \mathcal{O} , separated in time by τ . From Eq. (1), observer \mathcal{O}' measures a time lapse T' between these events given by

$$T' = A\tau. \quad (11)$$

Consider now the complementary situation in which an identical clock, at rest with respect to \mathcal{O}' , measures the same time interval τ between two events E_3 and E_4 that are at the same location in \mathcal{O}' . Correspondingly, from Eq. (4), the observer \mathcal{O} measures a time lapse T given by

$$T = \frac{D}{\Delta} \tau. \quad (12)$$

As stated above, if neither \mathcal{O} nor \mathcal{O}' is in some way privileged, the two time intervals T and T' must be identical, which implies

$$A = \frac{D}{\Delta}. \quad (13)$$

A similar argument applies in the case of a rod of rest length l_0 , oriented along the x -axis. If the rod is at rest with respect to \mathcal{O}' , Eq. (2) implies that the length l measured by observer \mathcal{O} satisfies

$$l_0 = Dl. \quad (14)$$

Analogously, for a rod at rest in \mathcal{O} , we have, from Eq. (5), that the length l' measured by \mathcal{O}' satisfies

$$l_0 = \frac{A}{\Delta} l'. \quad (15)$$

Again, the absence of privileged reference frames requires that $l = l'$ and therefore

$$\frac{A}{\Delta} = D. \quad (16)$$

Combining Eqs. (13) and (16), we have

$$\frac{A}{\Delta^2} = A. \quad (17)$$

The solution $A = 0$ is not acceptable on physical grounds, because it would lead to the meaningless result $x^{0'} = Bx$, so we are forced to choose

$$|\Delta| = 1. \quad (18)$$

If we restrict ourselves to proper transformations, we have simply

$$\Delta = 1, \quad (19)$$

so that

$$A = D. \quad (20)$$

These conditions imply that the transformation relating two different inertial frames is of the form

$$\Lambda = \begin{pmatrix} A & B \\ C & A \end{pmatrix}, \quad (21)$$

with

$$A^2 - BC = 1, \quad (22)$$

the latter condition being a consequence of Eq. (19).

IV. GROUP STRUCTURE

To further constrain the structure of the matrix Λ , we now add the natural requirement that the transformations connecting two inertial frames constitute a group, i.e., that the combination of two such transformations yields a third transformation of the same form.

We consider two transformations,

$$\Lambda_1 \equiv \begin{pmatrix} A_1 & B_1 \\ C_1 & A_1 \end{pmatrix}, \quad \Lambda_2 \equiv \begin{pmatrix} A_2 & B_2 \\ C_2 & A_2 \end{pmatrix}, \quad (23)$$

and their product

$$\begin{aligned} \Lambda_3 = \Lambda_1 \Lambda_2 &= \begin{pmatrix} A_1 & B_1 \\ C_1 & A_1 \end{pmatrix} \begin{pmatrix} A_2 & B_2 \\ C_2 & A_2 \end{pmatrix} \\ &= \begin{pmatrix} A_1 A_2 + B_1 C_2 & A_1 B_2 + B_1 A_2 \\ C_1 A_2 + A_1 C_2 & A_1 A_2 + C_1 B_2 \end{pmatrix}. \end{aligned} \quad (24)$$

From Eq. (21), we know that the matrix Λ_3 is an element of the set of admissible matrices only if its diagonal elements are equal, that is, if

$$B_1 C_2 = C_1 B_2. \quad (25)$$

In order to satisfy Eq. (25) for *all* transformations of the form (23), we have three different possibilities:

- (i) $B = \alpha C$, where α is a nonzero constant;
- (ii) $B = 0$ and $A = 1$, where the second equation follows from Eq. (22); or
- (iii) $C = 0$ and $A = 1$.

Case (iii) is recognized to be physically uninteresting, while case (ii) corresponds to the Galilean transformations:

$$x^0' = x^0, \quad (26)$$

$$x' = C x^0 + x, \quad (27)$$

the parameter C being the relative velocity of the two frames in units of \bar{v} .

To understand case (i), we introduce a rescaled time variable to replace x^0 (recall that $x^0 = \bar{v}t$)

$$\tilde{x}^0 = \bar{c}t = \frac{\bar{c}}{\bar{v}}x^0, \quad (28)$$

with

$$\bar{c} = \frac{\bar{v}}{\sqrt{|\alpha|}}. \quad (29)$$

Our transformation then takes the form

$$\begin{pmatrix} \tilde{x}'^0 \\ x' \end{pmatrix} = \begin{pmatrix} A & \frac{\alpha}{|\alpha|} C' \\ C' & A \end{pmatrix} \begin{pmatrix} \tilde{x}^0 \\ x \end{pmatrix}, \quad (30)$$

where

$$C' = C \sqrt{|\alpha|}, \quad (31)$$

and Eq. (22) now becomes

$$A^2 - \frac{\alpha}{|\alpha|} C'^2 = 1. \quad (32)$$

If α is negative, then Λ is an orthogonal rotation matrix. In this case, there are transformations Λ such that Λ^n is close to

the identity for some value of n , and this is clearly unphysical. Therefore, we can exclude this possibility and we are left with transformations of the form

$$\Lambda = \begin{pmatrix} A & C' \\ C' & A \end{pmatrix}, \quad (33)$$

where $A^2 - C'^2 = 1$. In this case, the observer \mathcal{O}' moves with speed w with respect to observer \mathcal{O} , where w is determined by

$$C'/A = w/\bar{c}. \quad (34)$$

The condition $A^2 - C'^2 = 1$ implies

$$A = \gamma = \frac{1}{\sqrt{1 - w^2/\bar{c}^2}}, \quad (35)$$

$$C' = \gamma w/\bar{c}, \quad (36)$$

so that Λ is a generic Lorentz transformation, with the speed of light identified with \bar{c} .

V. CONCLUSIONS

The Lorentz and Galilean transformations are the only structures compatible with the principles of inertia and relativity, i.e., the nonexistence of privileged reference frames. This result, which is easily stated, is usually obtained by means of rather complicated proofs. In this paper, we have presented an elementary derivation, suitable for presentation in undergraduate courses in special relativity.

^aElectronic mail: andrea.pelissetto@roma1.infn.it

^bElectronic mail: massimo.testa@roma1.infn.it

¹J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1998).

²W. von Ignatowsky, “Das relativitätsprinzip,” *Arch. Math. Phys.* **17**, 1–24 (1911).

³H. Bacry and J. M. Lévy-Leblond, “Possible kinematics,” *J. Math. Phys.* **9**, 1605–1614 (1968).

⁴V. Berzi and V. Gorini, “Reciprocity principle and the Lorentz transformations,” *J. Math. Phys.* **10**, 1518–1524 (1969).

⁵J. M. Lévy-Leblond, “One more derivation of the Lorentz transformation,” *Am. J. Phys.* **44**, 271–277 (1976).

⁶J. M. Levy-Leblond and J. P. Provost, “Additivity, rapidity, relativity,” *Am. J. Phys.* **47**, 1045–1049 (1979).

⁷D. Mermin, “Relativity without light,” *Am. J. Phys.* **52**, 119–124 (1984).

⁸R. Torretti, *Relativity and Geometry* (Dover, New York, 1996).

⁹S. Liberati, S. Sonego, and M. Visser, “Faster-than-c signals, special relativity, and causality,” *Ann. Phys.* **298**, 167–185 (2002).

¹⁰S. Sonego and M. Pin, “Foundations of anisotropic relativistic mechanics,” *J. Math. Phys.* **50**, 042902 (2009).

¹¹V. Baccetti, K. Tate, and M. Visser, “Inertial frames without the relativity principle: breaking Lorentz invariance,” in Proceedings of the Thirteenth Marcel Grossmann Meeting, Stockholm, Sweden, 1–7 July, 2012, edited by K. Rosquist, R. T. Jantzen, and R. Ruffini (World Scientific, Singapore, 2015); Also available as e-print [arXiv:1302.5989](https://arxiv.org/abs/1302.5989) [gr-qc].

¹²For a review and a complete set of references, see S. Liberati, “Tests of Lorentz invariance: A 2013 update,” e-print [arXiv:1304.5795](https://arxiv.org/abs/1304.5795) [gr-qc].

¹³In other words, we are considering systems in motion along the x -axis: we assume that the most general transformation can be obtained by combining a boost along the x -axis with a pair of independent orthogonal transformations, i.e., rotations, in the two reference frames \mathcal{O} and \mathcal{O}' .

¹⁴L. J. Eisenberg, “Necessity of the linearity of relativistic transformations between inertial systems,” *Am. J. Phys.* **35**, 649–649 (1967).