American Journal of Physics

NOTES AND DISCUSSIONS | AUGUST 01 2024

Comment on "On the linearity of the generalized Lorentz transformation" [Am. J. Phys. 90(6), 425–429 (2022)]

Justo Pastor Lambare

Am. J. Phys. 92, 635–638 (2024)
https://doi.org/10.1119/5.0185737

Comment on "On the linearity of the generalized Lorentz transformation" [Am. J. Phys. 90(6), 425–429 (2022)]

Justo Pastor Lambare[a)]
Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Asunción, Ruta Mcal. J. F. Estigarribia, Km 11 Campus de la UNA, San Lorenzo, Paraguay

(Received 1 November 2023; accepted 26 April 2024)

In an article in this journal [Verheest, Am. J. Phys. **90**, 425–429 (2022)], Frank Verheest presented a proof for the linearity of the Lorentz transformation. We fill in some gaps in his derivation and analyze the role of the light postulate that some physicists, including Verheest, have criticized as a necessary hypothesis for formulating the theory of relativity.
https://doi.org/10.1119/5.0185737

I. INTRODUCTION

In Ref. 1, Verheest derived the linearity property of the Lorentz transformation. The reasons for Lorentz transformations must be linear often mentioned only in passing and without due rigor. Although that attitude is justified from a heuristic viewpoint, a rigorous derivation employing only elementary mathematical tools can be educationally fruitful.

Occasionally, linearity is assumed by requiring the fulfillment of the law of inertia.[2] Although linearity preserves motions with constant velocity, such preservation does not require linearity, as the following example shows.[3] Let $\mathcal{O}(x_0, x_1, x_2, x_3)$ and $\mathcal{O}'(x'_0, x'_1, x'_2, x'_3)$ be two inertial frames, whose coordinates are related through

$$x'_i = \frac{\sum_k a_{ik} x_k + e_i}{\sum_k b_k x_k + f}, \quad i, k = 0, 1, 2, 3, \tag{1}$$

where a_{ik}, b_k, e_i, f are constants and x_0 stands for time. When $x_k(\alpha) = r_k \alpha + h_k$, with α being a continuous parameter and r_k and h_k as the constants, putting $\dot{x}_k = dx_k/d\alpha$, we have

$$v_k = \frac{dx_k}{dx_0} = \frac{\dot{x}_k}{\dot{x}_0} = \frac{r_k}{r_0} = const., \quad k = 1, 2, 3, \tag{2}$$

so it describes a rectilinear motion with uniform velocity. The student can prove as an exercise that (1) and (2) implies $v'_k = \dot{x}'_k/\dot{x}'_0 = const$. Transformation (1) can be eliminated on issues of differentiability or by noting that it takes infinite coordinates values into finite ones. We have presented it only as a counterexample.[4]

On the other hand, homogeneity and isotropy of space and homogeneity of time imply that linearity is often mentioned without giving further references or comments, leaving the impression that it is a trivial implication.[5,6]

Rigorous proofs of linearity can be given using different approaches and techniques.[7–12] Some authors[8,10,11] prove linearity only from spacetime homogeneity, leaving out isotropy. Whether a detailed proof is necessary or a heuristic justification suffices, as well as the method employed, is a matter of personal taste and philosophical attitude towards a rigorous formulation of the foundations of physics.

In any case, Verheest's approach is a valid contribution for those interested in the foundations of relativity. Although his derivation is generally correct, there are some issues that we consider omissions or gaps in the reasoning rather than mistakes. We explain that problem in Sec. III. However, first, in Sec. II, we delve into another fundamental question that Verheest brought about.

In addition to the linearity issue, the author of Ref. 1 expressed a long-standing concern that some physicists have pointed out regarding the central role that the speed of light seems to play in the principles of relativity theory.[12–15] That uneasiness is justified since relativity constitutes a central pillar in the theories of modern physics. Notwithstanding the importance of electromagnetic theory, it seems odd that a particular type of phenomenon should play such a central role. In the following section, we explain that the crucial role that light purportedly plays in Einstein's formulation is only apparent and is owed to historical and practical reasons.

II. THE LIGHT PRINCIPLE

As observed in Ref. 1, Einstein based his special theory of relativity on two principles (i) the laws of physics are invariant in all inertial frames of reference and (ii) light is always propagated in empty space with a definite velocity c which is independent of the state of motion of the emitting body.[16]

Principle (i) is an extension of the equivalence of inertial reference frames from mechanics to all physical phenomena,[17] while (ii) is also known as the light principle.

It is relevant to note that (i) and (ii) together imply that the speed of light has the same value for every inertial observer. In modern literature, postulate (ii) is sometimes replaced by stating straightly that light speed is the same in all inertial frames.[18,19]

In 1905, only two fundamental interactions were known, gravitational and electromagnetic. Newtonian gravity is described by an action at a distance law, i.e., instantaneous interaction. On the other hand, light was known to be an electromagnetic phenomenon with a finite speed, while all attempts to find evidence of a light-carrying medium had failed. That historical prospect explains why Einstein gave light such a central role, notwithstanding that principle (i) encompasses all physical laws.

The tradition of teaching relativity through the light principle continues to this day. As Verheest has observed, from a conceptual viewpoint, it is more compelling to derive the Lorentz transformations without mentioning the speed of light at all. The first to do that was Vladimir Ignatowski, as early as 1910.[13] Then, many such formulations followed using different approaches and techniques.[10,12,14,15,20]

Regarding the role of light in the formulation of the theory of relativity, it is relevant to observe the following:

(a) The light principle can be replaced by the more general principle, (ii′) the principle of finiteness of the speed of propagation of interactions.[5,21]
(b) As a result of (i) and (ii′), we obtain that interactions taking place as a consequence of the fundamental physical laws, such as electromagnetic, gravitational, weak, and strong or any eventual not yet known law, must take place at the same speed which is, therefore, a universal constant. So, this approach explains why, according to relativity all fundamental interactions, and not just light (electromagnetic), occurs at speed c.

Since arbitrary interactions must occur as a consequence of the fundamental ones, the universal speed represents the upper limit for the speed transmission of any other influence. A typical example is that one hears the thunder much later than one sees the lightning, with light being a fundamental interaction while sound is not.

Principles (i) and (ii′), plus the homogeneity and isotropy assumptions, can lead us to Lorentz transformations through the usual derivations replacing light speed with a finite universal limiting speed based on the exclusion of unobservable instantaneous interactions.

Also, as done by Verheest, we can hold only to principle (i), which puts Galilean and Lorentz transformations on the same basis. Ironically, such an approach has the conceptual advantage of making more evident the essential difference between Galilean and Einstein's relativity, namely, the existence of a universal finite speed limit and the exclusion of instantaneous interactions.

Indeed, when we assume (i), instantaneous interactions and Newton's absolute time are inextricably related. To express this point formally, we shall consider spacetime transformations between two inertial frames in the so-called standard configuration using the same notation as in[1]

$$x' = F(x, t; v), \tag{3}$$

$$t' = G(x, t; v), \tag{4}$$

where v represents the velocity of frame $\mathcal{O}'$ with respect to $\mathcal{O}$. Let an object A in $\mathcal{O}$, at $x = x_a$, cause an instantaneous effect at time $t = t_1$ through a fundamental interaction on a distant object B, located at $x = x_b$. In the other inertial frame $\mathcal{O}'$, since the laws of physics are the same in $\mathcal{O}$ and $\mathcal{O}'$, the effect must also occur at the same time on A and B, say $t' = t'_1$. The time coordinate transformation (4) gives

$$t'_1 = G(x_1, t_1; v), \tag{5}$$

$$t'_1 = G(x_2, t_1; v). \tag{6}$$

Since x_1 and x_2 are arbitrary, the time variable must be independent of the spatial coordinate, $t' = G(t; v)$. Homogeneity of time requires that the ratio dt'/dt be also independent of time,

$$\frac{dt'}{dt} = \frac{\partial}{\partial t} G(t; v) = a(v), \tag{7}$$

then by integration,

$$t' = a(v)t + b(v). \tag{8}$$

We can take $b(v) = 0$ by adequate initial conditions, for instance, by choosing $t' = 0$ when $t = 0$,

$$t' = a(v)t. \tag{9}$$

Space isotropy requires $a(v) = a(-v)$ and, since the inverse relation is obtained when changing v to $-v$,

$$t = a(-v)t' = a(v)t'. \tag{10}$$

Replacing Eq. (9) in Eq. (10),

$$t = [a(v)]^2 t. \tag{11}$$

The former equation leads us to $[a(v)]^2 = 1 \rightarrow a(v) = \pm 1$. Since for $v = 0$, it must reduce to the identity transformation, we are left with $t' = t$.

Thus, when we assume there is no limit to the speed of the transmission of interactions, Newton's absolute time is not optional but a necessary imposition.

If II stands for "infinite speed of interactions" and AT for "absolute time," we have proved the following implication:

$$II \Rightarrow AT. \tag{12}$$

Vice versa, if we set $\neg II = FI$ (finite speed of interactions) and $\neg AT = RT$ (relative time), Einstein's train-embankment thought experiment[22] proves

$$FI \Rightarrow RT. \tag{13}$$

The contrapositive of Eq. (13) is equivalent to the converse of Eq. (12).

$$AT \Rightarrow II. \tag{14}$$

Equations (12) and (14) establish the complete equivalence of infinite interactions with the absolute character of time.

III. LINEARITY

Here, we address two issues that were not sufficiently clarified in Sec. **B** of Ref. 1. In the following, F and G refer to the spacetime transformations (3) and (4), which we shall assume constitute a twice differentiable bijection. We include equation numbers in Ref. 1 with an asterisk. From Eqs. (3) and (4),

$$p' = \frac{dx'}{dt'} = \frac{p\dfrac{\partial F}{\partial x} + \dfrac{\partial F}{\partial t}}{p\dfrac{\partial G}{\partial x} + \dfrac{\partial G}{\partial t}} \quad (6^*), \tag{15}$$

where p and p' are the velocities in $\mathcal{O}$ and $\mathcal{O}'$ respectively. When $p=0$ we must have $p'=-v$ and from Eq. (15) we obtain

$$\frac{\partial F}{\partial t}+v\frac{\partial G}{\partial t}=0 \quad (7^*). \tag{16}$$

Similarly, when $p'=0$, we have $p=v$ and (15) reduces to

$$v\frac{\partial F}{\partial x}+\frac{\partial F}{\partial t}=0 \quad (8^*). \tag{17}$$

Equations (15)–(17) form the basis of Verheest's formulation.

A. First issue

The first issue arises after equation (8*). There Verheest asserts, "This implies that F is a function of the combined argument $x - vt$ as well as of v" without further explanation. Note that v enters the equation as a parameter, F being a function of the variables x and t.

It is clear that if F has the functional form $F(x-vt;v)$, (17) is satisfied. However, the former argument constitutes only a sufficient condition, and Verheest's derivation requires F to have that functional form necessarily.

Luckily that has an elegant solution. As observed in Ref. 1, from Eqs. (16) and (17) we obtain

$$\frac{\partial G}{\partial t}=\frac{\partial F}{\partial x} \quad (9^*). \tag{18}$$

Taking derivatives with respect to t in Eqs. (16) and (18)

$$\frac{\partial^2 F}{\partial t^2}+v\frac{\partial^2 G}{\partial^2 t}=0, \tag{19}$$

$$\frac{\partial^2 G}{\partial^2 t}=\frac{\partial^2 F}{\partial x\partial t}, \tag{20}$$

eliminating from the former two equations $\partial^2 G/\partial^2 t$,

$$\frac{\partial^2 F}{\partial t^2}+v\frac{\partial^2 F}{\partial t\partial x}=0. \tag{21}$$

Taking derivative with respect to x in Eq. (17),

$$v\frac{\partial^2 F}{\partial x^2}+\frac{\partial^2 F}{\partial x\partial t}=0. \tag{22}$$

Eliminating the cross derivatives in Eqs. (21) and (22),

$$\frac{\partial^2 F}{\partial x^2}-\frac{1}{v^2}\frac{\partial^2 F}{\partial t^2}=0. \tag{23}$$

The former wave equation and its general solution is well-known to physics students,

$$F(x,t;v)=f_v(x-vt)+g_v(x+vt). \tag{24}$$

The g_v part of the solution does not satisfy (17) so we must have $g_v=0$ and we obtain the necessary solution

$$F(x,t;v)=f_v(x-vt). \tag{25}$$

Another approach to finding (25) is to employ the usual substitution for solving the wave equation in Eq. (17)

B. Second issue

The second issue arises when solving a homogeneous linear system that Verheest's method requires,

$$\frac{\partial^2 F}{\partial t\partial x}+v\frac{\partial^2 G}{\partial t\partial x}=0,$$
$$C(v)\frac{\partial^2 F}{\partial t\partial x}-\frac{\partial^2 G}{\partial t\partial x}=0 \quad (12^*), \tag{26}$$

where

$$C(v)=\frac{\partial G/\partial x}{\partial F/\partial x} \quad (11^*). \tag{27}$$

Verheest's derivation is based on the vanishing of the second order derivatives in Eq. (26), which is justified assuming a nonzero determinant of the coefficients, $[1+vC(v)]\neq 0$. However, a rigorous treatment requires that we also analyze the case when $1+vC(v)=0$, giving

$$C(v)=-\frac{1}{v}. \tag{28}$$

When this happens, we cannot assume that both cross derivatives vanish. In this case, the system (26) reduces to a single equation,

$$\frac{\partial^2 G}{\partial t\partial x}=-\frac{1}{v}\frac{\partial^2 F}{\partial t\partial x}, \tag{29}$$

so we cannot assume that both sides vanish. To solve this case, from Eqs. (16), (28), and (27)

$$\frac{\partial G}{\partial t}=-\frac{1}{v}\frac{\partial F}{\partial t}, \tag{30}$$

$$\frac{\partial G}{\partial x}=-\frac{1}{v}\frac{\partial F}{\partial x}. \tag{31}$$

Replacing (25) in the former two equations,

$$\frac{\partial G}{\partial t}=f'_v, \tag{32}$$

$$\frac{\partial G}{\partial x}=-\frac{1}{v}f'_v. \tag{33}$$

By integration, we have

$$G(x,t;v)=-\frac{1}{v}f_v+h(x), \tag{34}$$

$$G(x,t;v)=-\frac{1}{v}f_v+l(t). \tag{35}$$

Therefore, $h(x)=l(t)=k=const.$ and we are left with the following spacetime transformation:

$$t'=-\frac{1}{v}f_v(x-vt)+k, \tag{36}$$

$$x'=f_v(x-vt). \tag{37}$$

However, this transformation is inadmissible because it does not have an inverse. Really, when $t'\neq -(1/v)x'+k$, it does not have solution in (x,t).

IV. CONCLUSIONS

We have complemented Verheet's linearity proof with two observations that may ease its detailed understanding for the reader. However, the foundationally relevant points were discussed in Sec. II. From a conceptual viewpoint, we have stressed that it is better to base the derivation of Lorentz transformation using axioms (i) and (ii′), replacing the light principle with a more physically compelling one. In this respect, we highlight two authoritative references, Landau and Lifshitz[21] and Jackson;[5] both postulate (ii′) instead of the light principle. In particular, Jackson explicitly spells out, *Because special relativity applies to everything, not just light, it is desirable to express the second postulate in terms that convey its generality:*

> In every inertial frame, there is a finite universal limiting speed C for physical entities.

Thus, by using postulate (ii′) instead of the light principle, we gain physical insight regarding the finite character of the speed of interactions avoiding any particular reference to light.

On the other hand, if to avoid any reference to light, we try to derive the spacetime transformations using only postulate (i), it becomes hard to motivate the puzzling abandonment of the absolute character of time and we leave unanswered Verheest justified question:

> How do we incorporate the transformation of time from one inertial observer to the next? Of course, we all know the Lorentz transformation, but how to get there?

The above-mentioned question was duly responded to in Sec. II. The form of incorporating time into the transformation is necessarily subject to our assumption about the existence of instantaneous interactions or their impossibility. Instantaneous interactions necessarily imply absolute Newtonian time (12). Hence, the time transformation cannot include spatial variables.

On the contrary, the existence of a universal finite limiting speed for physical interactions requires abandoning absolute time (13), which demands that time enter the transformation as a fourth coordinate depending on the spatial variables.

Finally, we observe that instantaneous interactions can be considered a special case of finite speed interactions by setting the universal constant $c = \infty$.

Thus, the mathematical reduction of the Lorentz transformations to the Galilean case when $c \to \infty$ (in practice, when $v/c \ll 1$) is naturally justified. This is according to "The Hierarchy of Theories." When an old theory (Newtonian physics) is replaced by a new one (special relativity), the new one does not disprove the former. What actually happened was that the old one continued to provide the correct predictions, but only for a limited set of phenomena.[23]

ACKNOWLEDGMENTS

The author thanks the reviewers for significantly improving the article.

AUTHOR DECLARATIONS

Conflict of Interest

The author has no conflicts to disclose.

[a)]Electronic mail: jupalam@gmail.com, ORCID: 0000-0002-3146-4786.

[1]F. Verheest, "On the linearity of the generalized Lorentz transformation," Am. J. Phys. **90**(6), 425–429 (2022).
[2]D. Lawden, *Introduction to Tensor Calculus, Relativity and Cosmology* (Dover Publications, Inc., New York, 1982).
[3]J. Aharoni, *The Special Theory of Relativity* (Oxford U. P., London, 1965).
[4]If we restrict ourselves to classical Newtonian mechanics, similar transformations to (1) but more stringent ones are also possible.[24]
[5]J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (John Wiley & Sons, Inc, New York, 1999).
[6]J. M. Charap, *Covariant Electrodynamics* (Johns Hopkins U. P., Baltimore, 2011).
[7]V. Fock, *The Theory of Space, Time & Gravitation*, 2nd ed. (Pergamon Press Ltd., Oxford, England, 1966).
[8]Y. P. Terletskii, *Paradoxes in the Theory of Relativity* (Springer Science+Business Media, New York, 1968).
[9]S. Weinberg, *Gravitation and Cosmology* (John Wiley and Sons, New York, 1972).
[10]V. Berzi and V. Gorini, "Reciprocity principle and the Lorentz transformations," J. Math. Phys. **10**(8), 1518–1524 (2003).
[11]R. Resnick, *Introduction to Special Relativity* (John Wiley and Sons, New York, 1968).
[12]J. M. Lévy-Leblond, "One more derivation of the Lorentz transformation," Am. J. Phys. **44**(3), 271–277 (1976).
[13]B. Ignatowsky, Phys. Z. **11**, 972 (1910), see https://en.wikisource.org/w/index.php?title=Translation:Some_General_Remarks_on_the_Relativity_Principle&oldid=10821770.
[14]P. B. Pal, "Nothing but relativity," Eur. J. Phys. **24**(3), 315–319 (2003).
[15]H. Almström, "Derivation of the Lorentz transformation without use of light," J. Phys. A **1**(3), 331–333 (1968).
[16]He also assumed the usual properties of inertial frames, homogeneity of time and homogeneity and isotropy of space.
[17]The first to formulate the general "postulate of relativity" was Henri Poincaré.[25]
[18]J. Reitz, F. Milford, and R. Christy, *Foundations of Electromagnetic Theory*, Addison-Wesley Series in Physics (Addison-Wesley Publishing Company, Reading, MA, 1993).
[19]B. Schutz, *A First Course in General Relativity*, 2nd ed. (Cambridge U. P., New York, 2009).
[20]J.-P. Anker and F. Ziegler, "Relativity without light: A new proof of Ignatowski's theorem," J. Geometry Phys. **158**, 103871 (2020).
[21]L. Landau and E. Lifshitz, *Course of Theoretical Physics, Volume 2, the Classical Theory of Fields*, 4th ed. (Elsevier, Oxford, UK, 1975).
[22]A. Einstein, *Relativity: The Special and General* (Pearson Education, Inc., New York, 2005).
[23]F. Rohrlich, *Classical Charged Particles*, 3rd ed. (World Scientific, Singapore, 2007).
[24]O. Jahn and V. V. Sreedhar, "The maximal invariance group of Newton's equations for a free point particle," Am. J. Phys. **69**(10), 1039–1043 (2001).
[25]H. Lorentz, "Deux mémoires de Henri Poincaré sur la physique mathématique," Acta Math. **38**, 293–308 (1921).