Why special relativity?

Shan Gao
Research Center for Philosophy of Science and Technology,
Shanxi University, Taiyuan 030006, P. R. China
E-mail: gaoshan2017@sxu.edu.cn.

March 9, 2019

Abstract

It has been shown that Lorentz-like transformations with an invariant speed can be derived based on the homogeneity and isotropy of space and time and the principle of relativity. However, since the transformations can be Lorentzian or Galilean, depending on whether the invariant speed is finite, a real connection has not been established between the transformations and special relativity. In this paper, I present an argument supporting the finiteness of the invariant speed in the Lorentz-like transformations. The new analysis suggests that special relativity can indeed be derived based not on the particular light postulate but on the universal properties of space and time.

Special relativity is originally based on two main postulates: the principle of relativity and the constancy of the speed of light (Einstein 1905). But, as Einstein later admitted to some extent (Einstein 1935), it is an incoherent mixture (Stachel 1995); the first principle is universal in scope, while the second is only a particular property of light, which has obvious electrodynamical origins in Maxwell's theory. In view of this potential issue, there have been attempts to drop the light postulate from special relativity, which can be traced back to Ignatowski (1910) (see also Torretti 1983; Brown 2005). It turns out that based on the homogeneity and isotropy of space and time and the principle of relativity one can derive Lorentz-like transformations with an undetermined invariant speed. Unlike special relativity that needs to assume the constancy of the speed of light, an invariant speed naturally appears in the Lorentz-like transformations. This is a surprise indeed.

However, since the value of the invariant speed can be finite or infinite, the Lorentz-like transformations actually allows two possible transformations: Lorentzian and Galilean. It seems that an empirical element is still needed to determine the invariant speed and further eliminate the Galilean transformations, although it may not refer to any properties of light in an essential way (Lévy-Leblond 1976; Mermin 1984).[1] This raises serious doubts about the connection between the Lorentz-like transformations and special relativity. Some authors insisted that the light postulate is still needed to derive the Lorentz transformations (Pauli 1921; Resnick 1967; Miller 1981; Drory 2015). Others doubted that the Lorentz-like transformations are really relativistic in nature (Brown 2005). In this paper, I will argue that the value of the invariant speed in the Lorentz-like transformations should be finite, and thus the light postulate can indeed be dropped and replaced by the universal properties of space and time such as homogeneity and isotropy. Before presenting my argument, let me first introduce a concise derivation of the Lorentz-like transformations (see also Pal 2003; Drory 2015).

Consider two inertial frames S and S', where S moves relative to S' with a speed v directed along the x axis and when $t = 0$ the origins of the two frames coincide. The general coordinate transformations between S and S' can be written as follows:

$$x' = F(x, t, v) \quad (1)$$
$$t' = G(x, t, v) \quad (2)$$

where (x', t') denotes the space and time coordinates in the frame S', and (x, t) denotes the space and time coordinates in the frame S.

(1). Homogeneity of space and time

The homogeneity of space and time requires that the space and time intervals between two events should not depend on the positions and instants of the events in each inertial frame. Consider two events separated by a space interval dx and a time interval dt. Let the first event's coordinates be (x_i, y_i, z_i, t_i) and the second's $(x_i + dx, y_i, z_i, t_i + dt,)$ in the frame S. Then the separations between these two events as observed from S' are

$$dx' = \frac{\partial F(x, t_i, v)}{\partial x}|_{x=x_i} dx + \frac{\partial F(x_i, t, v)}{\partial t}|_{t=t_i} dt \quad (3)$$
$$dt' = \frac{\partial G(x, t_i, v)}{\partial x}|_{x=x_i} dx + \frac{\partial G(x_i, t, v)}{\partial t}|_{t=t_i} dt \quad (4)$$

Since dx' and dt' do not depend on x_i and t_i in the frame S' as required by the homogeneity of space and time, and dx and dt are independent of each other, the partial derivatives in the above two equations should be only functions of v. This means that the homogeneity of space and time requires that the transformations are linear with respect to both space and time. Considering that the origins of the two frames S and S' coincide when $t = 0$, the linear transformations can be writen in matrix form as follows:

[1]See also Drory (2015) for a different view.

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \begin{pmatrix} A(v) & B(v) \\ C(v) & D(v) \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} \tag{5}$$

Since the origin of S' moves at a speed v relative to the origin of S along the x axis, i.e., $x' = 0$ when $x = vt$, we also have the following relation:

$$B(v) = -vA(v). \tag{6}$$

(2). Isotropy of space

The isotropy of space requires that the transformations should not change when the x axis is reversed, i.e., when x, x' and v change sign. By this requirement we have:

$$\begin{cases} A(-v) = A(v) \\ B(-v) = -B(v) \\ C(-v) = -C(v) \\ D(-v) = D(v) \end{cases} \tag{7}$$

(3). Principle of relativity

The principle of relativity requires that the inverse transformations assume the same form as the original transformations. This means that the transformations from S' to S assume the same functional forms as the transformations from S to S'. Moreover, the combination of the principle of relativity with isotropy of space further implies reciprocity (Berzi and Gorini 1969; Torretti 1983; Budden 1997), namely that the speed of S relative to S' is the negative of the speed of S' relative to S. Thus we have:

$$\begin{cases} A(-v) = \frac{D(v)}{A(v)D(v)-B(v)C(v)} \\ B(-v) = \frac{-B(v)}{A(v)D(v)-B(v)C(v)} \\ C(-v) = \frac{-C(v)}{A(v)D(v)-B(v)C(v)} \\ D(-v) = \frac{A(v)}{A(v)D(v)-B(v)C(v)} \end{cases} \tag{8}$$

When combining the conditions (7) and (8) we obtain:

$$D(v) = A(v). \tag{9}$$

$$C(v) = \frac{A^2(v) - 1}{B(v)}. \tag{10}$$

Then considering (6) the transformations can be formulated in terms of only one unknown function $A(v)$, namely

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \begin{pmatrix} A(v) & -vA(v) \\ -\frac{A^2(v)-1}{vA(v)} & A(v) \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} \tag{11}$$

or

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = A(v) \begin{pmatrix} 1 & -v \\ -\frac{A^2(v)-1}{vA^2(v)} & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} \tag{12}$$

In order to determine the form of $A(v)$, we may consider a third frame S'' which moves with a speed u relative to S' along the x axis. Then we have:

$$\begin{aligned} \begin{pmatrix} x'' \\ t'' \end{pmatrix} &= A(u)A(v) \begin{pmatrix} 1 & -u \\ -\frac{A^2(u)-1}{uA^2(u)} & 1 \end{pmatrix} \begin{pmatrix} 1 & -v \\ -\frac{A^2(v)-1}{vA^2(v)} & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} \\ &= A(u)A(v) \begin{pmatrix} 1+u\frac{A^2(v)-1}{vA^2(v)} & -(u+v) \\ -\frac{A^2(u)-1}{uA^2(u)} - \frac{A^2(v)-1}{vA^2(v)} & 1+v\frac{A^2(u)-1}{uA^2(u)} \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} \end{aligned} \tag{13}$$

The principle of relativity requires that this transformation assumes the same form as the transformation from S to S', and thus the two diagonal elements of the matrix also satisfy (9), namely they are equal. Then we have:

$$1+u\frac{A^2(v)-1}{vA^2(v)} = 1+v\frac{A^2(u)-1}{uA^2(u)}. \tag{14}$$

or

$$\frac{A^2(v)-1}{v^2A^2(v)} = \frac{A^2(u)-1}{u^2A^2(u)}. \tag{15}$$

Since u and v are arbitrary, this equation means that its both sides are constants. Denoting this constant by K and considering the condition $A(v)=1$ when $v=0$, we have:

$$A(v) = \frac{1}{\sqrt{1-Kv^2}} \tag{16}$$

Therefore, the final transformations are

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \frac{1}{\sqrt{1-Kv^2}} \begin{pmatrix} 1 & -v \\ -Kv & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} \tag{17}$$

Note that the isotropy of time and reciprocity require that the coordinate transformations for y and z are $y' = y$ and $z' = z$ (see Dey, 2018).

The velocity addition law can be further derived based on the above analysis. Suppose the speed of the frame S'' relative to S is w. Then using (16) and (13) in which the first diagonal element of the matrix is $A(w)$ by definition, we can obtain the velocity addition law:

$$w = \frac{u+v}{1+Kuv} \tag{18}$$

It can be seen that $1/\sqrt{K}$ is an invariant speed, independent of any inertial frame.

The possible values of K can be determined as follows. The first diagonal element of the matrix in (13) demands $A(v) \geq 1$, since if $A(v) < 1$ then for some values of u and v (e.g. $u \ll v$) we will obtain $A(w) < 0$, which contradicts (16). Thus we have $K \geq 0$ according to (16). In addition, since $A(v)$ is real, (16) also demands that K has a finite upper limit (unless in a motionless world where v can only be zero). For example, we have $K < 1/v^2$ for any possible velocity v, which also means that $1/\sqrt{K}$ is the maximum speed.

When $K = 0$ the Lorentz-like transformations (17) become the Galilean transformations, and when $K > 0$ they become the Lorentz transformations. This means that the Lorentz-like transformations are the most general transformations consistent with the principle of relativity, which can accommodate both Galilean and Einsteinian relativity. In this sense, they are not (Einsteinian) relativistic in nature (Brown 2005). Certainly, we may resort to experience to eliminate the possibility of $K = 0$. But there is still a deep question that has not been answered, namely why $K > 0$ or why the invariant, maximum speed is finite. Without answering this question, there is still one theoretical step left from the Lorentz-like transformations to special relativity.

In the following, I will try to fill the gap between the Lorentz-like transformations and special relativity by arguing that K should be larger than zero. First, no fundamental principles consistent with experience require that the value of K should be zero, since $K = 0$ leads to the Galilean transformations that contradict experience. Second, if there are fundamental principles that require the value of K should be larger than zero, then the gap between the Lorentz-like transformations and special relativity will be filled. Note that this is independent of whether we actually know what these principles are. If only these principles exist, the Lorentz-like transformations will become the Lorentz transformations. I will not discuss this possibility in this paper.

Third, consider the last possibility that no fundamental principles determine the value of K. If we can argue that the value of K is also larger than zero in this case, then the Lorentz-like transformations will be equivalent to special relativity. If no principles determine the value of K, then K can be zero or larger than zero. The question is: what are the possibilities of $K = 0$ and $K > 0$, respectively? Since K has a finite upper limit, we may consider a finite range of the possible values of K, $[0, \epsilon]$. The question will be what the probability of K being zero is and what the probability of K being larger than zero is. Obviously the sum of the two probabilities is one. Since no fundamental principles determine the value of K, it seems reasonable to assume that the probability of K assuming any possible value is the same; if the probability distribution of the possible values of K is not an even distribution, then a principle will be needed to explain why. A much weaker assumption is that the probability distribution of the possible values of K is a continuous function. At the very least, since no principles require that the value of K should be zero, the probability distribution of the possible values of K is not a δ function in the position $K = 0$, namely $p(K) \neq \delta(K)$. Then, we can find that the probability of $K = 0$ is zero, while the probability of $K > 0$ is one. In other words, the value of K in the Lorentz-like transformations will be larger than zero and the transformations will be Lorentz transformations.

Note that if the value of K is randomly selected from infinitely many possible values, $K = 1/c^2$ (where $c \approx 3 \times 10^8 m/s$) in our universe will be a contingent fact. In other possible universes, the values of K may be different, but they are arguably not zero either according to the above analysis. On the other hand, why the invariant speed $1/\sqrt{K}$ is equal to the speed of a massless particle such as photon in our universe is arguably not a contingent fact, but a result of the fundamental dynamics.

To sum up, I have argued that the invariant speed in the Lorentz-like transformations should be finite. The new analysis suggests that special relativity can indeed be derived based not on the particular light postulate but on the universal properties of space and time such as homogeneity and isotropy.

Acknowledgments

This work is supported by the National Social Science Foundation of China (Grant No. 16BZX021).

References

[1] Berzi, V. and Gorini, V. (1969). Reciprocity principle and the Lorentz transformations. J. Math. Phys., 10, 1518-24.

[2] Brown, H. (2005). Physical Relativity: Spacetime structure from a dynamical perspective. Oxford: Clarendon Press.

[3] Budden, T. (1997). A star in the Minkowskian sky: Anisotropic special relativity. Studies in History and Philosophy of Modern Physics 28(3), 325-361.

[4] Dey, S. (2018). Time isotropy, Lorentz transformation and inertial frames. Studies in History and Philosophy of Modern Physics, 63, 123-127.

[5] Drory, A. (2015). The necessity of the second postulate in special relativity. Studies in the History and Philosophy of Modern Physics, 51, 57-67.

[6] Einstein, A. (1905). Zur Elektrodynamik bewegter Korper, Annalen der Physik, 17, pp. 891-921. English Translation in: The Collected Papers of Albert Einstein, Volume 2, English translation. Princeton University Press, New Jersey (1989).

[7] Einstein, A. (1935). An elementary derivation of the equivalence of mass and energy. Bulletin of the American Mathematical Society. 41, 223-230.

[8] Ignatowski, W. V. (1910). Einige allgemeine Bemerkungen zum Relativittsprinzip. Phys. Zeits. 11, 972-976.

[9] Lévy-Leblond, J. M. (1976). One more derivation of the Lorentz transformation. Am. J. Phys. 44, 271-277.

[10] Miller, A.I. (1981). Albert Einstein's Special Theory of Relativity: Emergence (1905) and Early Interpretation (1905-1911). New York: Addison-Wesley.

[11] Pal, P. B. (2003). Nothing but relativity. Eur. J. Phys. 24, 315-319.

[12] Pauli, W. (1921). Theory of Relativity, Encyclopedia of Mathematics, V19. Leipzig: Teubner.

[13] Resnick, R. (1968). Introduction to Special Relativity. New York: John Wiley & Sons.

[14] Stachel, J. (1995). History of relativity. In L. M. Brown, A. Pais, and B. Pippard, (eds.) Twentieth Century Physics, vol.1, pp. 249-356. New York: American Institute of Physics.

[15] Torretti, R. (1983). Relativity and Geometry. Oxford: Pergamon Press, 1983.