

On the formal statement of the special principle of relativity

Márton Gömöri and László E. Szabó

Department of Logic, Institute of Philosophy

Eötvös University, Budapest

<http://phil.elte.hu/logic>

Preprint (12 September 2011)

Abstract

The aim of the paper is to develop a proper mathematical formalism which can help to clarify the necessary conceptual plugins to the special principle of relativity and leads to a deeper understanding of the principle in its widest generality.

1 Introduction

The aim of this paper is to spell out the (special) relativity principle (RP) in a precise mathematical form. There are various verbal formulations of the principle. In its shortest form it says that “All the laws of physics take the same form in any inertial frame of reference.” The laws of physics *in* a reference frame K are meant to be the laws of physics as they are ascertained by an observer being at rest relative to the reference frame K ; less anthropomorphically, as they appear in the results of the measurements, such that both the measuring equipments and the objects to be measured are co-moving with K .

For example, consider the following simple application of the principle in Einstein’s 1905 paper:

Let there be given a *stationary* rigid rod; and let its length be l as *measured by a measuring-rod which is also stationary*. We now imagine the axis of the rod lying along the axis of x of the stationary system of co-ordinates, and that a uniform motion of parallel translation with velocity v along the axis of x in the direction of increasing x is then imparted to the rod. We now inquire as to the length of the *moving* rod, and imagine its length to be ascertained by the following *two* operations:

- (a) The observer *moves together with the given measuring-rod and the rod to be measured*, and *measures the length* of the rod directly by superposing the measuring-rod, *in just the same way as if all three were at rest*.

- (b) By means of *stationary* clocks set up in the *stationary* system and synchronizing in accordance with [the light-signal synchronization], the observer ascertains at what points of the *stationary* system the two ends of the rod to be measured are located at a definite time. The distance between these two points, measured by the measuring-rod already employed, which in this case is *at rest*, is also a length which may be designated “the length of the rod.”

In accordance with the principle of relativity the length to be discovered by the operation (a)—we will call it “the length of the rod in the moving system”—must be equal to the length l of the stationary rod.

The length to be discovered by the operation (b) we will call “the length of the (*moving*) rod in the *stationary* system.” This we shall determine *on the basis of our two principles*, and we shall find that it *differs* from l . [all italics added]

Thus, with the following formulation (Szabó 2004) one can express in more detail how the principle is actually understood:

- (RP) The physical description of the behavior of a system co-moving as a whole with an inertial frame K , expressed in terms of the results of measurements obtainable by means of measuring equipments co-moving with K , takes the same form as the description of the similar behavior of the same system when it is co-moving with another inertial frame K' , expressed in terms of the measurements with the same equipments when they are co-moving with K' .

Our main concern in this paper is to unpack the verbal statement (RP) and to provide its general mathematical formulation. We are hopeful that the formalism we develop here helps to clarify the required conceptual plugins to the RP and leads to a deeper understanding of the principle in its widest generality.

2 The statement of the RP

Consider an arbitrary collection of physical quantities $\xi_1, \xi_2, \dots, \xi_n$ in K , operationally defined by means of some equipments being at rest in K . Let $\xi'_1, \xi'_2, \dots, \xi'_n$ denote another collection of physical quantities that are defined by the *same operations* with the *same equipments*, but in *different state of motion*, namely, in which they are all moving with constant velocity \mathbf{V} relative to K , co-moving with K' . Since, for all $i = 1, 2, \dots, n$, both ξ_i and ξ'_i are measured by the same equipment—although in different physical conditions—with the same pointer scale, it is plausible to assume that the possible values of ξ_i and ξ'_i range over the same $\sigma_i \subseteq \mathbb{R}$. We introduce the following notation: $\Sigma = \times_{i=1}^n \sigma_i \subseteq \mathbb{R}^n$.

It must be emphasized that quantities $\xi_1, \xi_2, \dots, \xi_n$ and $\xi'_1, \xi'_2, \dots, \xi'_n$ are, a priori, *different* physical quantities, due to the fact that the operations by which the quantities are defined are performed under *different* physical conditions; with measuring equipments of *different* states of motion. Any objective (non-conventional) relationship between them must be a contingent law

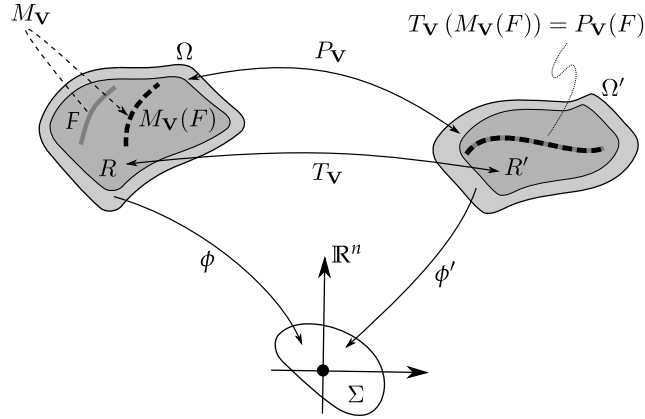


Figure 1: The relativity principle

of nature. Thus, the same numeric values, say, $(5, 12, \dots, 61) \in \mathbb{R}^n$ correspond to different states of affairs when $\zeta_1 = 5, \zeta_2 = 12, \dots, \zeta_n = 61$ versus $\zeta'_1 = 5, \zeta'_2 = 12, \dots, \zeta'_n = 61$. Consequently, $(\zeta_1, \zeta_2, \dots, \zeta_n)$ and $(\zeta'_1, \zeta'_2, \dots, \zeta'_n)$ are *not* elements of the *same* “space of physical quantities”; although the numeric values of the physical quantities, in both cases, can be represented in $\Sigma = \times_{i=1}^n \sigma_i \subseteq \mathbb{R}^n$.

Mathematically, one can express this fact by means of two different n -dimensional manifolds, Ω and Ω' , each covered by one global coordinate system, ϕ and ϕ' respectively, such that $\phi : \Omega \rightarrow \Sigma$ assigns to every point of Ω one of the possible n -tuples of numerical values of physical quantities $\zeta_1, \zeta_2, \dots, \zeta_n$ and $\phi' : \Omega' \rightarrow \Sigma$ assigns to every point of Ω' one of the possible n -tuples of numerical values of physical quantities $\zeta'_1, \zeta'_2, \dots, \zeta'_n$ (Fig. 1). In this way, a point $\omega \in \Omega$ represents the *class* of physical constellations in which the quantities $\zeta_1, \zeta_2, \dots, \zeta_n$ take the values $\zeta_1 = \phi_1(\omega), \zeta_2 = \phi_2(\omega), \dots, \zeta_n = \phi_n(\omega)$; similarly, a point $\omega' \in \Omega'$ represents the physical constellation characterized by $\zeta'_1 = \phi'_1(\omega'), \zeta'_2 = \phi'_2(\omega'), \dots, \zeta'_n = \phi'_n(\omega')$.¹ Again, these physical constellations are generally different, even in case of $\phi(\omega) = \phi'(\omega') \in \mathbb{R}^n$.

In the above sense, the points of Ω and the points of Ω' range over all possible value combinations of physical quantities $\zeta_1, \zeta_2, \dots, \zeta_n$ and $\zeta'_1, \zeta'_2, \dots, \zeta'_n$. It might be the case however that some combinations are impossible, in the sense that they never come to existence in the physical world. Let us denote by $R \subseteq \Omega$ and $R' \subseteq \Omega'$ the physically admissible parts of Ω and Ω' . Note that $\phi(R)$ is not necessarily identical with $\phi'(R')$.²

We shall use a bijection $P_V : \Omega \rightarrow \Omega'$ (“putting primes”; Bell 1987, p. 73) defined by means of the two coordinate maps ϕ and ϕ' :

$$P_V \stackrel{\text{def}}{=} (\phi')^{-1} \circ \phi \quad (1)$$

In contrast with P_V , we now introduce the concept of what we call the

¹ $\phi_i = \pi_i \circ \phi$, where π_i is the i -th coordinate projection in \mathbb{R}^n .

²One can show however that $\phi(R) = \phi'(R')$ if the RP, that is (7), holds.

“transformation” of physical quantities. It is conceived as a bijection

$$T_V : \Omega \supseteq R \rightarrow R' \subseteq \Omega' \quad (2)$$

determined by the contingent fact that whenever a physical constellation belongs to the class represented by some $\omega \in R$ then it also belongs to the class represented by $T_V(\omega) \in R'$, and vice versa. Since $\zeta_1, \zeta_2, \dots, \zeta_n$ can be various physical quantities in the various contexts, nothing guarantees that such a bijection exists. We assume however the existence of T_V .

Remark 1. It is worthwhile to consider several examples.

- (a) Let (ζ_1, ζ_2) be (p, T) , the pressure and the temperature of a given (equilibrium) gas; and let (ζ'_1, ζ'_2) be (p', T') , the pressure and the temperature of the same gas, measured by the moving observer in K' . In this case, there exists a one-to-one T_V :

$$p' = p \quad (3)$$

$$T' = T\gamma^{-1} \quad (4)$$

where $\gamma = \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}}$ (Tolman 1949, pp. 158–159).³ A point $\omega \in \Omega$ of coordinates, say, $p = 101325$ and $T = 300$ (in units Pa and $^\circ K$) represents the class of physical constellations—the class of possible worlds—in which the gas in question has pressure of $101325 Pa$ and temperature of $300^\circ K$. Due to (4), this class of physical constellations is different from the one represented by $P_V(\omega) \in \Omega'$ of coordinates $p' = 101325$ and $T' = 300$; but it is identical to the class of constellations represented by $T_V(\omega) \in \Omega'$ of coordinates $p' = 101325$ and $T' = 300\gamma^{-1}$.

- (b) Let $(\zeta_1, \zeta_2, \dots, \zeta_{10})$ be $(t, x, y, z, E_x, E_y, E_z, r_x, r_y, r_z)$, the time, the space coordinates where the electric field strength is taken, the three components of the field strength, and the space coordinates of a particle. And let $(\zeta'_1, \zeta'_2, \dots, \zeta'_{10})$ be $(t', x', y', z', E'_x, E'_y, E'_z, r'_x, r'_y, r'_z)$, the similar quantities obtainable by means of measuring equipments co-moving with K' . In this case, there is no suitable one-to-one T_V , as the electric field strength in K does not determine the electric field strength in K' , and vice versa.
- (c) Let $(\zeta_1, \zeta_2, \dots, \zeta_{13})$ be $(t, x, y, z, E_x, E_y, E_z, B_x, B_y, B_z, r_x, r_y, r_z)$ and let $(\zeta'_1, \zeta'_2, \dots, \zeta'_{13})$ be $(t', x', y', z', E'_x, E'_y, E'_z, B'_x, B'_y, B'_z, r'_x, r'_y, r'_z)$, where B_x, B_y, B_z and B'_x, B'_y, B'_z are the magnetic field strengths in K and K' . In this case, in contrast with (b), the well known Lorentz transformations of the spatio-temporal coordinates and the electric and magnetic field strengths constitute a proper one-to-one T_V . \lrcorner

Next we turn to the general formulation of the concept of the *description of a particular behavior* of a physical system, say, in K . We are probably not far from

³There is a debate over the proper transformation rules (Georgieiu 1969).

the truth if we assume that such a description is, in its most abstract sense, a *relation* between physical quantities $\zeta_1, \zeta_2, \dots, \zeta_n$; in other words, it can be given as a subset $F \subset R$.

Remark 2. Consider the above example (a) in Remark 1. An isochoric process of the gas can be described by the subset F that is, in coordinates, determined by the following single equation:

$$F \quad \{p = \kappa T \quad (5)$$

with a certain constant κ .

To give another example, consider the case (b). The relation F given by equations

$$F \quad \begin{cases} E_x = E_0 \\ E_y = 0 \\ E_z = 0 \\ r_x = x_0 + v_0 t \\ r_y = 0 \\ r_z = 0 \end{cases} \quad (6)$$

with some specific values of E_0, x_0, v_0 describes a neutral particle moving with constant velocity in a static homogeneous electric field. \lrcorner

Of course, one may not assume that an arbitrary relation $F \subset R$ has physical meaning. Let $\mathcal{E} \subset 2^R$ be the set of those $F \subset R$ which describe a particular behavior of the system. We shall call \mathcal{E} the *set of equations* describing the physical system in question. The term is entirely justified. In practical calculations, two systems of equations are regarded to be equivalent if and only if they have the same solutions. Therefore, a system of equations can be identified with the set of its solutions. In general, the equations can be algebraic equations, ordinary and partial integro-differential equations, linear and nonlinear, whatever. So, in its most abstract sense, a system of equations is a set of subsets of R .

Now, consider the following subsets⁴ of Ω' , determined by an $F \in \mathcal{E}$:

$P_V(F) \subseteq \Omega'$ which formally is the “primed F ”, that is a relation of exactly the same “form” as F , but in the primed variables $\zeta'_1, \zeta'_2, \dots, \zeta'_n$. Note that relation $P_V(F)$ does not necessarily describe a true physical situation, as it can be not realized in nature.

$T_V(F) \subseteq R'$ which is the same description of the same physical situation as F , but *expressed* in the primed variables.

We need one more concept. The RP is about the connection between two situations: one is in which the system, as a whole, is at rest relative to inertial frame K , the other is in which the system shows the similar behavior, but being in a collective motion relative to K , co-moving with K' . In other words, we assume the existence of a map $M_V : \mathcal{E} \rightarrow \mathcal{E}$, assigning to each $F \in \mathcal{E}$, stipulated to describe the situation in which the system is co-moving as a whole with inertial frame K , another relation $M_V(F) \in \mathcal{E}$, describing the similar behavior of the

⁴We denote the map of type $\Omega \rightarrow \Omega'$ and its direct image maps of type $2^\Omega \rightarrow 2^{\Omega'}$ and $2^{2^\Omega} \rightarrow 2^{2^{\Omega'}}$ or their restrictions by the same symbol.

same system when it is, as a whole, co-moving with inertial frame K' , that is, when it is in a collective motion with velocity \mathbf{V} relative to K .

Now, applying all these concepts, what the RP states is the following:

$$T_{\mathbf{V}}(M_{\mathbf{V}}(F)) = P_{\mathbf{V}}(F) \quad \text{for all } F \in \mathcal{E} \quad (7)$$

or equivalently,

$$P_{\mathbf{V}}(F) \subset R' \text{ and } M_{\mathbf{V}}(F) = T_{\mathbf{V}}^{-1}(P_{\mathbf{V}}(F)) \quad \text{for all } F \in \mathcal{E} \quad (8)$$

Remark 3. Notice that, for a given fixed F , everything on the right hand side of the equation in (8), $P_{\mathbf{V}}$ and $T_{\mathbf{V}}$, are determined *only by* the physical behaviors of *the measuring equipments* when they are in various states of motion. In contrast, the meaning of the left hand side, $M_{\mathbf{V}}(F)$, depends on the physical behavior of *the object physical system* described by F and $M_{\mathbf{V}}(F)$, when it is in various states of motion. That is to say, the two sides of the equation reflect the behaviors of *different parts* of the physical reality; and the RP expresses a law-like regularity between the behaviors of these different parts. \lrcorner

Remark 4. Let us illustrate these concepts with a well-known textbook example of a static versus uniformly moving charged particle. The static field of a charge q being at *rest* at point (x_0, y_0, z_0) in K is the following:

$$F \left\{ \begin{array}{l} E_x = \frac{q(x - x_0)}{\left((x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2\right)^{3/2}} \\ E_y = \frac{q(y - y_0)}{\left((x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2\right)^{3/2}} \\ E_z = \frac{q(z - z_0)}{\left((x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2\right)^{3/2}} \\ B_x = 0 \\ B_y = 0 \\ B_z = 0 \end{array} \right. \quad (9)$$

The stationary field of a charge q *moving* at constant velocity $\mathbf{V} = (V, 0, 0)$ relative to K can be obtained by solving the equations of electrodynamics (in K) with the time-dependent source (for example, Jackson 1999, pp. 661–665):

$$M_{\mathbf{V}}(F) \left\{ \begin{array}{l} E_x = \frac{qX_0}{\left(X_0^2 + (y - y_0)^2 + (z - z_0)^2\right)^{3/2}} \\ E_y = \frac{\gamma q (y - y_0)}{\left(X_0^2 + (y - y_0)^2 + (z - z_0)^2\right)^{3/2}} \\ E_z = \frac{\gamma q (z - z_0)}{\left(X_0^2 + (y - y_0)^2 + (z - z_0)^2\right)^{3/2}} \\ B_x = 0 \\ B_y = -c^{-2}VE_z \\ B_z = c^{-2}VE_y \end{array} \right. \quad (10)$$

where where (x_0, y_0, z_0) is the initial position of the particle at $t = 0$, $X_0 = \gamma(x - (x_0 + Vt))$.

Now, we form the same expressions as (9) but in the *primed* variables of the co-moving reference frame K' :

$$P_{\mathbf{V}}(F) \left\{ \begin{array}{l} E'_x = \frac{q'(x' - x'_0)}{\left((x' - x'_0)^2 + (y' - y'_0)^2 + (z' - z'_0)^2\right)^{3/2}} \\ E'_y = \frac{q'(y' - y'_0)}{\left((x' - x'_0)^2 + (y' - y'_0)^2 + (z' - z'_0)^2\right)^{3/2}} \\ E'_z = \frac{q'(z' - z'_0)}{\left((x' - x'_0)^2 + (y' - y'_0)^2 + (z' - z'_0)^2\right)^{3/2}} \\ B'_x = 0 \\ B'_y = 0 \\ B'_z = 0 \end{array} \right. \quad (11)$$

By means of the Lorentz transformation rules of the space-time coordinates, the field strengths and the electric charge (e.g. Tolman 1949), one can express (11) in terms of the original variables of K :

$$T_{\mathbf{V}}^{-1}(P_{\mathbf{V}}(F)) \left\{ \begin{array}{l} E_x = \frac{qX_0}{\left(X_0^2 + (y - y_0)^2 + (z - z_0)^2\right)^{3/2}} \\ E_y = \frac{\gamma q (y - y_0)}{\left(X_0^2 + (y - y_0)^2 + (z - z_0)^2\right)^{3/2}} \\ E_z = \frac{\gamma q (z - z_0)}{\left(X_0^2 + (y - y_0)^2 + (z - z_0)^2\right)^{3/2}} \\ B_x = 0 \\ B_y = -c^{-2}VE_z \\ B_z = c^{-2}VE_y \end{array} \right. \quad (12)$$

We find that the result is indeed the same as (10) describing the field of the moving charge: $M_{\mathbf{V}}(F) = T_{\mathbf{V}}^{-1}(P_{\mathbf{V}}(F))$. That is to say, the RP seems to be true in this particular case.

Reversely, *assuming* that the particle + electromagnetic field system satisfies the RP, that is, (8) holds for the equations of electrodynamics, one can *derive* the stationary field of a uniformly moving point charge (10) from the static field (9). ┘

3 Covariance

Now we have a strict mathematical formulation of the RP for a physical system described by a system of equations \mathcal{E} . Remarkably, however, we still have not encountered the concept of “covariance” of equations \mathcal{E} . The reason is that the RP and the covariance of equations \mathcal{E} are not equivalent—in contrast to what many believe. In fact, the logical relationship between the two conditions is much more complex. To see this relationship in more detail, we previously need to clarify a few things.

Consider the following two sets: $P_{\mathbf{V}}(\mathcal{E}) = \{P_{\mathbf{V}}(F)|F \in \mathcal{E}\}$ and $T_{\mathbf{V}}(\mathcal{E}) = \{T_{\mathbf{V}}(F)|F \in \mathcal{E}\}$. Since a system of equations can be identified with its set of solutions, $P_{\mathbf{V}}(\mathcal{E}) \subset 2^{\Omega'}$ and $T_{\mathbf{V}}(\mathcal{E}) \subset 2^{R'}$ can be regarded as two systems of equations for functional relations between $\xi'_1, \xi'_2, \dots, \xi'_n$. In the primed variables, $P_{\mathbf{V}}(\mathcal{E})$ has “the same form” as \mathcal{E} . Nevertheless, it can be the case that $P_{\mathbf{V}}(\mathcal{E})$ does not express a true physical law, in the sense that its solutions do not necessarily describe true physical situations. In contrast, $T_{\mathbf{V}}(\mathcal{E})$ is nothing but \mathcal{E} expressed in variables $\xi'_1, \xi'_2, \dots, \xi'_n$.

Now, covariance intuitively means that equations \mathcal{E} “preserve their forms against the transformation $T_{\mathbf{V}}$ ”. That is, in terms of the formalism we developed:

$$T_{\mathbf{V}}(\mathcal{E}) = P_{\mathbf{V}}(\mathcal{E}) \tag{13}$$

or, equivalently,

$$P_{\mathbf{V}}(\mathcal{E}) \subset 2^{R'} \text{ and } \mathcal{E} = T_{\mathbf{V}}^{-1}(P_{\mathbf{V}}(\mathcal{E})) \tag{14}$$

The first thing we have to make clear is that—even if we know or presume that it holds—covariance (14) is obviously *not sufficient* for the RP (8). For, (14) only guarantees the invariance of the set of solutions, \mathcal{E} , against $T_{\mathbf{V}}^{-1} \circ P_{\mathbf{V}}$, but it says nothing about which solution of \mathcal{E} corresponds to which solution. In Bell’s words:

Lorentz invariance alone shows that for any state of a system at rest there is a corresponding ‘primed’ state of that system in motion. But it does not tell us that if the system is set anyhow in motion, it will actually go into the ‘primed’ of the original state, rather than into the ‘prime’ of some *other* state of the original system. (Bell 1987, p. 75)

While it is the very essence of the RP that the solution $M_{\mathbf{V}}(F)$, describing the system *in motion* relative to K , corresponds to solution $T_{\mathbf{V}}^{-1} \circ P_{\mathbf{V}}(F)$. For example, what we use in the above mentioned textbook derivation of the stationary electromagnetic field of a uniformly moving point charge (end of Remark 4)

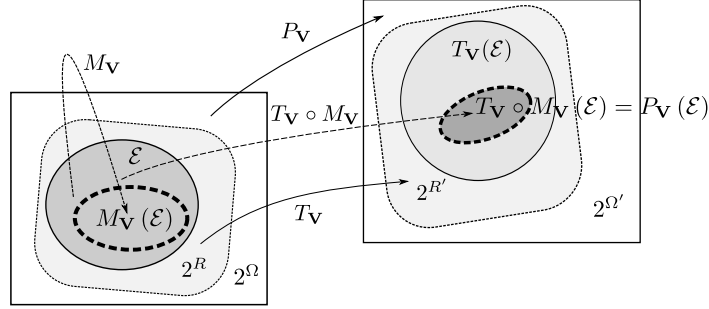


Figure 2: The RP only implies that $T_V(\mathcal{E}) \supseteq T_V \circ M_V(\mathcal{E}) = P_V(\mathcal{E})$. Covariance of \mathcal{E} would require that $T_V(\mathcal{E}) = P_V(\mathcal{E})$, which is generally not the case

is not the covariance of the equations—that would be not enough—but statement (8), that is, what the RP claims about the solutions of the equations in detail.

In a precise sense, covariance is not only not sufficient for the RP, but it is *not even necessary* (Fig. 2). The RP only implies that

$$T_V(\mathcal{E}) \supseteq T_V(M_V(\mathcal{E})) = P_V(\mathcal{E}) \quad (15)$$

(7) implies (13) only if we have the following *extra* condition:

$$M_V(\mathcal{E}) = \mathcal{E} \quad (16)$$

4 Initial and boundary conditions

Let us finally consider the situation when the solutions of a system of equations \mathcal{E} are specified by some extra conditions—initial and/or boundary value conditions, for example. In our general formalism, an extra condition for \mathcal{E} is a system of equations $\psi \subset 2^\Omega$ such that there exists exactly one solution $[\psi]_\mathcal{E}$ satisfying both \mathcal{E} and ψ . That is, $\mathcal{E} \cap \psi = \{[\psi]_\mathcal{E}\}$, where $\{[\psi]_\mathcal{E}\}$ is a singleton set. Since $\mathcal{E} \subset 2^R$, without loss of generality we may assume that $\psi \subset 2^R$.

Since P_V and T_V are injective, $P_V(\psi)$ and $T_V(\psi)$ are extra conditions for equations $P_V(\mathcal{E})$ and $T_V(\mathcal{E})$ respectively, and we have

$$P_V([\psi]_\mathcal{E}) = [P_V(\psi)]_{P_V(\mathcal{E})} \quad (17)$$

$$T_V([\psi]_\mathcal{E}) = [T_V(\psi)]_{T_V(\mathcal{E})} \quad (18)$$

for all extra conditions ψ for \mathcal{E} . Similarly, if $P_V(\mathcal{E}), P_V(\psi) \subset 2^{R'}$ then $T_V^{-1}(P_V(\psi))$ is an extra condition for $T_V^{-1}(P_V(\mathcal{E}))$, and

$$\left[T_V^{-1}(P_V(\psi)) \right]_{T_V^{-1}(P_V(\mathcal{E}))} = T_V^{-1} \left([P_V(\psi)]_{P_V(\mathcal{E})} \right) \quad (19)$$

Consider now a set of extra conditions $\mathcal{C} \subset 2^{2^R}$. Assume that \mathcal{C} is a *parametrizing set of extra conditions* for \mathcal{E} ; by which we mean that for all $F \in \mathcal{E}$ there exists exactly one $\psi \in \mathcal{C}$ such that $F = [\psi]_\mathcal{E}$; in other words,

$$\mathcal{C} \ni \psi \mapsto [\psi]_\mathcal{E} \in \mathcal{E} \quad (20)$$

is a bijection.

$M_{\mathbf{V}} : \mathcal{E} \rightarrow \mathcal{E}$ was introduced as a map between solutions of \mathcal{E} . Now, as there is a one-to-one correspondence between the elements of \mathcal{C} and \mathcal{E} , it generates a map $M_{\mathbf{V}} : \mathcal{C} \rightarrow \mathcal{C}$, such that

$$[M_{\mathbf{V}}(\psi)]_{\mathcal{E}} = M_{\mathbf{V}}([\psi]_{\mathcal{E}}) \quad (21)$$

Thus, from (17) and (21), the RP, that is (7), has the following form:

$$T_{\mathbf{V}}([M_{\mathbf{V}}(\psi)]_{\mathcal{E}}) = [P_{\mathbf{V}}(\psi)]_{P_{\mathbf{V}}(\mathcal{E})} \quad \text{for all } \psi \in \mathcal{C} \quad (22)$$

or, equivalently, (8) reads

$$[P_{\mathbf{V}}(\psi)]_{P_{\mathbf{V}}(\mathcal{E})} \subset R' \text{ and } [M_{\mathbf{V}}(\psi)]_{\mathcal{E}} = T_{\mathbf{V}}^{-1} \left([P_{\mathbf{V}}(\psi)]_{P_{\mathbf{V}}(\mathcal{E})} \right) \quad (23)$$

One might make use of the following theorem:

Theorem 1. *Assume that the system of equations $\mathcal{E} \subset 2^R$ is covariant, that is, (13) is satisfied. Then,*

- (i) *for all $\psi \in \mathcal{C}$, $T_{\mathbf{V}}(M_{\mathbf{V}}(\psi))$ is an extra condition for the system of equations $P_{\mathbf{V}}(\mathcal{E})$, and, (22) is equivalent to the following condition:*

$$[T_{\mathbf{V}}(M_{\mathbf{V}}(\psi))]_{P_{\mathbf{V}}(\mathcal{E})} = [P_{\mathbf{V}}(\psi)]_{P_{\mathbf{V}}(\mathcal{E})} \quad (24)$$

- (ii) *for all $\psi \in \mathcal{C}$, $P_{\mathbf{V}}(\psi) \subset 2^{R'}$, $T_{\mathbf{V}}^{-1}(P_{\mathbf{V}}(\psi))$ is an extra condition for the system of equations \mathcal{E} and (23) is equivalent to the following condition:*

$$[M_{\mathbf{V}}(\psi)]_{\mathcal{E}} = \left[T_{\mathbf{V}}^{-1}(P_{\mathbf{V}}(\psi)) \right]_{\mathcal{E}} \quad (25)$$

Proof. (i) Obviously, $T_{\mathbf{V}}(\mathcal{E}) \cap T_{\mathbf{V}}(M_{\mathbf{V}}(\psi))$ exists and is a singleton; and, due to (13), it is equal to $P_{\mathbf{V}}(\mathcal{E}) \cap T_{\mathbf{V}}(M_{\mathbf{V}}(\psi))$; therefore this latter is a singleton, too. Applying (18) and (13), we have

$$T_{\mathbf{V}}([M_{\mathbf{V}}(\psi)]_{\mathcal{E}}) = [T_{\mathbf{V}}(M_{\mathbf{V}}(\psi))]_{T_{\mathbf{V}}(\mathcal{E})} = [T_{\mathbf{V}}(M_{\mathbf{V}}(\psi))]_{P_{\mathbf{V}}(\mathcal{E})} \quad (26)$$

therefore, (24) implies (23).

(ii) Similarly, due to $P_{\mathbf{V}}(\psi) \subset 2^{R'}$ and (14), $\mathcal{E} \cap T_{\mathbf{V}}^{-1}(P_{\mathbf{V}}(\psi))$ exists and is a singleton. Applying (19) and (14), we have

$$T_{\mathbf{V}}^{-1} \left([P_{\mathbf{V}}(\psi)]_{P_{\mathbf{V}}(\mathcal{E})} \right) = \left[T_{\mathbf{V}}^{-1}(P_{\mathbf{V}}(\psi)) \right]_{T_{\mathbf{V}}^{-1}(P_{\mathbf{V}}(\mathcal{E}))} = \left[T_{\mathbf{V}}^{-1}(P_{\mathbf{V}}(\psi)) \right]_{\mathcal{E}} \quad (27)$$

that is, (25) implies (23). \square

Remark 5. Let us note a few important facts which can easily be seen in the formalism we developed:

- (a) The covariance of a set of equations \mathcal{E} does *not* imply the covariance of a subset of equations separately. It is because a smaller set of equations corresponds to an $\mathcal{E}^* \subset 2^R$ such that $\mathcal{E} \subset \mathcal{E}^*$; and it does not follow from (13) that $T_{\mathbf{V}}(\mathcal{E}^*) = P_{\mathbf{V}}(\mathcal{E}^*)$.

- (b) Similarly, the covariance of a set of equations \mathcal{E} does *not* guarantee the covariance of an arbitrary set of equations which is only satisfactory to \mathcal{E} ; for example, when the solutions of \mathcal{E} are restricted by some extra conditions. Because from (13) it does not follow that $T_{\mathbf{V}}(\mathcal{E}^*) = P_{\mathbf{V}}(\mathcal{E}^*)$ for an arbitrary $\mathcal{E}^* \subset \mathcal{E}$.
- (c) The same holds, of course, for the combination of cases (a) and (b); for example, when we have a smaller set of equations $\mathcal{E}^* \supset \mathcal{E}$ together with some extra conditions $\psi \subset 2^R$. For, (13) does not imply that $T_{\mathbf{V}}(\mathcal{E}^* \cap \psi) = P_{\mathbf{V}}(\mathcal{E}^* \cap \psi)$.
- (d) However, covariance is guaranteed if a covariant set of equations is restricted with a *covariant* set of extra conditions; because $T_{\mathbf{V}}(\mathcal{E}) = P_{\mathbf{V}}(\mathcal{E})$ and $T_{\mathbf{V}}(\psi) = P_{\mathbf{V}}(\psi)$ trivially imply that $T_{\mathbf{V}}(\mathcal{E} \cap \psi) = P_{\mathbf{V}}(\mathcal{E} \cap \psi)$. ┘

5 Concluding discussions and open problems

As we have seen, the notion of $M_{\mathbf{V}}$ plays a crucial role. Formally, one could say, the RP is *relative* to the definition of $M_{\mathbf{V}}$; the physical content of the RP depends on how this concept is physically understood. But, what does it mean to say that a physical system is the same and of the same behavior as the one described by F , except that it is, as a whole, in a collective motion with velocity \mathbf{V} relative to K ? Without answering this crucial question the RP is meaningless.

In fact, the same question can be asked with respect to the definitions of quantities $\xi'_1, \xi'_2, \dots, \xi'_n$ —and, therefore, with respect to the meanings of $T_{\mathbf{V}}$ and $P_{\mathbf{V}}$. For, $\xi'_1, \xi'_2, \dots, \xi'_n$ are not simply arbitrary variables assigned to reference frame K' , in one-to-one relations with $\xi_1, \xi_2, \dots, \xi_n$, but the physical quantities obtainable by means of the same operations with the same measuring equipments as in the operational definitions of $\xi_1, \xi_2, \dots, \xi_n$, except that everything is in a collective motion with velocity \mathbf{V} . Therefore, we should know what we mean by “the same measuring equipment but in collective motion”. From this point of view, it does not matter whether the system in question is the object to be observed or a measuring equipment involved in the observation.

These questions can be answered only within the given physical context; and, one must admit, in some situations the answers are non trivial and ambiguous (cf. Szabó 2004). At this level of generality we only want to point out two things.

First, whatever is the definition of $M_{\mathbf{V}} : \mathcal{E} \rightarrow \mathcal{E}$ in the given context, the following is a minimal requirement for it to have the assumed physical meaning:

- (M) Relations $F \in \mathcal{E}$ must describe situations which can be meaningfully characterized as such in which the system as a whole is at rest or in motion with some velocity relative to a frame of reference.

For example, in Remark 4, solutions (9) and (10) satisfy this condition, as in both cases the system of the charged particle + electromagnetic field qualifies as a system in collective rest or motion. The electromagnetic field is in collective motion with the point charge of velocity \mathbf{V} (Fig. 3) in the following sense:

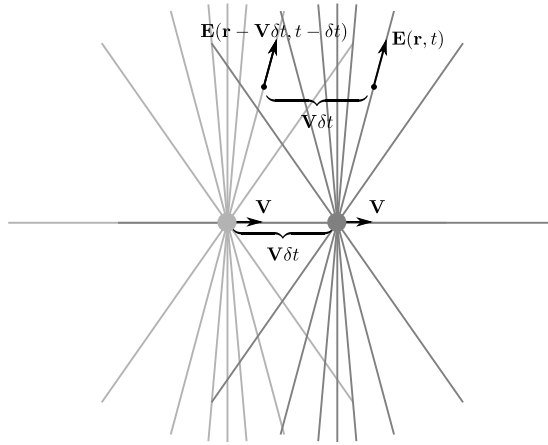


Figure 3: The stationary field of a uniformly moving point charge is in collective motion together with the point charge

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r} - \mathbf{V}\delta t, t - \delta t) \quad (28)$$

$$\mathbf{B}(\mathbf{r}, t) = \mathbf{B}(\mathbf{r} - \mathbf{V}\delta t, t - \delta t) \quad (29)$$

Notice that requirement (M) says nothing about whether and how the fact that the system as a whole is at rest or in motion with some velocity is reflected in the solutions $F \in \mathcal{E}$. It does not even require that this fact can be expressed in terms of $\xi_1, \xi_2, \dots, \xi_n$. It only requires that each $F \in \mathcal{E}$ belong to a physical situation in which it is meaningful to say—perhaps in terms of quantities different from $\xi_1, \xi_2, \dots, \xi_n$ —that the system is at rest or in motion relative to a reference frame. How a concrete physical situation can be characterized as such in which the system is at rest or in motion is a separate problem, which can be discussed in the particular contexts.

The second thing to be said about $M_{\mathbf{V}}(F)$ is that it is a notion determined by the concrete physical context; but it is *not* equal to the “Lorentz boosted solution” $T_{\mathbf{V}}^{-1}(P_{\mathbf{V}}(F))$ *by definition*—as a little reflection shows:

- (a) In this case, (8) would read

$$T_{\mathbf{V}}^{-1}(P_{\mathbf{V}}(F)) = T_{\mathbf{V}}^{-1}(P_{\mathbf{V}}(F)) \quad (30)$$

That is, the RP would become a tautology; a statement which is always true, independently of any contingent fact of nature; independently of the actual behavior of moving physical objects; and independently of the actual empirical meanings of physical quantities $\xi'_1, \xi'_2, \dots, \xi'_n$. But, the RP is supposed to be a fundamental *law of nature*. Note that a tautology is entirely different from a fundamental principle, even if the principle is used as a fundamental hypothesis or fundamental premise of a theory, from which one derives further physical statements. For, a fundamental premise, as expressing a contingent fact of nature, is potentially falsifiable by testing its consequences; a tautology is not.

- (b) Even if accepted, $M_{\mathbf{V}}(F) \stackrel{def}{=} T_{\mathbf{V}}^{-1}(P_{\mathbf{V}}(F))$ can provide physical meaning to $M_{\mathbf{V}}(F)$ only if we know the meanings of $T_{\mathbf{V}}$ and $P_{\mathbf{V}}$, that is, if we know the empirical meanings of the quantities denoted by $\xi'_1, \xi'_2, \dots, \xi'_n$. But, the physical meaning of $\xi'_1, \xi'_2, \dots, \xi'_n$ are obtained from the operational definitions: they are the quantities obtained by “the same measurements with the same equipments when they are, as a whole, co-moving with K' with velocity \mathbf{V} relative to K ”. Symbolically, we need, priory, the concepts of $M_{\mathbf{V}}(\xi_i\text{-equipment at rest})$. And this is a conceptual circularity: in order to have the concept of what it is to be an $M_{\mathbf{V}}(\text{brick at rest})$ the (size)' of which we would like to ascertain, we need to have the concept of what it is to be an $M_{\mathbf{V}}(\text{measuring rod at rest})$ —which is exactly the same conceptual problem.
- (c) One might claim that we do not need to specify the concepts of $M_{\mathbf{V}}(\xi_i\text{-equipment at rest})$ in order to know the *values* of quantities $\xi'_1, \xi'_2, \dots, \xi'_n$ we obtain by the measurements with the moving equipments, given that we can know the transformation rule $T_{\mathbf{V}}$ independently of knowing the operational definitions of $\xi'_1, \xi'_2, \dots, \xi'_n$. Typically, $T_{\mathbf{V}}$ is thought to be derived from the assumption that the RP (8) holds. If however $M_{\mathbf{V}}$ is, by definition, equal to $T_{\mathbf{V}}^{-1} \circ P_{\mathbf{V}}$, then in place of (8) we have the tautology (30), which does not determine $T_{\mathbf{V}}$.
- (d) Therefore, unsurprisingly, it is not the RP from which the transformation rules are routinely deduced, but the covariance (14). As we have seen, however, covariance is, in general, neither sufficient nor necessary for the RP. Whether (8) implies (14) hinges on the physical fact whether (16) is satisfied. But, if $M_{\mathbf{V}}$ is taken to be $T_{\mathbf{V}}^{-1} \circ P_{\mathbf{V}}$ by definition, the RP becomes true—in the form of tautology (30)—but does not imply covariance $T_{\mathbf{V}}^{-1} \circ P_{\mathbf{V}}(\mathcal{E}) = \mathcal{E}$.
- (e) Even if we assume that a “transformation rule” function $\phi' \circ T_{\mathbf{V}} \circ \phi^{-1}$ were derived from some independent premises—from the independent assumption of covariance, for example—how do we know that the $T_{\mathbf{V}}$ we obtained and the quantities of values $\phi' \circ T_{\mathbf{V}} \circ \phi^{-1}(\xi_1, \xi_2, \dots, \xi_n)$ are correct plugins for the RP? How could we verify that $\phi' \circ T_{\mathbf{V}} \circ \phi^{-1}(\xi_1, \xi_2, \dots, \xi_n)$ are indeed the values measured by a moving observer applying the same operations with the same measuring equipments, etc.?—without having an independent concept of $M_{\mathbf{V}}$, at least for the measuring equipments?
- (f) One could argue that we do not need such a verification; $\phi' \circ T_{\mathbf{V}} \circ \phi^{-1}(\xi_1, \xi_2, \dots, \xi_n)$ can be regarded as *the empirical definition* of the primed quantities:

$$(\xi'_1, \xi'_2, \dots, \xi'_n) \stackrel{def}{=} \phi' \circ T_{\mathbf{V}} \circ \phi^{-1}(\xi_1, \xi_2, \dots, \xi_n) \quad (31)$$

This is of course logically possible. The operational definition of the primed quantities would say: ask the observer at rest in K to measure $\xi_1, \xi_2, \dots, \xi_n$ with the measuring equipments at rest in K , and

then perform the mathematical operation (31). In this way, however, even the transformation rules would become tautologies; they would be true, no matter how the things are in the physical world.

- (g) Someone might claim that the identity of M_V with $T_V^{-1} \circ P_V$ is not a simple stipulation but rather an analytic truth which follows from the identity of the two *concepts*. Still, if that were the case, RP would be a statement which is true in all possible worlds; independently of any contingent fact of nature; independently of the actual behavior of moving physical objects.
- (h) On the contrary, as we have already pointed out in Remark 3, $M_V(F)$ and $T_V^{-1}(P_V(F))$ are *different concepts*, referring to different features of different parts of the physical reality. Any connection between the two things must be a contingent fact of the world.
- (i) $T_V^{-1} \circ P_V$ is a $2^R \rightarrow 2^R$ map which is completely determined by the physical behaviors of the *measuring* equipments. On the other hand, whether the elements of $\mathcal{E} \subset 2^R$ satisfy condition (M) and whether $T_V^{-1} \circ P_V(\mathcal{E}) \subseteq \mathcal{E}$ depend on the actual physical properties of the *object* physical system.
- (j) Let us note that in the standard textbook applications of the RP M_V is used as an independent concept, without any prior reference to the Lorentz boost $T_V^{-1} \circ P_V$. For example, we do not need to refer to the Lorentz transformations in order to understand the concept of ‘the stationary electromagnetic field of a uniformly moving point charge’; as we are capable to solve the electrodynamical equations for such a situation, within one single frame of reference, without even knowing of the Lorentz transformation rules. ┘

Acknowledgment

The research was partly supported by the OTKA Foundation, No. K 68043.

References

- Bell, J.S. (1987): How to teach special relativity, in *Speakable and unspeakable in quantum mechanics*. Cambridge, Cambridge University Press.
- Einstein, A (1905): Zur Elektrodynamik bewegter Körper, *Annalen der Physik* **17**, 891. (On the Electrodynamics of Moving Bodies, in H. A. Lorentz et al., *The principle of relativity: a collection of original memoirs on the special and general theory of relativity*. London, Methuen and Company 1923)
- Georgiou, A. (1969): Special relativity and thermodynamics, *Proc. Comb. Phil. Soc.* **66**, 423.

- Jackson, J.D. (1999): *Classical Electrodynamics (Third edition)*. Hoboken (NJ), John Wiley & Sons.
- Szabó, L.E. (2004): On the meaning of Lorentz covariance, *Foundations of Physics Letters* **17**, pp. 479–496.
- Tolman, R.C. (1949): *Relativity, Thermodynamics and Cosmology*. Oxford, Clarendon Press.