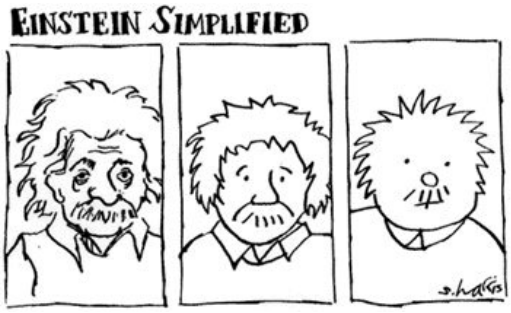
The Quintessence of Axiomatized Special Relativity Theory

Eugene Shubert
everythingimportant.org

December 6, 2013



Abstract

Albert Einstein made a simple yet significant contribution in physics by recognizing the key ideas in the original relativity theory of Henri Poincaré and Hendrik Lorentz, by dismissing the very complicated objectives of that theory and by only expounding on the easier and immediate consequences of the essential ideas. While Einstein's simplification of the overly-ambitious relativity theory of Poincaré and Lorentz was a noteworthy achievement, Einstein's synthesis was based on a very strange foundation, a seemingly impenetrable riddle. For that reason, to this day, most conceptualizations of special relativity are primarily anecdotal. That is a sufficient reason to rethink traditional explanations. The aim of *The Quintessence of Axiomatized Special Relativity* is to remove from Einstein's relativity theory everything that is confused, unnecessary and not amenable to experimental verification.

The most glaring misconception in relativity theory is the belief that nature requires linearity for the space and time transformation equations between inertial frames of reference. That mystifying, nonsensical, conceptual riddle, which began with Einstein, has been perpetuated by adoring physicists ever since. It should be dispensed with immediately. If spacetime is defined intuitively, tangibly, with minimal restrictions, then nonlinear Lorentz-equivalent transformation equations arise naturally.

Like the early Einstein, most authors of elementary textbooks on special relativity attach undue importance to how clocks should be synchronized. Requiring clocks to be synchronized is unnecessary. I believe that it's more enlightening to dispense with nonessentials and to pursue the elegance of greater simplicity and generality.

The approach taken here illustrates the essence of spacetime with a tangible model. It provides a clear foundation for not only Einstein's relativity but also admits the possibility that Einstein's postulates are too restrictive. It may be that all the laws of physics may be divided into two main categories. It is possible that some of the laws of physics are Lorentz invariant and that other physical laws are not. This generalization is required to allow for the possibility of motion faster than light and for our universe to have come into existence at some inexplicable event in the finite past.

Therefore, to understand the quintessence of relativity with greater clarity, I propose an entirely different definition of spacetime based on an altogether different anecdote: Why is the universe not infinitely old? Because waiting an infinite amount of time takes too much time, especially toward the end.

1 What is Axiomatization?

For mathematicians like David Hilbert, the axiomatization of physics means modeling all possible physical-theoretical frameworks in the same way that mathematicians typically define and investigate all fundamental mathematical structures. Consequently, it is to be understood that this derivation of special relativity theory represents a mathematician's perspective. It conforms to David Hilbert's philosophy of physics. [1].

Hilbert expressed the quintessence of his scientific philosophy as pure mathematics:

> "If geometry is to serve as a model for the treatment of physical axioms, we shall try first by a small number of axioms to include as large a class as possible of physical phenomena, and then by adjoining new axioms to arrive gradually at the more special theories. ...The mathematician will have also to take account not only of those theories coming near to reality, but also, as in geometry, of all logically possible theories. He must be always alert to obtain a complete survey of all conclusions derivable from the system of axioms assumed." David Hilbert, International Congress of Mathematicians, Paris France, 1900. [2].

2 A Minimal Axiom Set For Relativity Theory

There are a few essential preliminaries that I must address before I can specify my special axiom and then illustrate its meaning.

First and foremost: when I get to discussing spacetime, motion and the laws of physics, I will probably be referring to a specific mathematical model and not the physical universe that we actually inhabit. Please be attentive to the context to see if I mean the actual universe or just a discernable mathematical representation.

> "The sciences do not try to explain, they hardly even try to interpret, they mainly make models. By a model is meant a mathematical construct which, with the addition of certain verbal interpretations, describes observed phenomena. The justification of such a mathematical construct is solely and precisely that it is expected to work." — John von Neumann.

Second: Remember Hilbert's essential principle of axiomatization: physics should evolve from a small number of axioms. Therefore, you will be learning about spacetime in a series of steps. Each step features a toy model universe that inherits increasing complexity as you progress from one toy model universe to the next. Naturally, every model of spacetime that we can ever conceive of is only a piece of an increasingly complex hierarchy of spacetimes. My increasingly complex hierarchy will begin with the simplest spacetime imaginable.

Third: In general, spacetime is any mathematical model that combines space, time and motion into a single construct. Definitions vary and I have my own special definition. Spacetime in this tutorial is defined as a collection of inertial frames of reference. An inertial frame of reference is a Euclidean space with a mathematical definition of clock time defined at each point of that space. A concrete example of a mathematical clock is a geometric point that moves equal distances in equal times.

My fundamental axiom for relativity is:

> 1. The definition of clock time at each point in an inertial frame of reference is simply defined and mathematically well-defined. The ensuing time equations have the same mathematical form in all inertial frames of reference.

3 My First Toy Universe

The simplest spacetime imaginable, and upon which all the other spacetimes in this tutorial are built, is called Ξ_2.[1] Within it, only two kinds of motion are possible. You have to admit that spacetime of one spatial dimension where only two possible speeds are possible is easy to visualize.

It is common in mathematics and physics to call discrete levels of energy, position or configuration, *states*. Since I want to emphasize that only two kinds of motion are possible in the Ξ_2 universe, I shall label the two possible motions as state S_1 and state S_2 respectively. If Ξ_2 is to have only two possible states of motion in a simple one-dimensional space, then Ξ_2 consists of two one-dimensional Euclidean spaces.[2] Each state of motion S_i $(i = 1, 2)$ is assigned its own Euclidean space.

$\Gamma_2 \rightarrow$ $\cdots$ -9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9 $\cdots$

$\Gamma_1 \leftarrow$ $\cdots$ -9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9 $\cdots$

Figure 1

The Definition of Clock Time

How would you integrate the simplest kind of motion in a very elementary model of space so as to create a clear definition of clock time at each point of space?[3]

You can do that strictly visually or purely mathematically.

Take a careful look at figure 1. The diagram is meant to help you visualize Ξ_2 as two number lines Γ_1 and Γ_2. Γ_2 is moving to the right of Γ_1 or, equivalently, Γ_1 is moving to the left of Γ_2. Imagine that the two abstract lines Γ_1 and Γ_2 are infinitely long, frictionless, ethereal rulers. Suppose that Γ_2 is occupying the same space as Γ_1. There is no space between the rulers. The only objects that exist are these two nonmaterial rulers.

How should figure 1 be interpreted? We are required to create a definition of clock time for every point of Ξ_2. That is very easy to do. Little children know intuitively that a tiny arrow that moves steadily along a continuum of numbers is a clock. Simply notice that there is a very natural mathematical clock that exists at every point of Γ_1 and Γ_2. To see the clock at the point x_2 of Γ_2, for instance, imagine that the point x_2 is a pointer (visualize $\diamond$ or $\updownarrow$) that is moving along the continuum of numbers Γ_1. Except for the possible need of a sign change, that pointer is a clock. For the imaginary clock at the point x_1 of Γ_1, likewise, imagine that x_1 is a pointer that is moving along the continuum of numbers Γ_2.

Here is where the purely mathematical approach can speed things along and finish the construction. Clearly, every point of Γ_1 and Γ_2 is like a clock that moves equal distances in equal times. Consider the set of all ordered pairs of real numbers (x_1, x_2), where x_1 is an arbitrary element of Γ_1 and x_2 is an arbitrary element of Γ_2. The ordered pair (x_1, x_2) may be interpreted as the event where x_1 meets (touches, flies by or passes through) x_2, however you want to say it.

Next, define the clock time t_1 at the point x_1 of Γ_1, when the event (x_1, x_2) occurs, by the equation $t_1 = -x_2/\mu$. Likewise, define the clock time t_2 at the point x_2 of Γ_2, when the event (x_2, x_1) occurs, by the equation $t_2 = x_1/\mu$. The construction is now complete. Ξ_2 satisfies the definition of a spacetime. Curiously, the universe we inhabit is characterized by many fundamental constants. The universe Ξ_2 has the fundamental constant μ.

[1] Xi, which is written as Ξ and ξ, is the 14th letter of the Greek alphabet.

[2] Ξ_2 with the obvious topology, that of two topologically disconnected lines, satisfies the essential definition of a 1-manifold. Each point of Ξ_2 has an open set that is homeomorphic to the 1-dimensional space R^1.

[3] If you are a physicist, then you should be able to derive the time equations (1) and (2) on the next page without any help from me.

Clock Synchronization Schemes

As explained, $t_1 = -x_2/\mu$ and $t_2 = x_1/\mu$ are perfectly good definitions of clock time for the infinitely many clocks of Γ_1 and Γ_2, respectively. Suppose that you were to now replace all these clocks with new clocks point by point — the only difference between the old and new clocks being that for all time, past, present and future, the new clocks differ from the old by being set ahead one hour, or behind one hour or by any other fixed amount of time. Then $t_1 \mapsto t_1 + g_1(x_1)$ and $t_2 \mapsto t_2 + g_2(x_2)$. Consequently, the most general setting for mathematical clocks in Ξ_2 is given by the equations $t_1 = -x_2/\mu + g_1(x_1)$ and $t_2 = x_1/\mu + g_2(x_2)$. The functions $g_i(x_i)$, i=1,2 are called synchronization functions.

Proper Velocity

We have created the universe Ξ_2 with the understanding that every point of Ξ_2 is like the pointer of a clock that moves equal distances in equal times. This motion is best understood by first defining "proper velocity" and then doing a few calculations. For that we need the two fundamental equations:

$$t_1 = -x_2/\mu + g_1(x_1) \tag{1}$$

$$t_2 = x_1/\mu + g_2(x_2) \tag{2}$$

Proper velocity is defined as follows:

$$\mu_{12} = \frac{\Delta x_1}{\Delta t_2} = \left(\frac{x_1' - x_1}{t_2' - t_2}\right) \quad ; \quad \mu_{21} = \frac{\Delta x_2}{\Delta t_1} = \left(\frac{x_2' - x_2}{t_1' - t_1}\right) \tag{3}$$

To illustrate the meaning of these equations, pick a fixed point x_2 from Γ_2, a start time t_2 and an end time t_2'. We understand that the stationary point x_2 from Γ_2 is steadily moving through Γ_1. According to equation (2), the point x_2 at time t_2 is located at $x_1 = \mu t_2 - \mu g_2(x_2)$ on Γ_1. The same equation states that the point x_2 at a later time t_2' will be at x_1' where $x_1' = \mu t_2' - \mu g_2(x_2)$. We insert these values into the definition of μ_{12} and, with a little high school algebra, find that $\mu_{12} = \mu$. The result for μ_{12} is similar, except that we must use equation (1). Select a point x_1 from Γ_1, a beginning time t_1 and an end time t_1'. In that elapsed time $t_1' - t_1$, the point x_1, which is fixed on Γ_1, will move from $x_2 = -\mu t_1 + \mu g_1(x_1)$ to $x_2' = -\mu t_1' + \mu g_1(x_1)$. Plugging these values into the definition of μ_{21} gives $\mu_{21} = -\mu$.

Because all points of the line Γ_2 have equal proper velocities, and likewise for the line Γ_1, we shall say that the proper velocity of Γ_2 with respect to Γ_1 is μ_{12} and that the proper velocity of Γ_1 with respect to Γ_2 is μ_{21}. It is important to remember that $\mu_{21} = -\mu_{12}$ and that, regardless of the number of inertial frames of reference in any recognized universe, it is always true that $\mu_{ij} = -\mu_{ji}$.

That almost completes our study of Ξ_2. To make the transition from this universe to the more mathematically challenging ones easier, I will preemptively state a few necessary subtleties and observations.

It should be emphasized that, in my construction of Ξ_2, I invoked no axiom presupposing the existence of a cosmic everywhere present "now." It should also be recognized that I did not presuppose that the number lines Γ_1 and Γ_2 are equal in scale.

$\Gamma_2 \rightarrow$ $\cdots$ -9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9 $\cdots$

$\Gamma_1 \leftarrow$ $\cdots$ -9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9 $\cdots$

Figure 2

4 My Second Toy Universe

From your study of Ξ_2 you learned that moving coordinate systems may be interpreted as an infinite number of mathematical clocks. You saw that every point can be thought of as a pointer and that clock time is a function of ordered pairs of numbers (x_1, x_2), where x_1 is from Γ_1 and x_2 is from Γ_2 and that these tuples are events. The fundamental equations of Ξ_2 are:

$$t_1 = -x_2/\mu_{12} + g_1(x_1) \tag{4}$$

$$t_2 = -x_1/\mu_{21} + g_2(x_2) \tag{5}$$

In Ξ_2, $t_i(x_i, x_j)$ is the time on a clock positioned at x_i on the line Γ_i when the point x_j on the line Γ_j passes x_i. The proper velocity of the line Γ_j with respect to the line Γ_i is μ_{ij}. You are now ready to move beyond Ξ_2.

Consider the universe Ξ_3 featuring three inertial frames of reference, Γ_1, Γ_2, and Γ_3:

Γ_3	···	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	···
Γ_2	···	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	···
Γ_1	···	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	···

Figure 3

Our purpose here is to find a general definition of clock time on Ξ_3. We will use what we have learned about the universe Ξ_2. The mathematics can be extraordinary complex. The essential idea is very simple. Because we've added a third inertial frame of reference, there are now two different frames to select from in order to define time. We will require that every choice is inconsequential; the formula that defines time must give a consistent answer irrespective of the frame selected.

The obvious generalization of equations (4) and (5) would give us

$$t_i = -x_j/\mu_{ij} + g_i(\mu_{ij}, x_i) \qquad i \neq j \qquad \mu_{ij} = -\mu_{ji} \tag{6}$$

Nonlinear solutions to these clock time equations exist which in fact meet the mathematical constraints of our axioms but all known nonlinear solutions merely reset standard Lorentzian clock time. The Lorentz transformation equations come about from linear clock synchronization schemes and the reciprocity principle of special relativity. The reciprocity principle asserts that clocks should be synchronized in a simple manner so that if $x_j = 0$, then $x_i = vt_i$, and if $x_i = 0$, then $x_j = -vt_j$.

The most general equation for a linear and frame independent definition of clock time would be

$$t_i = -x_j/\mu_{ij} + g(\mu_{ij})x_i \qquad i \neq j \qquad \mu_{ij} = -\mu_{ji} \tag{7}$$

The constraint imposed on the function $g(\mu_{ij})$ by the reciprocity principle is easy to determine. From equation (7) please notice that $x_j = 0$ implies $x_i = t_i/g(\mu_{ij})$, and $x_i = 0$ implies $x_j = t_j/g(\mu_{ji})$. Reciprocity is therefore equivalent to the function g having the symmetric property $g(-\mu_{ij}) = -g(\mu_{ij})$ since $\mu_{ij} = -\mu_{ji}$.

As I have stated before, there are two ways to define clock time on any given line of Ξ_3. Our mathematically overdetermined system generates the following constraints:

$$t_1 = -x_2/\mu_{12} + g(\mu_{12})x_1 = -x_3/\mu_{13} + g(\mu_{13})x_1 \tag{8}$$

$$t_2 = -x_3/\mu_{23} + g(\mu_{23})x_2 = -x_1/\mu_{21} + g(\mu_{21})x_2 \tag{9}$$

$$t_3 = -x_1/\mu_{31} + g(\mu_{31})x_3 = -x_2/\mu_{32} + g(\mu_{32})x_3 \tag{10}$$

Solving these three equations for x_3 in terms of x_1 and x_2 yields

$$x_3 = x_1\mu_{13}(g(\mu_{13}) - g(\mu_{12})) + x_2\mu_{13}/\mu_{12} \tag{11}$$

$$x_3 = x_1\mu_{23}/\mu_{21} + x_2\mu_{23}(g(\mu_{23}) - g(\mu_{21})) \tag{12}$$

$$x_3 = x_1/(\mu_{31}(g(\mu_{31}) - g(\mu_{32}))) + x_2/(\mu_{32}(g(\mu_{32}) - g(\mu_{31}))) \tag{13}$$

Only two of the variables x_1, x_2 and x_3 are independent. Any two determine the third because all the proper velocities μ_{ij} have been specified. All the coefficients of x_1 in equations (11) to (13) are therefore equal. Similarly for x_2.

For time to be well-defined therefore in every Ξ_3 universe means that the following functional equations must be true:

$$\mu_{13}(g(\mu_{13}) - g(\mu_{12})) = \mu_{23}/\mu_{21} = 1/(\mu_{31}(g(\mu_{31}) - g(\mu_{32}))) \tag{14}$$

$$\mu_{23}(g(\mu_{23}) - g(\mu_{21})) = \mu_{13}/\mu_{12} = 1/(\mu_{32}(g(\mu_{32}) - g(\mu_{31}))) \tag{15}$$

If we were to solve equations (8) to (10), this time for x_2 in terms of x_1 and x_3 and then equate coefficients, we would find that

$$\mu_{12}(g(\mu_{12}) - g(\mu_{13})) = \mu_{32}/\mu_{31} = 1/(\mu_{21}(g(\mu_{21}) - g(\mu_{23}))) \tag{16}$$

$$\mu_{32}(g(\mu_{32}) - g(\mu_{31})) = \mu_{12}/\mu_{13} = 1/(\mu_{23}(g(\mu_{23}) - g(\mu_{21}))) \tag{17}$$

Equation (17) is identical to (15). Equations (14) to (16) contain many redundancies. Also, if we were to solve for x_1 in terms of x_2 and x_3 and proceed in the usual way, we wouldn't generate any new equations. When all the redundancies are eliminated, there are only three surviving equations:

$$g(\mu_{23}) - g(-\mu_{12}) = \frac{\mu_{13}}{\mu_{12}\mu_{23}} \tag{18}$$

$$g(\mu_{12}) - g(\mu_{13}) = \frac{\mu_{23}}{\mu_{12}\mu_{13}} \tag{19}$$

$$g(-\mu_{13}) - g(-\mu_{23}) = \frac{\mu_{12}}{\mu_{23}\mu_{13}} \tag{20}$$

These equations have a remarkably straightforward solution. Temporarily, let's assume that g is an odd function, i.e., $g(-\mu_{ij}) = -g(\mu_{ij})$. Therefore:

$$g(\mu_{23}) + g(\mu_{12}) = \frac{\mu_{13}}{\mu_{12}\mu_{23}} \tag{21}$$

$$g(\mu_{12}) - g(\mu_{13}) = \frac{\mu_{23}}{\mu_{12}\mu_{13}} \tag{22}$$

$$-g(\mu_{13}) + g(\mu_{23}) = \frac{\mu_{12}}{\mu_{23}\mu_{13}} \tag{23}$$

The amazing thing about these equations is that we can solve them in a few easy steps. Subtracting equation (23) from (22) gives us

$$g(\mu_{12}) - g(\mu_{23}) = \frac{\mu_{23}}{\mu_{12}\mu_{13}} - \frac{\mu_{12}}{\mu_{23}\mu_{13}} \tag{24}$$

Multiplying equations (21) and (24) together yields:

$$g^2(\mu_{12}) - g^2(\mu_{23}) = \frac{\mu_{13}}{\mu_{12}\mu_{23}}\left(\frac{\mu_{23}}{\mu_{12}\mu_{13}} - \frac{\mu_{12}}{\mu_{23}\mu_{13}}\right) = \frac{1}{(\mu_{12})^2} - \frac{1}{(\mu_{23})^2} \tag{25}$$

We conclude from the assumed independence of μ_{12} and μ_{23} that

$$g^2(\mu_{ij}) = \frac{1}{(\mu_{ij})^2} + k \tag{26}$$

Remembering the supposition that g is an odd function yields

$$g(\mu_{ij}) = \pm\frac{\sqrt{1+k\mu_{ij}^2}}{\mu_{ij}} \tag{27}$$

Taking a second look at our basic equations (18) to (20) shows that we can add a constant to any solution. Therefore:

$$g(\mu_{ij}) = \epsilon \pm \frac{\sqrt{1+k\mu_{ij}^2}}{\mu_{ij}} \tag{28}$$

I wish to write briefly on the physics of this function but prefer to do so using a simpler notation. Let's make the substitution $\mu_{12} = \mu$, $\mu_{23} = v$ and $\mu_{13} = \mu \oplus v$. Equation (18) and (28) suggests the possibility that

$$\mu \oplus v = -\mu v\left(\frac{\sqrt{1+k\mu^2}}{\mu} + \frac{\sqrt{1+kv^2}}{v}\right) \tag{29}$$

Is this a sensible formula for adding proper velocities? If μ is positive and we add to it an infinitesimally small positive v, then by continuity, we should expect a positive resultant velocity near to μ. The contradiction to continuity overthrows the possibility that a minus sign is a physically admissible solution of equation (28).

Consequently,

$$g(\mu_{ij}) = \epsilon + \frac{\sqrt{1+k\mu_{ij}^2}}{\mu_{ij}} \tag{30}$$

and

$$\mu \oplus v = \mu\sqrt{1+kv^2} + v\sqrt{1+k\mu^2} \tag{31}$$

Inserting the solution for $g(\mu_{ij})$ given by equation (30) into equation (7) yields

$$t_i = -\frac{x_j}{\mu_{ij}} + \left(\epsilon + \frac{\sqrt{1+k\mu_{ij}^2}}{\mu_{ij}}\right)x_i = -\frac{x_j}{\mu_{ij}} + \left(\frac{\sqrt{1+k\mu_{ij}^2}}{\mu_{ij}}\right)x_i + \epsilon x_i \tag{32}$$

It turns out that there is a physically meaningful property associated with the spacetime structure constant k in equation (32) but not so for the parameter ϵ. A nonzero epsilon universe only refers to an unusual way that clocks may be synchronized. I will prove this in section 6 by computing time dilation after resetting all spacetime clocks in all inertial frames of reference Γ_i $(i = 1, 2, 3)$ according to the general rule $t_i \mapsto t_i + S_i(x_i)$.

5 Nonlinear Lorentz Transformations

Clock time for all the spacetime clocks of Ξ_3 is described by the equation

$$t_i = -\frac{x_j}{\mu_{ij}} + \left(\frac{\sqrt{1 + k\mu_{ij}^2}}{\mu_{ij}} \right) x_i + S_i(x_i) \tag{33}$$

Possibly $S_i(x_i) = \epsilon x_i$ $(i = 1, 2, 3)$ for a specific Ξ_3 universe but an even more exotic synchronization scheme is possible. If you review the original constraints that were placed on equation (7), which resulted in the requirement that equation (8),(9) and (10) had to be true, you will notice that I could have just as easily started with a more general equation of the form

$$t_i = -x_j/\mu_{ij} + g(\mu_{ij})x_i + S_i(x_i) \tag{34}$$

It is easy to verify that there are no restrictions on the synchronization functions $S_i(x_i)$. Equation (33) is therefore well-defined. And just to be crystal clear in my use of equation (33) and (34) from this point onward, I shall define $g(\mu_{ij})$ to be

$$g(\mu_{ij}) = \frac{\sqrt{1 + k\mu_{ij}^2}}{\mu_{ij}} \tag{35}$$

and will let the ϵ (epsilon) in equation (30) be absorbed into the more general $S_i(x_i)$.

With those preliminaries out of the way, it is now a straightforward exercise to derive the Lorentz transformation equations easily and with great generality.

First, recall that equation (34) represents multiple equations. I need the following two:

$$t_i = -x_j/\mu + g(\mu)x_i + S_i(x_i) \tag{36}$$

$$t_j = x_i/\mu + g(-\mu)x_j + S_j(x_j) \tag{37}$$

As you can see, I have returned to the simpler notation introduced immediately after equation (28). Equation (36) is equation (34) without the subscripts on the proper velocity. You can write them in if you like or just remember that $\mu_{ij} = \mu$. Equation (37) is also represented in equation (34). You only need to replace i with j, and j with i and use the fact that $\mu_{ij} = -\mu_{ji}$.

The next step is to solve equation (36) for x_j. That step is trivial:

$$x_j = \mu g(\mu)x_i - \mu t_i + \mu S_i(x_i) \tag{38}$$

Inserting equation (38) into equation (37) yields

$$t_j = (1/\mu + \mu g(\mu)g(-\mu))\, x_i - \mu(g(-\mu)t_i + \mu g(-\mu)S_i(x_i) + S_j(\mu g(\mu)x_i - \mu t_i + \mu S_i(x_i)) \tag{39}$$

and that equation (39) simplifies to

$$t_j = \mu g(\mu)t_i - k\mu x_i - \mu g(\mu)S_i(x_i) + S_j(\mu g(\mu)x_i - \mu t_i + \mu S_i(x_i)) \tag{40}$$

Equations (38) and (40) are the Lorentz transformation equations if we adopt the standard synchronization, i.e., all $S_i(x_i) = 0$. To see these equations in their more traditional form, we need to change the parameter μ (the proper velocity) into just ordinary velocity ν. These two equally useful parameters are related by the following identities:

$$\nu = \frac{\mu}{\sqrt{1+k\mu^2}} \qquad \mu = \frac{\nu}{\sqrt{1-k\nu^2}} \tag{41}$$

Here is a related identity that is often very useful:

$$\sqrt{1+k\mu^2} = \frac{1}{\sqrt{1-k\nu^2}} \tag{42}$$

$$\text{Let } \ \gamma(\nu) = \frac{1}{\sqrt{1-k\nu^2}}. \ \text{ Then } \ \mu g(\mu) = \gamma(\nu) \ \text{ and } \ \mu = \nu\gamma(\nu). \tag{43}$$

Simply rewriting equation (38) and equation (40) in terms of the new parameter ν completes my derivation of the nonlinear Lorentz transformation:

$$x_j = \gamma(\nu)(x_i - \nu t_i + \nu S_i(x_i)) \tag{44}$$

$$t_j = \gamma(\nu)(t_i - k\nu x_i - S_i(x_i)) + S_j(\gamma(\nu)(x_i - \nu t_i + \nu S_i(x_i))) \tag{45}$$

Nonlinear Lorentz transformations, as displayed above, may seem unwieldy but they do have a beautiful mathematical representation. The beauty should persuade incredulous relativists that my math is correct.

Here is the beautiful result:

Define a function Θ_i from $R^2 \mapsto R^2$ as:

$$\Theta_i \begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} x \\ t + S_i(x) \end{pmatrix} \tag{46}$$

Consequently

$$\Theta_i{}^{-1} \begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} x \\ t - S_i(x) \end{pmatrix} \tag{47}$$

It is very easy to verify that Θ_i and $\Theta_i{}^{-1}$ are inverse functions of each other.
Next, let $L(\nu)$ denote the ordinary linear Lorentz transformation from $R^2 \mapsto R^2$:

$$L(\nu) \begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} \gamma(\nu)(x - \nu t) \\ \gamma(\nu)(t - k\nu x) \end{pmatrix} \tag{48}$$

Finally, define

$$\mathsf{C}_{ji}(\nu) = \Theta_j L(\nu) \Theta_i^{-1} \tag{49}$$

I can now state the theorem: A nonlinear Lorentz transformation is precisely the equation

$$\begin{pmatrix} x \\ t \end{pmatrix}' = \mathsf{C}_{ji}(\nu) \begin{pmatrix} x \\ t \end{pmatrix} \tag{50}$$

The proof requires nothing more than an explicit computation of 3 function compositions:

$$\begin{pmatrix} x \\ t \end{pmatrix}' = \Theta_j L(\nu) \Theta_i^{-1} \begin{pmatrix} x \\ t \end{pmatrix} = \Theta_j L(\nu) \begin{pmatrix} x \\ t - S_i(x) \end{pmatrix} \tag{51}$$

$$= \Theta_j \begin{pmatrix} \gamma(\nu)(x - \nu(t - S_i(x))) \\ \gamma(\nu)(t - S_i(x) - k\nu x) \end{pmatrix} = \Theta_j \begin{pmatrix} \gamma(\nu)(x - \nu t + \nu S_i(x)) \\ \gamma(\nu)(t - k\nu x - S_i(x)) \end{pmatrix} \tag{52}$$

$$= \begin{pmatrix} \gamma(\nu)(x - \nu t + \nu S_i(x)) \\ \gamma(\nu)(t - k\nu x - S_i(x)) + S_j(\gamma(\nu)(x - \nu t + \nu S_i(x))) \end{pmatrix} \tag{53}$$

In component form then,

$$x' = \gamma(\nu)(x - \nu t + \nu S_i(x)) \tag{54}$$

and

$$t' = \gamma(\nu)(t - k\nu x - S_i(x)) + S_j(\gamma(\nu)(x - \nu t + \nu S_i(x))) \tag{55}$$

If the functions $S_i(x_i)$ $(i = 1, 2, 3)$ in equation (44) and (45) only specify an impractical clock synchronization scheme that doesn't affect any of the physics of the universe Ξ_3, or if influential mainstream physicists don't appreciate it when elementary spacetime is defined with utmost generality, then we should set all the synchronization functions $S_i(x_i)$ equal to zero. Therefore, a simple derivation of the Lorentz transformation equations yields:

$$x_j = \gamma(x_i - \nu t_i) \tag{56}$$

$$t_j = \gamma(t_i - k\nu x_i) \tag{57}$$

$$\text{where } \gamma = \gamma(\nu) = \frac{1}{\sqrt{1 - k\nu^2}}. \tag{58}$$

The spacetime structure constant k may be approximated experimentally. It is typically assumed that $k = 1/c^2$, where c is the speed of light.

However, if you are interested in elegant mathematical structures, puzzles and useful theorems, then you might also enjoy generalizing Lorentzian physics just for the fun of it.

$$\text{If } \begin{pmatrix} x \\ t \end{pmatrix}' = \Theta_j L(\nu) \Theta_i^{-1} \begin{pmatrix} x \\ t \end{pmatrix} \tag{59}$$

$$\text{then } \begin{pmatrix} x \\ t \end{pmatrix} = \Theta_i L(-\nu) \Theta_j^{-1} \begin{pmatrix} x \\ t \end{pmatrix}' \tag{60}$$

$$\text{i.e., } \left(\Theta_j L(\nu) \Theta_i^{-1}\right)^{-1} = \Theta_i L(-\nu) \Theta_j^{-1} \tag{61}$$

$$\text{If } \begin{pmatrix} x \\ t \end{pmatrix}' = \Theta_j L(\nu_1) \Theta_i^{-1} \begin{pmatrix} x \\ t \end{pmatrix} \tag{62}$$

$$\text{and } \begin{pmatrix} x \\ t \end{pmatrix}'' = \Theta_k L(\nu_2) \Theta_j^{-1} \begin{pmatrix} x \\ t \end{pmatrix}' \tag{63}$$

$$\text{then } \begin{pmatrix} x \\ t \end{pmatrix}'' = \Theta_k L\left(\frac{\nu_1 + \nu_2}{1 + k\nu_1\nu_2}\right) \Theta_i^{-1} \begin{pmatrix} x \\ t \end{pmatrix} \tag{64}$$

To Poincarize the nonlinear Lorentz transformation to include translations in space and time, as in the Poincaré group, also known as the inhomogeneous Lorentz group, redefine Θ_i to be

$$\Theta_i \begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} x + \alpha_i \\ t + \beta_i + S_i(x) \end{pmatrix} \quad \text{Then} \quad {\Theta_i}^{-1} \begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} x - \alpha_i \\ t - \beta_i - S_i(x - \alpha_i) \end{pmatrix} \tag{65}$$

6 Time Dilation

It was promised in section 2 that this tutorial on spacetime would progress in a series of steps and that each subsequent toy universe would inherit greater complexity. I shall now add to the axioms of Ξ_3 the existence of movable clocks and a clock-carrying traveler that is able to jump instantly from one inertial frame of reference to another.

The usual result to moving clocks in a universe like Ξ_3 is that theoretically, if two or more clocks are synchronized at the same point in space and are then moved apart at a high rate of speed along dissimilar paths and are then moved back together again and compared, then, in general, the clocks will no longer be synchronized. This is not due to the failure of even ideals clock to keep accurate time but is the consequence of the structure of spacetime itself. The computational proof often has a traditional story to illustrate it. Consider the jump diagram below (figure 4). An adventurous clock-carrying traveler is situated at the point $x_2 = U$ in the inertial frame Γ_2. To make the thought experiment more interesting, it is customary to have a stay-at-home twin also located at the same spatial point U in the same inertial frame of reference.

The stay-at-home twin has a stay-at-home clock. The adventurous twin carries his own clock along the path hinted at in figure 4. Here are the details. At the instant that the point $x_3 = J$ on Γ_3 arrives at the point $x_2 = U$ on Γ_2, the adventurous twin will jump from his position at $x_2 = U$ on Γ_2 to $x_3 = J$ on Γ_3 and will ride that inertial frame at velocity ν all the way to his destination $x_2' = M$. The traveler will spend zero time at the point M and will then jump to the nearest point P on Γ_1 for a ride home at velocity $-\nu$. I shall now compute the total travel time that will elapse according to the two clocks.

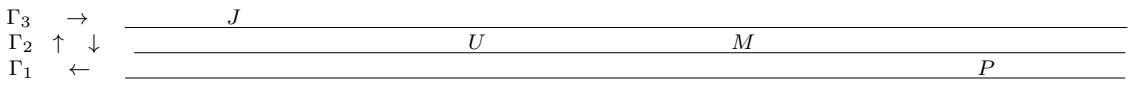


Figure 4

Since a rigorous computation is just as easy to accomplish with the nonlinear time equation as the linear version, for fun, I shall use the nonlinear equation. The nonlinear time equation is defined by solving equation (44) for t_i.

$$t_i = -x_j/\nu\gamma(\nu) + x_i/\nu + S_i(x_i) \tag{66}$$

What does this equation mean? Here is a quick review. The variable ν in equation (66) represents the velocity of the inertial frame of reference Γ_j with respect to the inertial frame Γ_i. According to the problem, $\nu_{12} = \nu_{23} = \nu$ and therefore $\nu_{21} = \nu_{32} = -\nu$. Also recall that the time equation represents multiple equations. To derive time dilation, I shall use the following four instances of equation (66).

$$t_1 = \frac{-x_2}{\nu\gamma(\nu)} + \frac{x_1}{\nu} + S_1(x_1) \tag{67}$$

$$t_2 = \frac{-x_3}{\nu\gamma(\nu)} + \frac{x_2}{\nu} + S_2(x_2) \tag{68}$$

$$t_2 = \frac{-x_1}{-\nu\gamma(-\nu)} + \frac{x_2}{-\nu} + S_2(x_2) \tag{69}$$

$$t_3 = \frac{-x_2}{-\nu\gamma(-\nu)} + \frac{x_3}{-\nu} + S_3(x_3) \tag{70}$$

Naturally $\gamma(\nu) = \gamma(-\nu)$. The rest is an elementary exercise in computing events.

$$\text{Event } (J,U) \text{ in } \Gamma_3 \text{ takes place at time } t_3 = \frac{U}{\nu\gamma(\nu)} - \frac{J}{\nu} + S_3(J) \tag{71}$$

$$\text{Event } (J,M) \text{ in } \Gamma_3 \text{ takes place at time } t_3' = \frac{M}{\nu\gamma(\nu)} - \frac{J}{\nu} + S_3(J) \tag{72}$$

Therefore $t_3' - t_3 =$ (Time at J when the traveler arrives at M) $-$ (Time at J when the traveler jumps to J from U). Thus the traveler's total travel time from U to M is

$$t_3' - t_3 = \left(\frac{M}{\nu\gamma(\nu)} - \frac{J}{\nu} + S_3(J)\right) - \left(\frac{U}{\nu\gamma(\nu)} - \frac{J}{\nu} + S_3(J)\right) = \frac{M-U}{\nu\gamma(\nu)} \tag{73}$$

Because speed in this problem is a fixed constant, it is reasonable to expect that the time it takes the traveler to reach his destination and the time spent on the return trip home are equal but the point of axiomatization is to verify all the details mathematically. Every assumption that can't be proved must be investigated carefully. There is always the possibility for intuition to fail. After all, the time equation represents a counterintuitive view of physics.

When does the traveler arrive at M according to the spacetime clock at M? According to the jump diagram, the events (M,J) and (M,P) happen at the same time. Consequently, this time may be computed in two different ways, using equation (68) and (69), and these two equations must be equal.

Therefore

$$\frac{-P}{\nu\gamma(\nu)} + \frac{M}{\nu} + S_2(M) = \frac{J}{\nu\gamma(\nu)} - \frac{M}{\nu} + S_2(M) \tag{74}$$

$$\text{Equation (74) implies } \; P + J = 2M\gamma(\nu). \text{ I will need that result later.} \tag{75}$$

I now compute the time spent for the traveler on his return trip home. The time in Γ_1 when the traveler jumps to P from M may be computed from the event (P,M).

$$t_1 = \frac{-M}{\nu\gamma(\nu)} + \frac{P}{\nu} + S_1(P) \tag{76}$$

Finally, the time when P arrives at U in Γ_1 is

$$t_1' = \frac{-U}{\nu\gamma(\nu)} + \frac{P}{\nu} + S_1(P) \tag{77}$$

$$\text{Therefore } \; t_1' - t_1 = \left(\frac{-U}{\nu\gamma(\nu)} + \frac{P}{\nu} + S_1(P)\right) - \left(\frac{-M}{\nu\gamma(\nu)} + \frac{P}{\nu} + S_1(P)\right) = \frac{M-U}{\nu\gamma(\nu)} \tag{78}$$

So the total travel time according to the traveler's clock is

$$\frac{M-U}{\nu\gamma(\nu)} + \frac{M-U}{\nu\gamma(\nu)} = \frac{2(M-U)}{\nu\gamma(\nu)} \tag{79}$$

The logic for computing the total stay-at-home time is similar. The time when the traveler at P gets back home, which is the event (U,P) according to the spacetime clock at U, is

$$t_2' = \frac{P}{\nu\gamma(\nu)} - \frac{U}{\nu} + S_2(U) \tag{80}$$

And the starting time for the trip according to the spacetime clock at U is

$$t_2 = \frac{-J}{\nu\gamma(\nu)} + \frac{U}{\nu} + S_2(U) \tag{81}$$

So the total stay-at-home time is

$$t_2' - t_2 = \left(\frac{P}{\nu\gamma(\nu)} - \frac{U}{\nu} + S_2(U)\right) - \left(\frac{-J}{\nu\gamma(\nu)} + \frac{U}{\nu} + S_2(U)\right) = \frac{P+J}{\nu\gamma(\nu)} - \frac{2U}{\nu} \tag{82}$$

It is customary to call the total stay-at-home time Δt and the total traveler's time $\Delta t'$. Therefore

$$\Delta t = \frac{P+J}{\nu\gamma(\nu)} - \frac{2U}{\nu} \tag{83}$$

$$\Delta t' = \frac{2(M-U)}{\nu\gamma(\nu)} \tag{84}$$

Next, I insert the expression for $P + J$ from equation (75) into equation (83) and get

$$\Delta t = \frac{2(M-U)}{\nu} \tag{85}$$

Finally, I compare equation (84) and (85) and that completes my rigorous derivation of the time dilation formula from the nonlinear time equation.

$$\Delta t' = \Delta t/\gamma(\nu) \tag{86}$$

7 Related Scientific Publications

1. Helmut Günther, *A note to the linearity proof of Lorentz transformation* (2003)

2. Piotr Kosiński, *Note on clock synchronization and Edwards transformations* (2006)

3. Guido Rizzi, Matteo Luca Ruggiero, Alessio Serafini, *Synchronization Gauges and the Principles of Special Relativity* (2004)