The Axiomatization of Physics - Step 1
A Derivation of the Lorentz Transformation From a Simple Definition of Time

Eugene Shubert
everythingimportant.org

updated January 11, 2009

Abstract

The aim of *The Axiomatization of Physics – Step 1* is to define physics and then axiomatize an increasingly complex hierarchy of mathematical models of spacetime. We begin by building toy models of spacetime and then progress slowing so that high school students who have mastered algebra can understand the theory of relativity easily. Spacetime in high dimensions begins by defining space and time in one spatial dimension.

We will derive the fundamental equations of spacetime without using unnecessary axioms. Einstein's two famous postulates for the special theory of relativity will not be assumed. There is a more fundamental theory than special relativity. One consequence is that Einstein's theory could be false and our theory could be true. However, if Einstein's theory is true, then our theory is also true.

The essence of spacetime is contained in the Lorentz transformation equations. For fun, we will compute time dilation using nonlinear Lorentz-equivalent transformations, consider physics in a spherical universe, and derive the correct transformation equations in the case where motion faster than light is possible. Our fundamental axioms are:

1. Newton's first law of motion.

2. There is a simple definition of clock time for each point in an inertial frame of reference. (i) This time is mathematically well-defined. (ii) Time is defined the same way in all inertial frames of reference.

1 Why Axiomatize Physics?

In the history of physics, ideas that were once seen to be fundamental, general, and inescapable parts of the theoretical framework are sometimes later seen to be consequent, special, and but one possibility among many in a yet more general theoretical framework. ... Examples are the earth-centered picture of the solar system, the Newtonian notion of time, the exact status of the laws of thermodynamics, the Euclidean laws of spatial geometry, and classical determinism. In view of this history, it is appropriate to ask of any current theory "which ideas are truly fundamental and which are 'excess baggage'." J.B. Hartle, *Classical physics and Hamiltonian quantum mechanics as relics of the Big Bang*, Physica Scripta T 36 (1991), 228-236.

"The reductionist approach — explaining physical phenomena in terms of simple, mathematically precise, quantities — has been extraordinarily successful in almost all areas of physics. It goes against everything we have learned about

> nature to propose a theory in which complicated macroscopic objects, whose precise definition must ultimately be arbitrary, are fundamental quantities.” A. Kent, *Against many-worlds interpretations*, Int. J. Mod. Phys. 5 (1990), 1745-1762.
>
> “A great physical theory is not mature until it has been put in a precise mathematical form, and it is often only in such a mature form that it admits clear answers to conceptual problems.” A. S. Wightman, *Hilbert’s sixth problem: mathematical treatment of the axioms of physics*, in: Proc. Sympos. Pure Math., Vol. 28, AMS, 1976, pp. 147-220.

2 David Hilbert’s Philosophy of Physics

Physics should evolve from a small number of axioms

“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. [1].

The definition of physics

Physics is the study of all mathematically consistent universes. A universe is a mathematical model that describes spacetime, matter, energy and their interactions. Think of each model universe as filling one page in the atlas of all possible universes.

Within the mathematicians’ worldview, physics is a mere subdiscipline in mathematics. Possibilities are greater than what actually exists. Mathematics is true throughout the multiverse. For us, the atlas includes all mathematically consistent fantasy universes that represent even toy features of space, time and matter.

“Philosophy is written in this grand book, the universe, ... But the book cannot be understood unless one first learns to comprehend the language and read the characters in which it is written. It is written in the language of mathematics.” — Galileo Galilei.

3 The Coordinatization of Space

Imagine two isolated points. Two points determine a line. Think of one point as being on your left. Assign it the number 0. Think of the other point as being on your right. Assign it the number 1. Now assume that there is a way to measure distance between points. Place the number 2 to the right of 1 the same distance as 0 is from 1. Imagine that procedure being repeated indefinitely for all the whole numbers. Notice that for every number, if we add 1 to it, we move to the right one step. Logically then, for every number, if we subtract 1 from it, we move to the left one step. That’s how we assign numbers to the left of 0. Here’s the picture (it’s up to you to imagine the integers extending indefinitely in both directions):

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

Figure 1

In mathematics, fractions are called rational numbers. They are easy to visualize. To understand what a quotient of two integers means geometrically, it suffices to consider the line segment [0,1]. The fraction 1/2 is the point halfway between 0 and 1. Halfway means that there are two equal distances. The two fractions 1/3, 2/3 imply 3 equal distances. The distance between 0 and 1/3, 1/3 to 2/3, and 2/3 to 1 are all equal. For any number n, it's easy to conceive of $n+2$ equally spaced points on the line segment [0,1]. These points are assigned the location coordinates $0 = 0/(n+1)$, $1/(n+1)$, $2/(n+1)$, $3/(n+1)$, $\ldots$ $(n+1)/(n+1) = 1$.

Intuitively, every point on the line is arbitrarily close to some fraction. We only need to make the number n large enough.

In ancient times, there were philosophers who believed that every point on a number line may be represented as the quotient of two integers. [2]. It was the Greek philosopher Hippasus of Metapontum, born circa 500 B.C. in Magna Graecia, who proved that presupposition that to be false. [3].

A consequence of the Pythagorean theorem is that the hypotenuse of a right triangle with equal sides of unit length is $\sqrt{2}$. If we place this hypotenuse on the number line with one end on zero 0 and the other end on the positive numbers, then that second endpoint will specify a point on the number line with the coordinate value $\sqrt{2}$. It is easy to prove that the $\sqrt{2}$ is not a quotient of two integers. Numbers like that, which evidently exist but are not rational, are said to be irrational.

Real numbers are a mildly abstract generalization of practical measurements. Real numbers are easy to grasp intuitively. The first step is to define convergence from a mathematical definition of closeness of points. Then the essential idea is remarkably simple: It is possible to get arbitrarily close to any point on the number line by specifying a properly behaved infinite sequence of rational numbers. For instance, the $\sqrt{2}$ may be defined as the limit of the infinite sequence 1.4, 1.41, 1.414, 1.4142, 1.41421, 1.414213, 1.4142135, ... Notice that the square of these terms is the sequence 1.96, 1.9881, 1.999396, 1.99996164, 1.9999899241, ... which is converging toward the number 2. It is a standard beginning point in math to define real numbers in terms of Cauchy sequences of rational numbers. [4].

It is often the case where real numbers are not thought of as representing points on a line but only as an abstract set having algebraic properties, where any two elements of the set can be added together or multiplied, and so on. When the real number line is in view, and the distance between points is important, then the real number line is often called a one-dimensional Euclidean space. A Euclidean space of one dimension has the following definition of distance between points: If r and s are real numbers, then the distance between r and s is defined by the absolute value of their numerical difference. In mathematical parlance, $d(r,s) = |r - s|$.

4 My First Toy Universe

Spacetime is any mathematical model that combines space, time and motion into a single construct. Amazingly, all the fantastic spacetime models that we will define in this tutorial are built upon Ξ_2. This the simplest spacetime imaginable. This first toy universe is a required object of study.

The importance of creating toy universes is they help us to identify and develop all the logical consequences of truly fundamental ideas. The universe Ξ_2 is remarkably simple. Within it, only two kinds of motion are possible. Can you imagine how simple the universe would be if there were only two possible speeds?

It is common in mathematics and physics to call discrete levels of energy, position or configuration, *states.* Following the usual conventions, we shall refer to the two kinds of motion that are possible in the Ξ_2 universe as state S_1 and state S_2 respectively. We may think of S_1 as a stationary state and S_2 as the moving state but these labels are arbitrary. As we will see in a moment, Ξ_2 has no intrinsic properties that distinguish between moving and not moving.

It is now time to grasp why Ξ_2 is such a natural beginning point in the study of spacetime. Mathematicians love to generalize and this is where we begin.

The Simplest Spacetime Imaginable

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?

Sir Isaac Newton explained the simplest kind of motion. Newton's first law of motion states: "An object at rest tends to stay at rest and an object in motion tends to stay in motion with the same speed and in the same direction unless acted upon by an unbalanced force."

The easiest space to work with mathematically is the very familiar one-dimensional Euclidean space, popularly known as the number line. It is helpful to visualize this simple space as an infinitely long ruler.

If Ξ_2 is to have two states of motion in a simple one-dimensional space, then Ξ_2 consists of two one-dimensional Euclidean spaces. Each state of motion S_i $i = 1, 2$ is assigned its own Euclidean space.

We are now ready for a definition of Ξ_2. See *Figure 2.* Imagine that the two abstract lines Γ_1 and Γ_2 are pristine, frictionless 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. We are also to imagine Γ_2 sliding on Γ_1 at a constant velocity.

$$\begin{array}{rl} \Gamma_2 \rightarrow & \underline{\cdots\ \text{-9}\ \ \text{-8}\ \ \text{-7}\ \ \text{-6}\ \ \text{-5}\ \ \text{-4}\ \ \text{-3}\ \ \text{-2}\ \ \text{-1}\ \ 0\ \ 1\ \ 2\ \ 3\ \ 4\ \ 5\ \ 6\ \ 7\ \ 8\ \ 9\ \cdots} \\ \Gamma_1 \leftarrow & \underline{\cdots\ \text{-9}\ \ \text{-8}\ \ \text{-7}\ \ \text{-6}\ \ \text{-5}\ \ \text{-4}\ \ \text{-3}\ \ \text{-2}\ \ \text{-1}\ \ 0\ \ 1\ \ 2\ \ 3\ \ 4\ \ 5\ \ 6\ \ 7\ \ 8\ \ 9\ \cdots} \end{array}$$

Figure 2

Because the number 2 is to the right of the number 1 on the number line, we will assign a meaning to the indices 1 and 2 and declare that Γ_2 is moving to the right of Γ_1.

Note that I have referred to "constant velocity" without defining the term. I will do so now by first defining clock time. For us, the simple definition will be a fundamental axiom. More importantly, it will assign a physical interpretation to the universe Ξ_2. From that clear and simple definition we will conclude that "proper velocity" is equal and opposite so that the points of Γ_2 move equal distances in equal times along Γ_1 and that the points of Γ_1 move equal distances in equal times in the opposite direction along Γ_2.

The Definition of Clock Time

The most distinct physical characteristic or property of the universe Ξ_2 is that a unique kind of motion exists. In addition to the spatial structure of this universe, there is also an interesting time structure as well. Observe 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. By definition, 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.

There is now a mathematical definition of clock time at each point of space. And this clock structure is easy to understand. Little children know intuitively that a tiny arrow that moves steadily along a continuum of numbers is a clock.

A clear definition should have a straightforward interpretation. From the definition of clock time, please note that all the mathematical clocks of Γ_2 assign a direction to time as progressing from smaller to larger numbers whereas all the mathematical clocks on Γ_1 run backwards. Is there any physical reason to reject this oddity? If anyone were to claim that this is a paradox, I would answer by asserting a pragmatic principle: If a clock is running backwards at the same rate that another clock is moving forward, then a sign change is recommended for one of the clocks.

It is important to know the difference between an arbitrary convention and a law of physics. It's an arbitrary choice that we assign the direction of time as not counting down toward more negative numbers. We will adopt the popular convention that all clocks should measure time as continually progressing toward larger numbers. So, to get all our mathematical clocks of Γ_1 to behave like the mathematical clocks of Γ_2, we will define the clock time at the general point x_1 of Γ_1 mathematically to be the negative of whatever number the arrow at x_1 is pointing to.

It is necessary to translate the meaning of these mathematical clocks, which we've defined up till now with pictures, into the preferred, universal language of functions. Recall the definition of a function: In mathematics, a function relates each of its inputs to exactly one output. A standard notation for the output of the function f with the input x is f(x).

In terms of functions then, $t_1 = f_1(x_1, x_2) = -x_2$ and $t_2 = f_2(x_1, x_2) = x_1$. The interpretation has already been given. Here is a restatement that might help to clarify the meaning of the mathematical symbols. The functions f_1 and f_2 specify every detail about the mathematical clocks on Γ_1 and Γ_2, respectively. The input for these functions is any ordered pair of numbers (x_1, x_2) where x_1 is a point on Γ_1 and x_2 is a point on Γ_2. These tuples say everything about Ξ_2. The ordered pair (x_1, x_2) is the event where x_1 meets (touches, flies by or passes through) x_2, however you want to say it. The function f_1 inputs the event (x_1, x_2) and assigns an output t_1 for the mathematical clock at x_1. The output says the event happens at time $t_1 = -x_2$. Likewise, the function f_2 inputs the same event (x_1, x_2) and assigns an output t_2 for the mathematical clock at x_2. This output for the mathematical clock on Γ_2 says the event happens at time $t_2 = x_1$.

Because mathematical clocks express time in terms of real numbers, we are then free to select whatever scale we desire for time. If a and b are positive real numbers, then $t_1 \mapsto t_1/a$ and $t_2 \mapsto t_2/b$ is merely changing the time scale such as hours to minutes or hours to seconds. Since we aim at defining time with consistent equations, point by point, everywhere in the universe, we will insist that $a = b$. For the purpose of noting the similarity between Ξ_2 and real world physics, we will assign the symbol μ to represent an appropriate time scaling factor and note that μ has dimensions of velocity.

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 we 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.

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 our 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}. Please note that $\mu_{21} = -\mu_{12}$ and that μ_{12} is a positive number.

I believe that we have now completed 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 definitions and observations so that the reader will recognize the connection. 1) In other universes, the moving lines Γ_1 and Γ_2 of Ξ_2 are called inertial frames of reference. 2) It should be emphasized that, in our construction of Ξ_2, we invoked no axiom presupposing the existence of a cosmic everywhere present "now." 3) It should also be recognized that we did not presuppose that the number lines Γ_1 and Γ_2 are equal in scale.

$$\Gamma_2 \rightarrow \underline{\ \cdots\ \text{-9}\ \ \text{-8}\ \ \text{-7}\ \ \text{-6}\ \ \text{-5}\ \ \text{-4}\ \ \text{-3}\ \ \text{-2}\ \ \text{-1}\ \ 0\ \ 1\ \ 2\ \ 3\ \ 4\ \ 5\ \ 6\ \ 7\ \ 8\ \ 9\ \cdots\ }$$
$$\Gamma_1 \leftarrow \underline{\ \cdots\ \text{-9}\ \ \text{-8}\ \ \text{-7}\ \ \text{-6}\ \ \text{-5}\ \ \text{-4}\ \ \text{-3}\ \ \text{-2}\ \ \text{-1}\ \ 0\ \ 1\ \ 2\ \ 3\ \ 4\ \ 5\ \ 6\ \ 7\ \ 8\ \ 9\ \cdots\ }$$

Figure 3

5 My Second Toy Universe

From our study of Ξ_2 we learned that moving coordinate systems may be interpreted as an infinite number of mathematical clocks. We 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}. We 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	$\cdots$	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	$\cdots$
Γ_2	$\cdots$	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	$\cdots$
Γ_1	$\cdots$	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	$\cdots$

Figure 4

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 Lorentizian 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 we 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. We will now use our preference for g being 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 $\mu_{12} = \mu$, $\mu_{23} = \nu$ and $\mu_{13} = \mu \oplus \nu$. Equation (18) and (28) suggests the possibility that

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

Is this a sensible formula for adding proper velocities? If μ is positive and we add to it an infinitesimally small positive ν, 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, the correct formula for $\mu \oplus \nu$ is

$$\mu \oplus \nu = \mu\sqrt{1 + k\nu^2} + \nu\sqrt{1 + k\mu^2} \tag{30}$$

Now consider the possibility that the spacetime structure constant k is negative. If k were negative, there would be a limit to proper velocity; otherwise, we could add two large proper velocities according to equation (30) and end up taking the square root of a negative number.

The safe interval for adding proper velocities in the case of a negative k is easily seen to be $[-\zeta, \zeta]$, where the maximum proper velocity $\zeta = 1/\sqrt{-k}$. More compactly, if μ, ν $\in [-\zeta, \zeta]$, then $\mu \oplus \nu \in [-\zeta, \zeta]$. One curious feature of a $k < 0$ universe is that $\zeta \oplus \zeta = 0$. This is no contradiction; it is just never seen in our ordinary experience. The real problem with this universe is that adding extra spatial dimensions to it also permits violations of causality.

These facts indicate that no significantly realistic spacetime exists having a negative spacetime structure constant. So, to emphasize that k is nonnegative, we write $k = 1/c^2$. At this point, the only meaning that I have ascribed to c is that c is a real number greater than zero. If the case where $k = 0$ needs to be considered, we may adopt the convention that $c = \infty$ in that instance. Equations (30) and (28) are therefore

$$\mu \oplus \nu = \mu\sqrt{1+\nu^2/c^2} + \nu\sqrt{1+\mu^2/c^2} \tag{31}$$

$$g(\mu) = \epsilon + \frac{\sqrt{1+\mu^2/c^2}}{\mu} \tag{32}$$

It turns out that there is a physically meaningful property associated with the constant c but not so for the parameter ϵ. A nonzero epsilon universe is only a reference to an unusual way that clocks may be synchronized.

At this stage we've completely solved the problem. Here is how you get the Lorentz transformation. First, recall that equation (7) represents multiple equations. We need the following two:

$$t_i = -x_j/\mu + g(\mu)x_i \tag{33}$$

$$t_j = x_i/\mu + g(-\mu)x_j \tag{34}$$

Equation (33) is equation (7) without the subscripts on the proper velocity. You can write them in if you like or just remember that $\mu_{ij} = \mu$. I thought that we should continue with the simpler notion. Equation (34) is equation (7) with i replacing j and j replacing i and remembering that $\mu_{ij} = -\mu_{ji}$.

Next, solve equation (33) for x_j. This gives us

$$x_j = \mu g(\mu)x_i - \mu t_i \tag{35}$$

Inserting equation (35) into equation (34) and simplifying the results gives us

$$t_j = -\mu g(-\mu)t_i + \left(\frac{1}{\mu} + \mu g(\mu)g(-\mu)\right)x_i \tag{36}$$

Calculating the coefficients of x_i and t_i explicitly for equations (35) and (36) is straightforward from equation (32):

$$x_j = \left(\epsilon + \sqrt{1+\mu^2/c^2}\right)x_i - \mu t_i \tag{37}$$

$$t_j = \left(\sqrt{1+\mu^2/c^2} - \epsilon\mu\right)t_i - \left(\epsilon^2 - 1/c^2\right)\mu x_i \tag{38}$$

Not surprisingly, we can solve equations (37) and (38) for x_i and t_i in terms of x_j and t_j by using the symmetry property mentioned previously: interchange i and j, and u with $-u$. Therefore:

$$x_i = \left(\epsilon + \sqrt{1+\mu^2/c^2}\right)x_j + \mu t_j \tag{39}$$

$$t_i = \left(\sqrt{1+\mu^2/c^2} + \epsilon\mu\right)t_j + \left(\epsilon^2 - 1/c^2\right)\mu x_j \tag{40}$$

Equations (37) and (38) are the Lorentz transformation equations if we set $\epsilon = 0$. To see these equations in their more traditional form, we need to change the parameter μ (the proper velocity) into just ordinary velocity v. These two equally useful parameters are related by the following identities:

$$v = \frac{\mu}{\sqrt{1+\mu^2/c^2}} \qquad \mu = \frac{v}{\sqrt{1-v^2/c^2}} \tag{41}$$

Here is a related identity that is often very useful:

$$\sqrt{1+\mu^2/c^2} = \frac{1}{\sqrt{1-v^2/c^2}} \tag{42}$$

A straightforward and elementary substitution gives us the final form:

$$x_j = \frac{(1+\epsilon v)x_i - vt_i}{\sqrt{1-v^2/c^2}} \tag{43}$$

$$t_j = \frac{(1-\epsilon v)t_i + (\epsilon^2 - 1/c^2)vx_i}{\sqrt{1-v^2/c^2}} \tag{44}$$

Since the technical meaning of ϵ in these equations only refers to an impractical clock synchronization scheme that doesn't affect any of the physics of the universe Ξ_3, we can set it equal to zero. Therefore, we have finally arrived at the Lorentz transformation equations:

$$x_j = \gamma(x_i - vt_i) \tag{45}$$

$$t_j = \gamma(t_i - vx_i/c^2) \tag{46}$$

$$\text{where } \gamma = \gamma(v) = \frac{1}{\sqrt{1-v^2/c^2}}. \tag{47}$$

6 Can the Lorentz Transformation be Simplified?

In this section, I shall start with the Lorentz transformation and the clock synchronization scheme it implies and then reset all clocks in all but one inertial frame of reference. From there I shall rescale the meaning of distance between points *frame by frame* with a suitable scaling factor. I shall also adjust clock rates by applying a similar suitable change of time scale and thus finally arrive at the traditional and easy to understand Galilean group of transformations:

$$\tilde{X} = X - vT \tag{48}$$

$$\tilde{T} = T \tag{49}$$

The equations of the Lorentz transformation may be written as

$$x' = \gamma(x - vt) \tag{50}$$

$$x \ = \gamma(x' + vt') \tag{51}$$

These two equations are algebraically equivalent to our definition of time

$$t \ = -x'/v\gamma + x/v \tag{52}$$

$$t' = \ x/v\gamma - x'/v \tag{53}$$

Let Γ', $\Gamma_{\prime}$ or Γ_p denote the primed frame and Γ the unprimed frame. According to equation (2) in our study of Ξ_2, we may reset clock time in the primed frame by adding any function of x' to the existing synchronization function. Let the change in clock synchronization be $t' \mapsto t' + q(v)x'$. The equation for the new time T' will then become

$$T' = x/v\gamma - x'/v + q(v)x' \tag{54}$$

Inserting x' from equation (50) into the preceding equation yields

$$T' = x/v\gamma - \gamma(x - vt)\big(1/v - q(v)\big) \tag{55}$$

Multiplying and then rearranging terms to see the coefficients of x and t yields

$$T' = \gamma(v)\big(1 - vq(v)\big)t \, - \, \gamma(v)\big(v/c^2 - q(v)\big)x \tag{56}$$

If we set $q(v) = 0$ in the preceding equation then $T' = \gamma(v)(t - vx/c^2)$, as expected. For a more enlightening and easy to interpret form, select $q(v)$ so that the coefficient of x vanishes. Therefore $q(v) = v/c^2$ and we finally arrive at the simpler Lorentz-equivalent equations

$$x' = \gamma(x - vT) \tag{57}$$

$$T' = T/\gamma \tag{58}$$

Equations (57) and (58) are called the Tangherlini transform. The physics of this transform is easy to justify. It is nothing but an elementary consequence of resetting clocks. And the simplification is that events that are considered simultaneous in one inertial frame will be simultaneous in all inertial frames. Now let $\tilde{X} = x'/\gamma$ and $\tilde{T} = \gamma T'$. Then the Galilean equations $\tilde{X} = X - vT$ and $\tilde{T} = T$ follow immediately.

The idea of rescaling clock time and distance *frame by frame* to make all spacetime events describable by the Galilean equations is physically permissible for one simple reason. The essential purpose of a spacetime transformation is to have a set of equations that convert clock times and space coordinates back and forth between different observers in different inertial frames of reference.

These are important facts not to be forgotten:

1. The Lorentz transform standing alone, as an axiom, comes with its own frame-dependent definition of time and distance. Without additional assumptions, there appears to be no way to determine how relative space and time scales differ frame by frame. We conclude that there is a principle of space incommensurability and time incommensurability. We have assumed that all distances and all times only share a *mathematical similarity*, respectively. Thus, if we begin with the Lorentz transformation in one spatial dimension and interpret all the inertial frames of reference as moving lines with clocks attached at each point, then it is impossible to prove that all the lines use the same unit of distance and it's impossible to prove that all clocks tick at identical rates.

2. If you perform a careful review, it will confirm a very unsurprising fact. No hidden or explicit assumption presupposing the existence of a universally applicable definition of a fundamental distance was used in our derivation of the Lorentz transformation. And there was no use made of a fundamental unit of time. The derivation presupposed no transportable metersticks, movable physical clocks or any presumed clock synchronization scheme.

3. The Lorentz transformation alone doesn't imply time dilation. Other axioms must be added to our extraordinarily weak axiom set for time dilation computations to follow logically. These supplemental axioms will be enumerated in section 8.

What comes next?

This is a work in progress. Check back for major additions, revisions and developments.