Pergamon

Stud. Hist. Phil. Mod. Phys., Vol. 32, No. 1, pp. 83–99, 2001
© 2001 Elsevier Science Ltd. All rights reserved.
Printed in Great Britain
1355-2198/01 $ - see front matter

Spacetime without Reference Frames: An Application to the Velocity Addition Paradox

T. Matolcsi*, A. Goher [†]

Much conceptualisation in contemporary physics is bogged down by unnecessary assumptions concerning a specific choice of coordinates which often leads to misunderstandings and paradoxes. Considering an absolute (coordinate-free) formulation of special relativistic spacetime, we show clearly that the velocity addition paradox emerged because the use of coordinates obscures that the space of relativistic observers is 'more relative' than the space of non-relativistic observers. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Special Relativity; Observers; Lorentz Boost; Thomas Rotation.

1. Introduction

Paradoxes in physics appear mostly because tacit assumptions, true in some special domain, are applied in another domain where they are not valid. The only tool for ruling out tacit assumptions and the possibility of paradoxes is to construct complete mathematical models for physical objects, mathematical models in which every notion related to the object in question is formalised mathematically.

The usual treatment of special relativity is in terms of coordinates with respect to inertial observers. Inertial observers, coordinate axes and several other

(*Received 2 January 2000; revised 19 June 2000*)
* Department of Applied Analysis, Eötvös Loránd University, 1088 Budapest, Múzeum krt. 6-8, Hungary (*e-mail*: matolcsi@ludens.elte.hu).
[†] Finit-2 Limited, 1043 Budapest, Martirok u 58, Hungary (*e-mail*: goher@finit.hu).

PII: S1355-2198(00)00037-X

notions in that treatment are intuitive—not formalised—notions. Incorrect tacit assumptions deriving from the use of coordinates have resulted in new paradoxes in the last decade: the velocity addition paradox (Mocanu, 1992) and the light speed paradox (Selleri, 1997).

It has been shown (Matolcsi, 1998) that the light speed paradox is based on the tacit assumption that a particular value of an unmeasured relative velocity makes sense, an assumption which is acceptable only in the non-relativistic case where absolute time exists. Taking a positivistic approach in which only directly measured quantities make sense, we realise that a particular value of a relative velocity in the relativistic case presupposes a synchronisation. The use of coordinates hides that *the relative velocity of a material point or a light signal with respect to an observer depends on the synchronisations and thus has meaning only if a synchronisation is given*—the paradox emerged because this fact was neglected.

The velocity addition paradox was discussed by Ungar (1989) and Good (1995). However, their arguments and mathematical formulas in terms of coordinates do not give an evident physical explanation of the paradox, though it became clear that the paradox was related somehow to the Thomas rotation.

We shall now demonstrate that the paradox is based on the tacit assumption that the space of observers in special relativity is as relative as the space of observers in the non-relativistic case. The use of coordinates hides that *the space of relativistic observers is 'more relative' than the space of non-relativistic observers*—disregarding this fact one encounters the velocity addition paradox.

It is frequently emphasised, as a main feature of special relativity, that time is relative (not absolute); but, in general, no hint is made regarding space. Of course, space is related to observers, too; this holds even in non-relativistic physics,[1] which is reflected by the Galilean transformation $t' = t$, $x' = x - vt$. The Lorentz transformation in its usual form $t' = \kappa(t - xv)$, $x' = \kappa(x - vt)$ shows that both time and space are relative but it produces the impression that the degree of relativity of space is the same as in the non-relativistic case.

As mentioned previously, tacit assumptions and the possibility of paradoxes can be ruled out by constructing mathematical models of spacetime in which every notion is mathematically defined.

General relativity is a mathematically developed physical theory, whose modern setting is based on the global objects of manifolds (see Wald, 1983): vector fields, differential forms, covariant derivations, etc. These global objects can be called *absolute* from a physical point of view because they are not related to observers (reference frames, coordinate systems). In the last years several attempts have been made to formalise non-relativistic spacetime in a similar mathematical way (Appleby and Kadianakis, 1986; Rodrigues *et al.*, 1995), which clearly shows the demand for an absolute formulation of physical

[1] A trivial fact from a physical point of view: the space relative to someone sitting near a road is constituted by the houses, the trees, etc.; a car travelling on the road is not a part of that space; the seats, the dashboard, etc. constitute the space for someone sitting in the car.

theories. It is evident nowadays that the mathematical structure of spacetime can (and must) be formulated without observers.

Let us take this opportunity to emphasise the following: the frequently stated assertion that special relativity is the theory of inertial observers and general relativity is the theory of arbitrary observers (see e.g. Møller, 1972) is to be replaced by the statement that general relativity describes gravity and special relativity concerns the lack of gravity (Synge, 1955, 1964).

A *general relativistic spacetime model* is a triplet $(M, \mathbf{I}, g)$ where M is a four-dimensional manifold, $\mathbf{I}$ is the measure line of spacetime distances and g is an $\mathbf{I}\otimes\mathbf{I}$ valued Lorentz form on M. A *special relativistic spacetime model* is a particular general relativistic one in which M is an affine space and g is constant (as Weyl (1922) stated some seventy years ago).

A thorough treatment of non-relativistic spacetime models and special relativistic spacetime models based on the affine structure has been given by Matolcsi (1984, 1993). These models will serve to clarify how one is misled by the use of coordinates in the velocity addition paradox.

2. Fundamentals of Spacetime Models

We recapitulate briefly the fundamental notions of spacetime models as given by Matolcsi (1993).

An affine space V over the vector space $\mathbf{V}$ is a non-void set with a given map $V \times V \to \mathbf{V}$, $(x, y) \mapsto x - y$ such that $(x - y) + (y - z) + (z - x) = 0$ for all $x, y, z \in V$ and $V \to \mathbf{V}$, $x \mapsto x - y$ is bijective for all $y \in V$.

A map $A{:}V \to U$ between affine spaces is affine if there is a linear map $\mathbf{A}{:}\mathbf{V} \to \mathbf{U}$ such that $Ax - Ay = \mathbf{A}(x - y)$ for all $x, y \in V$.

The tensorial quotient of a vector space $\mathbf{V}$ by a one-dimensional vector space $\mathbf{I}$ is defined to be a vector space $\mathbf{V}/\mathbf{I}$ whose dimension equals the dimension of $\mathbf{V}$ and the map $\mathbf{V} \times (\mathbf{I}\backslash\{0\}) \to \mathbf{V}/\mathbf{I}$, $(\mathbf{x},\mathbf{t}) \mapsto \mathbf{x}/\mathbf{t}$ obeys the usual rules of division by numbers.

2.1. Non-relativistic spacetime models

A *non-relativistic spacetime model* is a quintet $(M, I, \tau, \mathbf{D}, \cdot)$ where:

– M is a four-dimensional oriented affine space (*spacetime*) (over the vector space $\mathbf{M}$),

– I is a one-dimensional oriented affine space (*absolute time*) (over the vector space $\mathbf{I}$, the *measure line of time periods*), $\tau{:}M$ to I is affine surjection (*time evaluation*) (over the linear map $\boldsymbol{\tau}$: $\mathbf{M}$ to $\mathbf{I}$),

– $\mathbf{D}$ is a one-dimensional oriented vector space (*measure line of distances*),

– $\cdot{:}\mathbf{E} \times \mathbf{E} \to \mathbf{D}\otimes\mathbf{D}$ is a positive definite symmetric bilinear map (*Euclidean structure*), where

$$\mathbf{E} := \operatorname{Ker} \boldsymbol{\tau} \tag{1}$$

is the (three-dimensional) linear subspace of *spacelike vectors*.

A *worldline function* is a twice continuously differentiable function $r:I \to M$ such that $\tau(r(t)) = t$ for all $t \in I$. A *worldline* is the range of a worldline function; a worldline is a curve in M. A worldline represents the history of a classical masspoint.

Absolute velocities are the derivatives of worldline functions; their set is

$$V(1) := \left\{ \mathbf{u} \in \frac{\mathbf{M}}{\mathbf{I}} \middle| \tau(\mathbf{u}) = 1 \right\}. \tag{2}$$

Given a $\mathbf{u} \in V(1)$, every spacetime vector can be uniquely split into the sum of a timelike vector, parallel to $\mathbf{u}$, and a spacelike vector; in other words, we can give the $\mathbf{u}$-splitting

$$\mathbf{M} \to \mathbf{I} \times \mathbf{E}, \quad \mathbf{x} \mapsto (\tau(\mathbf{x}), \pi_{\mathbf{u}}(\mathbf{x})), \tag{3}$$

where

$$\pi_{\mathbf{u}}(\mathbf{x}) := \mathbf{x} - \tau(\mathbf{x})\mathbf{u}. \tag{4}$$

An observer,[2] physically, is a collection of masspoints (a continuous medium) which could be given as a collection of worldlines. It is more convenient, however, to describe it by the tangent vectors of the corresponding worldlines. Thus we accept that an *observer* is a velocity field, i.e. a smooth mapping

$$\mathbf{U} : M \to V(1). \tag{5}$$

The maximal integral curves of such a vector field are worldlines, representing the histories of the material points constituting the observer; thus it is quite evident that an integral curve of $\mathbf{U}$ is a space point of the observer; the set of the maximal integral curves of $\mathbf{U}$ is the *space* of the observer, or $\mathbf{U}$-space in brief. This is the most important—but trivial—fact concerning observers: *a space point of an observer is a curve in spacetime.*

Observers and their spaces are well-defined simple and straightforward notions. *The spaces of different observers are evidently different.*

An observer is *inertial* if it is a constant mapping. We shall consider only inertial observers (so the term inertial will occasionally be omitted). An inertial observer can be given by its constant value, that is why we find it convenient to say 'an observer $\mathbf{u} \in V(1)$'. The space points of the inertial observer $\mathbf{u}$ are straight lines in spacetime, parallel to $\mathbf{u}$.

[2] The terms 'observer' and 'reference frame' are used in several senses in the literature. 'Observer' frequently means a single worldline, and 'reference frame' refers to a collection of worldlines; but 'observer' can refer to a collection of worldlines, too, and 'reference frame' involves, implicitly or explicitly, coordinates or a basis ('tetrad') (Synge, 1955, Secs I.9, I.8, II.7; Synge, 1964, Sec. III.5). Since worldline is a customary notion, there is no need to rename it 'observer', so a collection of worldlines is accepted to be an observer and the name 'reference frame' will be retained for observers with given coordinates (see Matolcsi, 1993).

The vector between two **u**-space points q_1 and q_2 is defined to be the vector between simultaneous world points of the straight lines in question, i.e. the space of the observer **u**, endowed with the subtraction

$$q_1 - q_2 := x_1 - x_2 \quad (x_1 \in q_1,\ x_2 \in q_2,\ \boldsymbol{\tau}(x_1 - x_2) = 0) \tag{6}$$

is an affine space over the vector space **E**; note that we have

$$q_1 - q_2 = \boldsymbol{\pi}_{\mathbf{u}}(x_1 - x_2) \quad (x_1 \in q_1, x_2 \in q_2). \tag{7}$$

Thus in a non-relativistic spacetime model the *different spaces* of different inertial observers are different affine spaces over the *same* vector space. Consequently, it has an '*a priori* meaning' that a vector in the space of an observer equals a vector in the space of another observer, and this equality is a symmetric and transitive relation.

2.2. Special relativistic spacetime models

A *special relativistic spacetime model* is a triplet $(M, \mathbf{I}, \cdot)$ where:
– M is a four-dimensional oriented affine space (*spacetime*) (over the vector space **M**),
– **I** is a one-dimensional oriented vector space (*measure line of spacetime distances*),
– $\cdot : \mathbf{M} \times \mathbf{M} \to \mathbf{I} \otimes \mathbf{I}$ is a *Lorentz product* endowed with an *arrow orientation* which determines the *future directed* timelike and lightlike vectors.

The set of *absolute velocities* is

$$V(1) := \left\{ \mathbf{u} \in \frac{\mathbf{M}}{\mathbf{I}} \,\middle|\, \mathbf{u} \cdot \mathbf{u} = -1, \ \mathbf{u} \text{ is future directed} \right\}. \tag{8}$$

A *worldline function* is a twice continuously differentiable function $r : I \to M$ such that $\dot{r}(t) \in V(1)$ for all $t \in I$. A *worldline* is the range of a worldline function; a worldline is a curve in M. A worldline represents the history of a classical masspoint.

Given a $\mathbf{u} \in V(1)$, we define

$$\mathbf{E}_{\mathbf{u}} := \{ \mathbf{x} \in \mathbf{M} \,|\, \mathbf{u} \cdot \mathbf{x} = 0 \}, \tag{9}$$

which is a three-dimensional linear subspace of **M**. The restriction of the Lorentz product to $\mathbf{E}_{\mathbf{u}}$ is a Euclidean product. The corresponding norm (length of vectors) is denoted by $|\ |$.

Every spacetime vector can be uniquely split into the sum of a timelike vector, parallel to **u**, and a spacelike vector in $\mathbf{E}_{\mathbf{u}}$; in other words, we can give the **u**-splitting

$$\mathbf{M} \to \mathbf{I} \times \mathbf{E}_{\mathbf{u}}, \quad \mathbf{x} \mapsto (-\mathbf{u} \cdot \mathbf{x}, \boldsymbol{\pi}_{\mathbf{u}}(\mathbf{x})), \tag{10}$$

where

$$\boldsymbol{\pi}_{\mathbf{u}}(\mathbf{x}) := \mathbf{x} + (\mathbf{u} \cdot \mathbf{x})\mathbf{u}. \tag{11}$$

An *observer* is a velocity field, i.e. a smooth mapping

$$\mathbf{U}{:}M \to V(1). \tag{12}$$

The set of maximal integral curves of $\mathbf{U}$ (which are worldlines) is the *space* of the observer, or $\mathbf{U}$-space in brief; thus the space points of an observer are curves in spacetime. *The spaces of different observers are evidently different.*

An observer is *inertial* if it is a constant mapping. We shall consider only inertial observers (so the term inertial will occasionally be omitted). An inertial observer can be given by its constant value, that is why we find it convenient to say 'an observer $\mathbf{u}\in V(1)$'.

The space points of the inertial observer $\mathbf{u}$ are straight lines in spacetime, parallel to $\mathbf{u}$.

According to Einstein's standard synchronisation corresponding to the observer $\mathbf{u}$, the world points x and y are $\mathbf{u}$-simultaneous if and only if $\mathbf{u}\cdot(x-y)=0$, in other words, $x-y\in\mathbf{E}_{\mathbf{u}}$.

$\mathbf{u}$-simultaneous world points form an affine hyperplane over $\mathbf{E}_{\mathbf{u}}$. Such a hyperplane is considered a $\mathbf{u}$-instant, and the set $I_{\mathbf{u}}$ of hyperplanes parallel to $\mathbf{E}_{\mathbf{u}}$ is the time of the observer, or $\mathbf{u}$-time in brief. The time interval between the $\mathbf{u}$-instants t_1 and t_2 is defined to be the proper time passed along the $\mathbf{u}$-space points (straight lines parallel to $\mathbf{u}$) between the hyperplanes t_1 and t_2; thus $I_{\mathbf{u}}$, endowed with the subtraction

$$t_1 - t_2 := -\mathbf{u}\cdot(x_1 - x_2) \quad (x_1\in t_1, x_2\in t_2), \tag{13}$$

is an affine space over $\mathbf{I}$.

The vector between two $\mathbf{u}$-space points q_1 and q_2 is defined to be the vector between $\mathbf{u}$-simultaneous world points of the straight lines in question. Thus the space of the observer $\mathbf{u}$, denoted by $E_{\mathbf{u}}$, endowed with the subtraction

$$q_1 - q_2 := x_1 - x_2 \quad (x_1\in q_1,\ x_2\in q_2,\ -\mathbf{u}\cdot(x_1-x_2)=0), \tag{14}$$

is an affine space over the vector space $\mathbf{E}_{\mathbf{u}}$; note that we have

$$q_1 - q_2 = \boldsymbol{\pi}_{\mathbf{u}}(x_1 - x_2) \quad (x_1\in q_1, x_2\in q_2). \tag{15}$$

Thus in the special relativistic spacetime model the *different spaces* of different inertial observers are different affine spaces over *different* vector spaces. Consequently, it has no '*a priori* meaning' that a vector in the space of an observer equals a vector in the space of another observer.

3. Relative Velocities in Special Relativity

The history of a masspoint is observed by an observer $\mathbf{u}$ as a motion and is described as a function assigning $\mathbf{u}$-space points to $\mathbf{u}$-instants. The *motion* relative to the observer $\mathbf{u}$, corresponding to the masspoint history represented by the worldline C, is the mapping $r_{C,\mathbf{u}}{:}I_{\mathbf{u}} \to E_{\mathbf{u}}$ which assigns to t the $\mathbf{u}$-space point (straight line parallel to $\mathbf{u}$) containing the unique intersection point $C\star t$ of the line C and the hyperplane t, i.e.

$$r_{C,\mathbf{u}}(t) = C \bigstar t + \mathbf{u} \otimes \mathbf{I}. \tag{16}$$

The derivative of $r_{C,\mathbf{u}}$ yields the relative velocity of the masspoint with respect to the observer. If C is a straight line parallel to $\mathbf{u}' \in V(1)$, then there is a $\mathbf{t}' \in \mathbf{I}$ such that $C \bigstar t_1 - C \bigstar t_2 = \mathbf{u}'\mathbf{t}'$, thus we infer from (13) and (15) that

$$\frac{r_{C,\mathbf{u}}(t_1) - r_{C,\mathbf{u}}(t_2)}{t_1 - t_2} = \frac{\pi_{\mathbf{u}}(C \bigstar t_1 - C \bigstar t_2)}{-\mathbf{u}\cdot(C \bigstar t_1 - C \bigstar t_2)} = \frac{\mathbf{u}' + (\mathbf{u}\cdot\mathbf{u}')\mathbf{u}}{-\mathbf{u}\cdot\mathbf{u}'}. \tag{17}$$

That is why we define

$$\mathbf{v}_{\mathbf{u}'\mathbf{u}} := \frac{\mathbf{u}'}{-\mathbf{u}\cdot\mathbf{u}'} - \mathbf{u} \tag{18}$$

as the *relative velocity* of $\mathbf{u}'$ with respect to $\mathbf{u}$.[3]

We easily derive that

(i) $\mathbf{v}_{\mathbf{u}'\mathbf{u}} \in \mathbf{E}_{\mathbf{u}}/\mathbf{I}$,

(ii) $|\mathbf{v}_{\mathbf{u}'\mathbf{u}}|^2 = |\mathbf{v}_{\mathbf{u}\mathbf{u}'}|^2 = 1 - 1/(\mathbf{u}'\cdot\mathbf{u})^2$, implying

$$-\mathbf{u}'\cdot\mathbf{u} = \frac{1}{\sqrt{1 - |\mathbf{v}_{\mathbf{u}'\mathbf{u}}|^2}}, \tag{19}$$

(iii) if $\mathbf{u} \neq \mathbf{u}'$, then $\mathbf{E}_{\mathbf{u}} \cap \mathbf{E}_{\mathbf{u}'}$ is a two-dimensional linear subspace of $\mathbf{M}$, and both $\mathbf{v}_{\mathbf{u}'\mathbf{u}}$ and $\mathbf{v}_{\mathbf{u}\mathbf{u}'}$ are orthogonal to $\mathbf{E}_{\mathbf{u}} \cap \mathbf{E}_{\mathbf{u}'}$,

(iv) $\mathbf{v}_{\mathbf{u}'\mathbf{u}} = -\mathbf{v}_{\mathbf{u}\mathbf{u}'}$ if and only if $\mathbf{u} = \mathbf{u}'$ which is equivalent to $\mathbf{v}_{\mathbf{u}'\mathbf{u}} = \mathbf{v}_{\mathbf{u}\mathbf{u}'} = 0$.

Thus we have the following important and far-reaching fact: if $\mathbf{u} \neq \mathbf{u}'$, then *the relative velocity of* $\mathbf{u}$ *with respect to* $\mathbf{u}'$ *is not the opposite of the relative velocity of* $\mathbf{u}'$ *with respect to* $\mathbf{u}$.

A *light signal* is a straight line parallel to a lightlike vector. The motion of a light signal according to an observer is defined similarly to the motion of a masspoint, and we get in an analogous way that a light signal parallel to the lightlike vector $\mathbf{k} \in \frac{\mathbf{M}}{\mathbf{I}}$ has the relative velocity

$$\mathbf{v}_{\mathbf{k}\mathbf{u}} := \frac{\mathbf{k}}{-\mathbf{k}\cdot\mathbf{u}} - \mathbf{u} \tag{20}$$

with respect to the observer $\mathbf{u} \in V(1)$. It is a simple fact that $|\mathbf{v}_{\mathbf{k}\mathbf{u}}| = 1$.

4. Physical Equality (Parallelism) of Vectors in Different Observer Spaces

The space vectors of different observers $\mathbf{u}$ and $\mathbf{u}'$ constitute different three-dimensional vector spaces $\mathbf{E}_{\mathbf{u}}$ and $\mathbf{E}_{\mathbf{u}'}$, respectively. Thus it has no '*a priori*'

[3] We underline that this concerns the standard synchronisation; using another synchronisation, we get another relative velocity.

meaning, in general, that a vector (straight line) in the space of an observer is parallel to a vector (straight line) in the space of another observer; incidentally, this is obvious from a physical point of view, too. Consequently, 'taking two reference frames moving to each other and having parallel axes', a starting point in the usual treatments based on coordinates, is dubious. Moreover, in these usual treatments it is taken for granted that the relative velocities of observers with respect to each other are opposite; we have seen, however, that the relative velocity of $\mathbf{u}$ with respect to $\mathbf{u}'$ is not the opposite of the relative velocity of $\mathbf{u}'$ with respect to $\mathbf{u}$.

Now we establish a physical procedure which establishes a correspondence, called physical equality, between the space vectors of different observers in such a way that relative velocities of observers become physically opposite to each other.

First of all we show that a light signal moving in the $\mathbf{u}$-space in the direction of $\mathbf{v}_{\mathbf{uu}'}$ moves in the $\mathbf{u}'$-space in the direction of $-\mathbf{v}_{\mathbf{u}'\mathbf{u}}$. Let us introduce the unit vectors in the directions of the relative velocities:

$$\mathbf{n}_{\mathbf{u}'\mathbf{u}} := \frac{\mathbf{v}_{\mathbf{u}'\mathbf{u}}}{v}, \quad \mathbf{n}_{\mathbf{uu}'} := \frac{\mathbf{v}_{\mathbf{uu}'}}{v}, \tag{21}$$

where $v := |\mathbf{v}_{\mathbf{u}'\mathbf{u}}| = |\mathbf{v}_{\mathbf{uu}'}|$.

Proposition 1. *If* $\mathbf{k} \in \mathbf{M}/\mathbf{I}$ *is a lightlike vector and*

$$\mathbf{v}_{\mathbf{ku}} = \mathbf{n}_{\mathbf{u}'\mathbf{u}} \tag{22}$$

then

$$\mathbf{v}_{\mathbf{ku}'} = -\mathbf{n}_{\mathbf{uu}'}. \tag{23}$$

Proof. Multiplying equality (22) by $\mathbf{k}$, we easily deduce that

$$\frac{-\mathbf{k}\cdot\mathbf{u}}{-\mathbf{k}\cdot\mathbf{u}'} = \sqrt{\frac{1-v}{1+v}} \tag{24}$$

which yields equality (23). □

Next we suggest that two observers relate their spaces to each other by the following procedure. Take a vector $\mathbf{a}$ in the $\mathbf{u}$-space, orthogonal to $\mathbf{v}_{\mathbf{u}'\mathbf{u}}$. Send light signals of different colours from the starting point and the end point of the vector, respectively, in the direction of $\mathbf{v}_{\mathbf{u}'\mathbf{u}}$. The light signals arrive at the observer $\mathbf{u}'$ in the direction $-\mathbf{v}_{\mathbf{uu}'}$, and hit a plane orthogonal to $\mathbf{v}_{\mathbf{uu}'}$; these two points determine a vector $\mathbf{a}'$ which is considered physically equal to $\mathbf{a}$.

Take a vector $\mathbf{b}$ in the $\mathbf{u}$-space, parallel to $\mathbf{v}_{\mathbf{u}'\mathbf{u}}$. Send light signals $\mathbf{u}$-simultaneously from the starting point and the end point of the vector, in the direction of $\mathbf{v}_{\mathbf{u}'\mathbf{u}}$. The light signals arrive at the observer $\mathbf{u}'$ in the direction $-\mathbf{v}_{\mathbf{uu}'}$, and hit a plane orthogonal to $\mathbf{v}_{\mathbf{uu}'}$ with a time delay $\mathbf{t}'$; the vector $\mathbf{b}' := -\mathbf{t}'\sqrt{1+v/1-v}\,\mathbf{n}_{\mathbf{uu}'}$ in the $\mathbf{u}'$-space is considered physically equal to $\mathbf{b}$.[4]

To formalise the above procedure, we introduce the following definition:

Definition 1. A vector $\mathbf{x}'$ in $\mathbf{E}_{\mathbf{u}'}$ is considered to be *physically equal* to a vector $\mathbf{x}$ in $\mathbf{E}_{\mathbf{u}}$ if

– the orthogonal projection of $\mathbf{x}'$ onto $\mathbf{E}_{\mathbf{u}} \cap \mathbf{E}_{\mathbf{u}'}$ (which is the plane in $\mathbf{E}_{\mathbf{u}'}$ orthogonal to $\mathbf{v}_{\mathbf{uu}'}$) equals the orthogonal projection of $\mathbf{x}$ onto $\mathbf{E}_{\mathbf{u}} \cap \mathbf{E}_{\mathbf{u}'}$ (which is the plane in $\mathbf{E}_{\mathbf{u}}$ orthogonal to $\mathbf{v}_{\mathbf{u}'\mathbf{u}}$), i.e.

$$\mathbf{x}' - (\mathbf{n}_{\mathbf{uu}'} \cdot \mathbf{x}')\mathbf{n}_{\mathbf{uu}'} = \mathbf{x} - (\mathbf{n}_{\mathbf{u}'\mathbf{u}} \cdot \mathbf{x})\mathbf{n}_{\mathbf{u}'\mathbf{u}}, \tag{25}$$

– the orthogonal projection of $\mathbf{x}'$ onto the direction of $\mathbf{v}_{\mathbf{uu}'}$ is opposite to the orthogonal projection of $\mathbf{x}$ onto the direction of $\mathbf{v}_{\mathbf{u}'\mathbf{u}}$, i.e.

$$\mathbf{n}_{\mathbf{uu}'} \cdot \mathbf{x}' = -\mathbf{n}_{\mathbf{u}'\mathbf{u}} \cdot \mathbf{x}. \tag{26}$$

A vector $\mathbf{x}'$ in $\mathbf{E}_{\mathbf{u}'}$ is *physically parallel* to a vector $\mathbf{x}$ in $\mathbf{E}_{\mathbf{u}}$ if there is a real number λ such that $\lambda\mathbf{x}$ is physically equal to $\mathbf{x}'$ or $\mathbf{x}$ is physically equal to $\lambda\mathbf{x}'$.

It is quite evident that physical equality (parallelism) is a symmetric relation but *is not transitive*, as can be seen from the following transparent mathematical formulae.

5. Lorentz Boosts

The agreement about physical equality establishes a linear bijection $\mathbf{E}_{\mathbf{u}} \to \mathbf{E}_{\mathbf{u}'}$, $\mathbf{x} \mapsto \mathbf{x}'$, $\mathbf{x}'$ is physical equal to $\mathbf{x}$, which can be extended to a linear bijection $\mathbf{M} \to \mathbf{M}$ by the requirement $\mathbf{u} \mapsto \mathbf{u}'$. This linear bijection is uniquely determined by the prescribed properties because they fix its values on vectors spanning $\mathbf{M}$. The explicit form of this linear bijection is given as follows.

Let $\mathbf{u}' \otimes \mathbf{u}$ denote the linear map $\mathbf{M} \to \mathbf{M}$, $\mathbf{x} \mapsto \mathbf{u}'(\mathbf{u} \cdot \mathbf{x})$ and let $\mathbf{1}$ be the identity map of $\mathbf{M}$. Then the *Lorentz boost* from $\mathbf{u}'$ to $\mathbf{u}$,

$$\mathbf{B}(\mathbf{u},\mathbf{u}') := \mathbf{1} + \frac{(\mathbf{u}+\mathbf{u}')\otimes(\mathbf{u}+\mathbf{u}')}{1-\mathbf{u}\cdot\mathbf{u}'} - 2\mathbf{u}\otimes\mathbf{u}', \tag{27}$$

establishes the physical equality of vectors in the $\mathbf{u}$-space to vectors in the $\mathbf{u}'$-space, because we have the following easily verifiable relations (see Matolcsi, 1993):

(i) $\mathbf{B}(\mathbf{u}',\mathbf{u})\mathbf{u} = \mathbf{u}'$,
(ii) $\mathbf{B}(\mathbf{u}',\mathbf{u})\mathbf{x} = \mathbf{x}$ if $\mathbf{x} \in \mathbf{E}_{\mathbf{u}} \cap \mathbf{E}_{\mathbf{u}'}$,
(iii) $\mathbf{B}(\mathbf{u}',\mathbf{u})\mathbf{v}_{\mathbf{u}'\mathbf{u}} = -\mathbf{v}_{\mathbf{uu}'}$,

[4] If $\mathbf{b} = \beta\mathbf{n}_{\mathbf{u}'\mathbf{u}}$, then the time delay is determined by the fact that $\beta\mathbf{n}_{\mathbf{u}'\mathbf{u}} + \mathbf{u}'\mathbf{t}'$ is lightlike, from which we infer that $\beta = \mathbf{t}'\sqrt{\frac{1+v}{1-v}}$.

which imply that $\mathbf{B}(\mathbf{u}',\mathbf{u})$ preserves the Lorentz product, its arrow orientation and the orientation of $\mathbf{M}$; furthermore,

(iv) $\mathbf{B}(\mathbf{u},\mathbf{u}')^{-1} = \mathbf{B}(\mathbf{u}',\mathbf{u})$.

Accordingly, the phrase '$\mathbf{x}$ boosted from $\mathbf{E}_{\mathbf{u}}$ into $\mathbf{E}_{\mathbf{u}'}$' will mean $\mathbf{B}(\mathbf{u}',\mathbf{u})\mathbf{x}$, which is the vector in $\mathbf{E}_{\mathbf{u}'}$ physically equal to $\mathbf{x}$.

In the usual treatments of special relativity, spacetime is considered to be $\mathbb{R}\times\mathbb{R}^3$ in which $\mathbb{R}$ is 'time' and $\mathbb{R}^3$ is 'space'. All space vectors—space vectors of different observers—are taken to be elements of the same vector space $\mathbb{R}^3$. This corresponds to the fact that an observer $\mathbf{u}$ and an orthonormal basis in the $\mathbf{u}$-space are chosen to coordinatise spacetime, i.e. an observer is 'hidden' in the coordinates and all space vectors are tacitly boosted into the space of the hidden observer.

The Lorentz boost above is the absolute counterpart of the usual 'pure Lorentz transformation' or 'Lorentz transformation without rotation': if $\mathbf{n}_1$, $\mathbf{n}_2$, $\mathbf{n}_3$ is an orthonormal basis (representing coordinate axes) in the $\mathbf{u}$-space, then $\mathbf{B}(\mathbf{u}',\mathbf{u})\mathbf{n}_1$, $\mathbf{B}(\mathbf{u}',\mathbf{u})\mathbf{n}_2$, $\mathbf{B}(\mathbf{u}',\mathbf{u})\mathbf{n}_3$ determine the coordinate axes in the $\mathbf{u}'$-space that are parallel to those in the $\mathbf{u}$-space. Moreover, the matrix of $\mathbf{B}(\mathbf{u}',\mathbf{u})$ in the basis $\mathbf{n}_0 := \mathbf{u}$, $\mathbf{n}_1$, $\mathbf{n}_2$, $\mathbf{n}_3$ becomes the well-known usual Lorentz matrix

$$\begin{pmatrix} \kappa & \kappa v_1 & \kappa v_2 & \kappa v_3 \\ \kappa v_1 & 1+\frac{\kappa^2}{1+\kappa}v_1^2 & \frac{\kappa^2}{1+\kappa}v_1 v_2 & \frac{\kappa^2}{1+\kappa}v_1 v_3 \\ \kappa v_2 & \frac{\kappa^2}{1+\kappa}v_2 v_1 & 1+\frac{\kappa^2}{1+\kappa}v_2^2 & \frac{\kappa^2}{1+\kappa}v_2 v_3 \\ \kappa v_3 & \frac{\kappa^2}{1+\kappa}v_3 v_1 & \frac{\kappa^2}{1+\kappa}v_3 v_2 & 1+\frac{\kappa^2}{1+\kappa}v_3^2 \end{pmatrix} \tag{28}$$

where $\kappa := 1/\sqrt{1-|\mathbf{v}_{\mathbf{u}'\mathbf{u}}|^2} = -\mathbf{u}\cdot\mathbf{u}'$ and $v_i := \mathbf{n}_i\cdot\mathbf{v}_{\mathbf{u}'\mathbf{u}} = \mathbf{n}_i\cdot\mathbf{u}'/\kappa$ ($i = 1,2,3$).

Note that the Lorentz boost in the present context refers to two observers, i.e. to *two absolute velocities*. The usual matrix of a Lorentz transformation refers to a *single relative velocity*. Nevertheless, that matrix form, too, refers to two observers, but one of them is 'hidden' in the coordinate axes and the relative velocity of another observer is taken with respect to the hidden observer.

Our treatment rules out hidden observers and coordinates.

The simple form (27) of the Lorentz boost allows us to exhibit a simple form for two successive Lorentz boosts: let $\mathbf{u},\mathbf{u}',\mathbf{u}''\in V(1)$. Then

$$\mathbf{B}(\mathbf{u}'',\mathbf{u}')\mathbf{B}(\mathbf{u}',\mathbf{u}) = \mathbf{1} + \frac{(\mathbf{u}''+\mathbf{u}')\otimes(\mathbf{u}''+\mathbf{u}')}{1-\mathbf{u}''\cdot\mathbf{u}'} + \frac{(\mathbf{u}'+\mathbf{u})\otimes(\mathbf{u}'+\mathbf{u})}{1-\mathbf{u}'\cdot\mathbf{u}} - \frac{(\mathbf{u}''+\mathbf{u}')\otimes(\mathbf{u}'+\mathbf{u})(1-\mathbf{u}''\cdot\mathbf{u}'-\mathbf{u}'\cdot\mathbf{u}'-\mathbf{u}\cdot\mathbf{u}'')}{(1-\mathbf{u}''\cdot\mathbf{u}')(1-\mathbf{u}'\cdot\mathbf{u})}, \tag{29}$$

from which we easily derive that the Lorentz boost from $\mathbf{u}$ to $\mathbf{u}'$ followed by the Lorentz boost from $\mathbf{u}'$ to $\mathbf{u}''$ is, in general, not the Lorentz boost from $\mathbf{u}$ to $\mathbf{u}''$. More precisely, we have

$$\mathbf{B}(\mathbf{u}'',\mathbf{u}')\mathbf{B}(\mathbf{u}',\mathbf{u}) = \mathbf{B}(\mathbf{u}'',\mathbf{u}) \quad \text{iff } \mathbf{u},\ \mathbf{u}' \text{ and } \mathbf{u}'' \text{are coplanar.} \tag{30}$$

This shows that the *physical equality* (and physical parallelism) of vectors in different observer spaces *is not a transitive relation*: it may be that

(i) $\mathbf{x}' \in \mathbf{E}_{\mathbf{u}'}$ is physically equal to $\mathbf{x} \in \mathbf{E}_{\mathbf{u}}$ i.e. $\mathbf{x}' = \mathbf{B}(\mathbf{u}',\mathbf{u})\mathbf{x}$, and
(ii) $\mathbf{x}'' \in \mathbf{E}_{\mathbf{u}''}$ is physically equal to $\mathbf{x}' \in \mathbf{E}_{\mathbf{u}'}$ i.e. $\mathbf{x}'' = \mathbf{B}(\mathbf{u}'',\mathbf{u}')\mathbf{x}'$, but
(iii) $\mathbf{x}''$ is not physically equal to $\mathbf{x}$: $\mathbf{x}'' \neq \mathbf{B}(\mathbf{u}'',\mathbf{u})\mathbf{x}$.

Let us remark that in the usual treatments of special relativity, using relative velocities and coordinates, one states that the product of two pure Lorentz transformations is a pure Lorentz transformation if and only if the corresponding relative velocities are colinear. There, all the space vectors—relative velocities as well—are tacitly boosted in the space of the observer hidden in the coordinates. The two relative velocities in question would be $\mathbf{v}_{\mathbf{u}'\mathbf{u}}$ and $\mathbf{v}_{\mathbf{u}''\mathbf{u}'}$; the latter one, however, is to be boosted into the $\mathbf{u}$-space. Thus

$$\mathbf{v} := \mathbf{v}_{\mathbf{u}'\mathbf{u}}, \qquad \mathbf{w} := \mathbf{B}(\mathbf{u},\mathbf{u}')\mathbf{v}_{\mathbf{u}''\mathbf{u}'} \tag{31}$$

correspond to the relative velocities in the usual treatments. Then the statement mentioned above derives from the following:

Proposition 2. $\mathbf{u},\mathbf{u}',\mathbf{u}''$ *are coplanar if and only if* $\mathbf{v}_{\mathbf{u}'\mathbf{u}}$ *and* $\mathbf{B}(\mathbf{u},\mathbf{u}')\mathbf{v}_{\mathbf{u}''\mathbf{u}'}$ *are colinear.*

6. Thomas Rotation

We can reformulate (30) as follows:

$$\mathbf{R}_{\mathbf{u}}(\mathbf{u}',\mathbf{u}'') := \mathbf{B}(\mathbf{u},\mathbf{u}'')\mathbf{B}(\mathbf{u}'',\mathbf{u}')\mathbf{B}(\mathbf{u}',\mathbf{u}) \tag{32}$$

is the identity transformation if and only if $\mathbf{u}$, $\mathbf{u}'$ and $\mathbf{u}''$ are coplanar.

We easily find from the properties of the Lorentz boosts:

Proposition 3.

(i) $\mathbf{R}_{\mathbf{u}}(\mathbf{u}',\mathbf{u}'')\mathbf{u} = \mathbf{u}$,
(ii) *the restriction of* $\mathbf{R}_{\mathbf{u}}(\mathbf{u}',\mathbf{u}'')$ *onto* $\mathbf{E}_{\mathbf{u}}$ *is an orientation and Euclidean product preserving linear bijection from* $\mathbf{E}_{\mathbf{u}}$ *onto* $\mathbf{E}_{\mathbf{u}}$, *i.e. it is a rotation*,
(iii) *excluding the trivial case when the three absolute velocities are coplanar, the axis of rotation* (*the set of invariant vectors*) *is the one-dimensional linear subspace* $\mathbf{E}_{\mathbf{u}} \cap \mathbf{E}_{\mathbf{u}'} \cap \mathbf{E}_{\mathbf{u}''}$.

We continue to consider the linear bijection defined on $\mathbf{M}$ rather than its restriction to $\mathbf{E}_{\mathbf{u}}$. That is why we accept the following definition:

Definition 2. $\mathbf{R}_{\mathbf{u}}(\mathbf{u}',\mathbf{u}'')$ is called the *Thomas rotation* of $\mathbf{u}$ corresponding to $\mathbf{u}'$ and $\mathbf{u}''$.

We emphasise that the Thomas rotation is defined without coordinate axes, so its fundamental meaning is not connected with the rotation of axes and it corresponds to no real rotation. *The Thomas rotation measures the deviation of the physical equality from being transitive.*

Note that *three absolute velocities* are involved in the above definition. In the usual formulation based on coordinates, the matrix of the Thomas rotation (see Ungar, 1989) involves two relative velocities because an observer is hidden in the coordinate axes.

Of course, we should like to deduce from the previous definition the expression for the Thomas rotation in terms of relative velocities. Since $\mathbf{R}_{\mathbf{u}}(\mathbf{u}',\mathbf{u}'')$ is a rotation in the $\mathbf{u}$-space, and space in the usual matrix formalism always means the space of the hidden observer, $\mathbf{u}$ now corresponds to the hidden observer and the two relative velocities in question would be $\mathbf{v}_{\mathbf{u}'\mathbf{u}}$ and $\mathbf{v}_{\mathbf{u}''\mathbf{u}'}$; however, the latter one is to be boosted in the $\mathbf{u}$-space, i.e. we have to take the velocities defined in (31). It is easy to check that $\mathbf{v}$ and $\mathbf{w}$ are in the rotation plane of the Thomas rotation, i.e. they are orthogonal to the rotation axis $\mathbf{E}_{\mathbf{u}} \cap \mathbf{E}_{\mathbf{u}'} \cap \mathbf{E}_{\mathbf{u}''}$. Introducing

$$\alpha := -\mathbf{u}' \cdot \mathbf{u} = \frac{1}{\sqrt{1 - |\mathbf{v}|^2}}, \tag{33}$$

$$\beta := -\mathbf{u}'' \cdot \mathbf{u}' = \frac{1}{\sqrt{1 - |\mathbf{w}|^2}}, \tag{34}$$

$$\gamma := -\mathbf{u}'' \cdot \mathbf{u} = \alpha\beta(1 + \mathbf{v} \cdot \mathbf{w}), \tag{35}$$

we can recover the absolute velocities $\mathbf{u}'$ and $\mathbf{u}''$ from $\mathbf{u}$ and the relative velocites $\mathbf{v}$ and $\mathbf{w}$:

$$\mathbf{u}' = \alpha(\mathbf{u} + \mathbf{v}), \tag{36}$$

$$\mathbf{u}'' = \beta(\mathbf{u}' + \mathbf{B}(\mathbf{u}',\mathbf{u})\mathbf{w}) = \gamma\mathbf{u} + \beta\mathbf{w} + \frac{\alpha(\beta + \gamma)}{1 + \alpha}\mathbf{v}. \tag{37}$$

The Thomas rotation in terms of relative velocities is

$$\mathbf{T}_{\mathbf{u}}(\mathbf{v},\mathbf{w}) := \mathbf{R}_{\mathbf{u}}\left(\alpha(\mathbf{u} + \mathbf{v}), \gamma\mathbf{u} + \beta\mathbf{w} + \frac{\alpha(\beta + \gamma)}{1 + \alpha}\mathbf{v}\right). \tag{38}$$

A lengthy but straightforward calculation yields the following result:

Proposition 4.

$$\mathbf{T}_{\mathbf{u}}(\mathbf{v},\mathbf{w}) = \mathbf{1} + \alpha^2\frac{1 - \beta}{(1 + \alpha)(1 + \gamma)}\mathbf{v}\otimes\mathbf{v} + \beta^2\frac{1 - \alpha}{(1 + \beta)(1 + \gamma)}\mathbf{w}\otimes\mathbf{w}$$

$$\alpha\beta\frac{(1 + \alpha)(1 + \gamma) + (\beta + \gamma)(1 - \alpha)}{(1 + \alpha)(1 + \beta)(1 + \gamma)}\mathbf{v}\otimes\mathbf{w} - \alpha\beta\frac{1}{1 + \gamma}\mathbf{w}\otimes\mathbf{v}. \tag{39}$$

This very nice form allows us to deduce easily all the results regarding the Thomas rotation which are difficult to obtain in the matrix formalism.

Proposition 5. *Let* ε *denote the angle of rotation of* $\mathbf{T_u}(\mathbf{v},\mathbf{w})$ *and let* θ *denote the angle between* $\mathbf{v}$ *and* $\mathbf{w}$. *Then*

$$\cos\varepsilon = 1 - \frac{(\alpha-1)(\beta-1)}{1+\gamma}\sin^2\theta. \tag{40}$$

Proof. It suffices to find the cosine of the angle between $\mathbf{x}$ and $\mathbf{T_u}(\mathbf{v},\mathbf{w})\mathbf{x}$ for some special $\mathbf{x}$ in the rotation plane. Let us choose $\mathbf{x}$ such that $|\mathbf{x}| = 1$, $\mathbf{w}\cdot\mathbf{x} = 0$ and $\mathbf{v}\cdot\mathbf{x} > 0$. Then $\mathbf{v}\cdot\mathbf{x} = |\mathbf{v}|\sin\theta$ and $\cos\varepsilon = \mathbf{x}\cdot\mathbf{T_u}(\mathbf{v},\mathbf{w})\mathbf{x}$, so we get the desired result immediately. □

The above formula is simple but it contains four quantities which are not independent:

$$\gamma = \alpha\beta + \sqrt{\alpha^2-1}\sqrt{\beta^2-1}\cos\theta. \tag{41}$$

Eliminating θ, we get

$$\cos\varepsilon = 1 - \frac{1+2\alpha\beta\gamma-(\alpha^2+\beta^2+\gamma^2)}{(1+\alpha)(1+\beta)(1+\gamma)}. \tag{42}$$

Eliminating γ and introducing

$$k := \sqrt{\frac{(\alpha+1)(\beta+1)}{(\alpha-1)(\beta-1)}}, \tag{43}$$

we get

$$\cos\varepsilon = 1 - \frac{2\sin^2\theta}{1+k^2+2k\cos\theta} = \frac{(k+\cos\theta)^2-\sin^2\theta}{(k+\cos\theta)^2+\sin^2\theta}. \tag{44}$$

The orientation (the positive direction of the rotation axis) of the Thomas rotation is the direction of $\mathbf{x}\times\mathbf{T_u}(\mathbf{v},\mathbf{w})\mathbf{x}$ where $\mathbf{x}$ is an arbitrary non-zero vector in the rotation plane and $\times$ denotes the vectorial product.

Proposition 6. *The orientation of the Thomas rotation* $\mathbf{T_u}(\mathbf{v},\mathbf{w})$ *is given by* $\mathbf{w}\times\mathbf{v}$.

Proof. The map $\lambda\mapsto\mathbf{T_u}(\mathbf{v},\mathbf{w}+\lambda\mathbf{v})$ $(\lambda\in\mathbb{R})$ is continuous. Since there are two disjoint orientations, a continuous map cannot change the orientation; consequently, the orientation is the same for all λ as for $\lambda = 0$. Let λ_0 be such that $(\mathbf{w}+\lambda_0\mathbf{v})\cdot\mathbf{v} = 0$. Then we easily find that

$$\mathbf{v}\times\mathbf{T_u}(\mathbf{v},\mathbf{w}+\lambda_0\mathbf{v})\mathbf{v} = -\frac{\alpha\beta|\mathbf{v}|^2}{1+\gamma}(\mathbf{v}\times\mathbf{w}), \tag{45}$$

which proves our assertion, because all the coefficients on the right-hand side are positive. □

7. The Velocity Addition Paradox

The velocity addition paradox can be described as follows. Let us consider three observers: me, you and him for the sake of easy formulation. Your velocity $\mathbf{v}$ relative to me and his velocity $\mathbf{w}$ relative to you determine his velocity $\mathbf{v}\oplus\mathbf{w}$ to me by the formula (see Mocanu, 1992)

$$\mathbf{v}\oplus\mathbf{w} = \frac{\alpha\beta}{\gamma}\left(\mathbf{v} + \mathbf{w} + \frac{\alpha}{1+\alpha}\mathbf{v}\times(\mathbf{v}\times\mathbf{w})\right) = \frac{\alpha(\beta+\gamma)}{\gamma(1+\alpha)}\mathbf{v} + \frac{\beta}{\gamma}\mathbf{w}. \tag{46}$$

Similarly, your velocity $\hat{\mathbf{w}}$ relative to him and my velocity $\hat{\mathbf{v}}$ relative to you determine my velocity $\hat{\mathbf{w}}\oplus\hat{\mathbf{v}}$ relative to him by the same formula. We 'evidently' have $\hat{\mathbf{w}} = -\mathbf{w}$, $\hat{\mathbf{v}} = -\mathbf{v}$ and $\hat{\mathbf{w}}\oplus\hat{\mathbf{v}} = -\mathbf{v}\oplus\mathbf{w}$; however, the actual formula for the addition $\oplus$ shows that, in general,

$$(-\mathbf{w})\oplus(-\mathbf{v}) \neq -(\mathbf{v}\oplus\mathbf{w}), \quad \text{or, equivalently,} \quad \mathbf{w}\oplus\mathbf{v} \neq \mathbf{v}\oplus\mathbf{w}. \tag{47}$$

We shall soon see that the paradox arose in the usual matrix formalism from the fact that instead of vectors in the spaces of different observers one tacitly considers the corresponding physically equal vectors in the space of the hidden observer (every space vector is tacitly boosted into the space of the hidden observer), which implies the incorrect tacit assumption that physical equality is a transitive relation. In fact, relative velocites $\mathbf{v}$ and $\mathbf{w}$ as well as $\hat{\mathbf{w}}$ and $\hat{\mathbf{v}}$ are considered to be elements of $\mathbb{R}^3$, their sum and vectorial product appear in the formulae $\mathbf{v}\oplus\mathbf{w}$ and $\hat{\mathbf{w}}\oplus\hat{\mathbf{v}}$, yielding elements of $\mathbb{R}^3$.

Now let us return to the notations $\mathbf{u}$ (me), $\mathbf{u}'$ (you) and $\mathbf{u}''$ (he). Then $\mathbf{v}_{\mathbf{u}''\mathbf{u}}$ would be $\mathbf{v}$ and $\mathbf{v}_{\mathbf{u}''\mathbf{u}'}$ would be $\mathbf{w}$. However, $\mathbf{v}_{\mathbf{u}'\mathbf{u}'}$ and $\mathbf{v}_{\mathbf{u}''\mathbf{u}'}$ are in the different three-dimensional vector spaces $\mathbf{E}_{\mathbf{u}}/\mathbf{I}$ and $\mathbf{E}'_{\mathbf{u}}/\mathbf{I}$ respectively, their linear combination does not lie in either $\mathbf{E}_{\mathbf{u}}/\mathbf{I}$ or $\mathbf{E}'_{\mathbf{u}}/\mathbf{I}$, and their vectorial product is not meaningful.

The velocity addition formula (46) *is meaningful and holds true only if the second relative velocity is boosted into the space of the observer to which the first velocity and the resulting one are related.*

Thus we have to take $\mathbf{v} = \mathbf{v}_{\mathbf{u}''\mathbf{u}'}$ and $\mathbf{w} = \mathbf{B}(\mathbf{u},\mathbf{u}')\mathbf{v}_{\mathbf{u}''\mathbf{u}'}$ in accordance with (31) and then

$$\mathbf{v}\oplus\mathbf{w} = \mathbf{v}_{\mathbf{u}''\mathbf{u}}. \tag{48}$$

Regarding the other addition in the paradox involving $\hat{\mathbf{w}}$ and $\hat{\mathbf{v}}$, we must be careful: since the velocity of $\mathbf{u}$ relative to $\mathbf{u}''$ is calculated by the addition formula from the velocity of $\mathbf{u}'$ relative to $\mathbf{u}''$ and from the velocity of $\mathbf{u}$ relative to $\mathbf{u}'$, this last relative velocity must be boosted into the $\mathbf{u}''$-space, so

$$\hat{\mathbf{w}} := \mathbf{v}_{\mathbf{u}'\mathbf{u}''}, \qquad \hat{\mathbf{v}} := \mathbf{B}(\mathbf{u}'',\mathbf{u}')\mathbf{v}_{\mathbf{u}\mathbf{u}'} \tag{49}$$

and then

$$\hat{\mathbf{w}}\oplus\hat{\mathbf{v}} = \mathbf{v}_{\mathbf{u}\mathbf{u}''}. \tag{50}$$

Note that $\hat{\mathbf{w}}$ and $-\mathbf{w}$ as well as $\hat{\mathbf{v}}$ and $-\mathbf{v}$ are in different spaces: the first ones in the $\mathbf{u}''$-space, the second ones in the $\mathbf{u}$-space. The 'evidence' used in the formulation of the paradox that $\hat{\mathbf{w}}$ equals $-\mathbf{w}$ and $\hat{\mathbf{v}}$ equals $-\mathbf{v}$ would indeed imply that $\hat{\mathbf{w}}$ and $\hat{\mathbf{v}}$ are physically equal to $-\mathbf{w}$ and $-\mathbf{v}$, respectively. But the vectors in the $\mathbf{u}$-space, physically equal to $\hat{\mathbf{w}}$ and $\hat{\mathbf{v}}$, respectively, are

$$\begin{aligned}\mathbf{B}(\mathbf{u},\mathbf{u}'')\hat{\mathbf{w}} &= \mathbf{B}(\mathbf{u},\mathbf{u}'')\mathbf{v}_{\mathbf{u}'\mathbf{u}''} = -\mathbf{B}(\mathbf{u},\mathbf{u}'')\mathbf{B}(\mathbf{u}'',\mathbf{u}')\mathbf{v}_{\mathbf{u}''\mathbf{u}'}\\ &= -\mathbf{B}(\mathbf{u},\mathbf{u}'')\mathbf{B}(\mathbf{u}'',\mathbf{u}')\mathbf{B}(\mathbf{u}',\mathbf{u})\mathbf{w} = -\mathbf{R}_{\mathbf{u}}(\mathbf{u}',\mathbf{u}'')\mathbf{w}, \qquad (51)\end{aligned}$$

$$\begin{aligned}\mathbf{B}(\mathbf{u},\mathbf{u}'')\hat{\mathbf{v}} &= \mathbf{B}(\mathbf{u},\mathbf{u}'')\mathbf{B}(\mathbf{u}'',\mathbf{u}')\mathbf{v}_{\mathbf{u}\mathbf{u}'}\\ &= -\mathbf{B}(\mathbf{u},\mathbf{u}'')\mathbf{B}(\mathbf{u}'',\mathbf{u}')\mathbf{B}(\mathbf{u}',\mathbf{u})\mathbf{v}_{\mathbf{u}''\mathbf{u}'} = -\mathbf{R}_{\mathbf{u}}(\mathbf{u}',\mathbf{u}'')\mathbf{v}. \qquad (52)\end{aligned}$$

We see that $\hat{\mathbf{w}}$ *is not physically equal* to $-\mathbf{w}$ and $\hat{\mathbf{v}}$ *is not physically equal* to $-\mathbf{v}$, contrary to the 'evidence' which leads to the paradox. Then it is not surprising that $\hat{\mathbf{w}}\oplus\hat{\mathbf{v}}$ is not physically equal to $\mathbf{v}\oplus\mathbf{w}$ either. All this is the consequence of the non-transitivity of physical equality.

Formulae (51) and (52) indicate that the velocity addition formula will be 'commutative' if we replace $\hat{\mathbf{w}}$ and $\hat{\mathbf{v}}$ by $-\mathbf{R}_{\mathbf{u}}(\mathbf{u}',\mathbf{u}'')\mathbf{w} = \mathbf{T}_{\mathbf{u}}(\mathbf{v},\mathbf{w})\mathbf{w}$ and $-\mathbf{R}_{\mathbf{u}}(\mathbf{u}',\mathbf{u}'')\mathbf{v} = \mathbf{T}_{\mathbf{u}}(\mathbf{v},\mathbf{w})\mathbf{v}$, respectively.

Proposition 7.

$$\mathbf{T}_{\mathbf{u}}(\mathbf{v},\mathbf{w})\mathbf{w}\oplus\mathbf{T}_{\mathbf{u}}(\mathbf{v},\mathbf{w})\mathbf{v} = \mathbf{T}_{\mathbf{u}}(\mathbf{v},\mathbf{w})(\mathbf{w}\oplus\mathbf{v}) = \mathbf{v}\oplus\mathbf{w}. \qquad (53)$$

Proof. In the next lemma we demonstrate that a Lorentz transformation (in particular, the Thomas rotation) is 'linear' with respect to the addition $\oplus$, which implies the first equality.

To prove the second equality, we apply the inverse of the Thomas rotation and we take into account the properties of Lorentz boosts, as well as equalities (48)–(50):

$$\begin{aligned}\mathbf{T}_{\mathbf{u}}(\mathbf{v},\mathbf{w})^{-1}(\mathbf{v}\oplus\mathbf{w}) &= \mathbf{B}(\mathbf{u},\mathbf{u}')\mathbf{B}(\mathbf{u}',\mathbf{u}'')\mathbf{B}(\mathbf{u}'',\mathbf{u})\mathbf{v}_{\mathbf{u}''\mathbf{u}}\\ &= -\mathbf{B}(\mathbf{u},\mathbf{u}')\mathbf{B}(\mathbf{u}',\mathbf{u}'')\mathbf{v}_{\mathbf{u}\mathbf{u}''}\\ &= -\mathbf{B}(\mathbf{u},\mathbf{u}')\mathbf{B}(\mathbf{u}',\mathbf{u}'')(\mathbf{v}_{\mathbf{u}'\mathbf{u}''}\oplus\mathbf{B}(\mathbf{u}'',\mathbf{u}')\mathbf{v}_{\mathbf{u}\mathbf{u}'})\\ &= (\mathbf{B}(\mathbf{u},\mathbf{u}')\mathbf{v}_{\mathbf{u}''\mathbf{u}'})\oplus\mathbf{v}_{\mathbf{u}'\mathbf{u}} = \mathbf{w}\oplus\mathbf{v}. \qquad (54)\end{aligned}$$

Lemma 1. *If* $\mathbf{L}$ *is a Lorentz transformation—i.e. a Lorentz product preserving linear bijection of* $\mathbf{M}$*—then*

$$\mathbf{L}(\mathbf{v}\oplus\mathbf{w}) = (\mathbf{L}\mathbf{v})\oplus(\mathbf{L}\mathbf{w}). \qquad (55)$$

Proof. According to (46), $\mathbf{v}\oplus\mathbf{w}$ is a linear combination of $\mathbf{v}$ and $\mathbf{w}$ with coefficients composed of $|\mathbf{v}|^2$, $|\mathbf{w}|^2$ and $\mathbf{v}\cdot\mathbf{w}$. Then our assertion is quite trivial, since $\mathbf{L}$ is linear and $|\mathbf{L}\mathbf{v}|^2 = |\mathbf{v}|^2$, $|\mathbf{L}\mathbf{w}|^2 = |\mathbf{w}|^2$, $(\mathbf{L}\mathbf{v})\cdot(\mathbf{L}\mathbf{w}) = \mathbf{v}\cdot\mathbf{w}$. $\square$

8. Discussion

A spacetime structure which is free of observers, reference frames and coordinates admits a treatment of special relativity based on *absolute* objects, i.e. objects not involving reference frames and coordinates. Such a treatment makes it evident that the spaces of different observers are different affine spaces over *different* vector spaces, in contrast to the non-relativistic case. A physical procedure has been given to relate the spaces of different observers to each other, establishing a notion of physical equality and physical parallelism of vectors in different observer spaces. The physical equality is described by Lorentz boosts, whose explicit form in terms of absolute velocities allows us to very easily make calculations; in particular, we easily obtain that the product of two successive Lorentz boosts is, in general, not a Lorentz boost, which shows that *physical equality is not a transitive relation*, in contrast to the non-relativistic cases. The Thomas rotation, simply defined by the succession of three Lorentz boosts, measures how much the relation of physical equality of vectors deviates from being transitive. An explicit form of the Thomas rotation has been deduced which is much more convenient to apply than the usual matrix forms.

The absolute formulation of spacetime illuminates that the velocity addition paradox is a consequence of the facts that in the treatments using coordinates

1. the space of every observer is tacitly considered through the corresponding physically equal vectors in the space of the observer hidden in the coordinates,
2. physical equality is tacitly taken to be a transitive relation.

References

Appleby, P. G. and Kadianakis, N. (1986) ‘A Frame-Independent Description of the Principles of Classical Mechanics’, *Archive for Rational Mechanics and Analysis* **95,** 1–22.

Good, I. J. (1995) ‘Lorentz Matrices: A Review’, *International Journal of Theoretical Physics* **74,** 779–799.

Matolcsi, T. (1984) *A Concept of Mathematical Physics, Models for Spacetime* (Budapest: Akadémiai Kiadó).

Matolcsi, T. (1993) *Spacetime without Reference Frames* (Budapest: Akadémiai Kiadó).

Matolcsi, T. (1998) ‘Spacetime without Reference Frames: An Application to Synchronization on the Rotating Disk’, *Foundations of Physics* **27,** 1685–1701.

Mocanu, C. I. (1992) ‘On the Relativistic Velocity Composition Paradox and Thomas Rotation’, *Foundations of Physics Letters* **5,** 443–456.

Møller, M. C. (1972) *The Theory of Relativity,* 2nd edn (Oxford: Clarendon Press).

Selleri, F. (1997) ‘Noninvariant One-Way Speed of Light and Locally Equivalent Reference Frames’, *Foundations of Physics Letters* **10,** 73–83.

Synge, J. L. (1955) *Relativity: The Special Theory* (Amsterdam: North-Holland).

Synge, J. L. (1964) *Relativity: The General Theory* (Amsterdam: North-Holland).
Rodrigues Jr, W. A., de Souza, Q. A. G. and Bozhkov, Y. (1995) 'The Mathematical Structure of Newtonian Spacetime: Classical Dynamics and Gravitation', *Foundations of Physics* **25,** 871–924.
Ungar, A. A. (1989) 'The Relativistic Velocity Composition Paradox and the Thomas Rotation', *Foundation of Physics* **19,** 1385–1396.
Wald, R. (1983) *Space, Time and Gravity* (Chicago: University of Chicago Press).
Weyl, H. (1922) *Space – Time – Matter* (New York: Dover).