International Journal of Theoretical Physics, Vol. 34, No. 5, 1995

Lorentz Matrices: A Review

I. J. Good[1]

Received December 22, 1994

This is an expository review of the Lorentz transformation, which is a change of coordinates used by one "inertial observer" to those used by another one. The transformation can be represented by a four-by-four matrix, the Lorentz matrix or the Minkowski–Lorentz matrix. The most familiar, or "special," case has the x axis of both observers parallel to their relative velocity. A more general transformation drops this constraint. But then a seeming "paradox" arises when there are three observers, and this has led to a challenge to the self-consistency of the special theory of relativity. It is shown here that this challenge is based on a misunderstanding. The properties of the more general Lorentz transformation are reviewed consistently in terms of the matrix approach, which the author believes is now the easiest approach to understand. The spectral analysis of the Lorentz matrix is also discussed. Several checks are included to "make assurance double sure."

1. INTRODUCTION

The *Lorentz matrix*, as understood herein, is a four-by-four matrix, defined below, that effects a transformation from one coordinate system in spacetime to another one. Actually there are two slightly different forms of the matrix. Matrices are now familiar to "everybody" (physicists, chemists, engineers, mathematicians, statisticians, econometricians, and programmers) so I believe it is well worthwhile, for the sake of clarity, to expound the "general" Lorentz transformation based entirely on the matrix approach, and also briefly to study the Lorentz matrix itself. Matrices have not always been so familiar: even Heisenberg, when he invented matrix mechanics in 1925, did not know that he was using matrices; while Whittaker (1953, pp. 255–257) still thought it appropriate to discuss the most elementary properties of ma-

[1]Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 24061-0439.

0020-7748/95/0500-0779$07.50/0

trices in a book written for physicists. The familiarity of matrices today is largely due to the computer revolution.

One point of the present paper is to show the fallacy in an attack on the *self-consistency* of the special theory of relativity (STR). See Galeczki (1993, p. 448; 1994a, p. 78), where he discusses "the incompatibility between Lorentz transformations and the inertial frame of reference" and again in Galeczki (1994b, p. 85, columns i and ii). I understand that he makes the point also in a book in German on a "requiem" for the STR. An English version is in press. Once such fallacious attacks are published in four or five places it is important to show, as clearly as possible, why these attacks are incorrect. Otherwise many people will be misled on the grounds that what is said four times is true. The fallacy arises largely because of the ambiguity of the word *rotation*, an ambiguity that is not made clear in the *Oxford English Dictionary*, for example. Before getting down to details, let us consider an extremely elementary example of this ambiguity.

For the last 300 years, physicists and other applied mathematicians have been choosing Cartesian coordinate systems that happen to be convenient for specific applications. For example, to specify the position of a point inside a brick-shaped auditorium one might choose the axes parallel to the walls and floor. It would be mathematically consistent, though perhaps inconvenient, to use some specific *nonrotating* rotation of that system. To say that one Cartesian system is a rotation of another one might mislead some listeners into thinking that one of the two systems is *rotating* if we do not specify otherwise. That is why I have used the oxymoron-sounding expression "non-rotating rotation." When one changes one's mind about what frame of reference to use for a specific purpose, it is unnecessary to rotate one's swivel chair. So much for the preliminaries.

2. THE THREE OBSERVERS

An *inertial system* is a four-dimensional spatial and temporal coordinate system in which a particle under no forces moves with a uniform velocity (a vector). An "inertial observer" is one at rest in an inertial system which I call a Natural Coordinate System or NCS for that observer. (It is often known, in less self-explanatory terminology, as a "proper" system, a usage that clashes with the definition of a proper orthogonal matrix.) Any *nonrotating* spatial rotation of an NCS is also an NCS for that observer, so each inertial observer has ∞^3 (not ∞^4) NCSs and can choose one of these arbitrarily or on grounds of convenience. A *rotating* rotation of an inertial system cannot be inertial.

Let Arthur, Bertha, and Chipso (a robot) be three inertial observers. We may call them A, B, and C when we do not want to be personal. Let the coordinates attributed to a generic *event* (point event) by these observers in their chosen NCSs be, respectively, (x_A, y_A, z_A, t_A) or (ξ_A^T, t_A), with a similar notation for B and C. (The superscript T denotes transposition and ξ_A^T denotes a row vector with three components.) For the sake of simplicity, but with a slight loss of generality, we assume that the "origin" (0, 0, 0, 0) is the same "event" for all three observers; for example, they might meet instantaneously at A's spatial origin and synchronize their clocks to zero thereat. The coordinates used by a given observer will depend on the NCS chosen by that observer. Any other NCS for that observer, with the same origin, is a spatial nonrotating rotation of the one chosen.

Our vectors will always be in three or four dimensions. Geometrically they have magnitude and direction and we sometimes find it convenient to think of them as being rooted at the origin.

As measured by Arthur in his chosen NCS, Bertha is moving with velocity $\mathbf{v}$ (denoted boldly). This velocity is necessarily uniform because she, too, is "inertial." The corresponding speed is denoted timidly by v. Suppose further that the NCS chosen by Bertha is *quasiparallel* to Arthur's chosen NCS, as measured in his NCS. For the moment I ask the reader to accept "quasiparallel" in a vague sense meaning something like "parallel." It will turn out later that Bertha, too, will then necessarily "regard" A's NCS as quasiparallel to hers. (Nearly all popular books and articles on physics, and many unpopular ones, use technical terms without definition, but I shall clarify my meaning in Section 4.) This to some extent justifies us in saying that the NCSs of A and B *are* quasiparallel, *tout court*. But the justification is incomplete because a third observer might "regard" these coordinate systems as not quasiparallel. We shall soon return to this point.

Let Chipso have a (uniform) velocity $\mathbf{u}$ with respect to Bertha, as "seen" by her, and suppose that its chosen NCS is quasiparallel to hers (in her NCS). Denote by $\mathbf{w}$ Chipso's velocity as seen by Arthur.

We now discuss the relationships among the coordinates used by A, B, and C.

3. THE MINKOWSKI–LORENTZ MATRIX

As usual, we denote the speed of light in a vacuum by c (as measured in an inertial system), and the square root of minus 1 by i.

The ratio $\mathbf{v}/c$ is denoted by uppercase bold $\mathbf{V}$ and v/c is denoted by V. The components of $\mathbf{V}$ are denoted by V_1, V_2, and V_3 in a specified spatial coordinate system. We think of $\mathbf{V}$ as a column vector with these three components. The scalar $\mathbf{V}^T\mathbf{V} = V^2$ equals $V_1^2 + V_2^2 + V_3^2$. It must "subceed" 1. We use similar notations in relation to $\mathbf{U}$ and $\mathbf{W}$. Let

$$\gamma = (1 - V^2)^{-1/2}, \qquad \delta = (1 - U^2)^{-1/2}, \qquad \epsilon = (1 - W^2)^{-1/2} \tag{1}$$

For the ordinary familiar or "special" Lorentz transformation, from A's coordinate system to B's, one assumes that both x axes are chosen to be parallel to **v**. But without this restriction we have the more general Lorentz transformation, said in the literature to be "without rotation" if B's NCS is quasiparallel to A's as seen by A. This quasiparallel transformation [for example, Arzeliès (1966, p. 74), who cites the original (1959) French edition of Tonnelat (1966); Møller (1972, p. 41); or, in effect, Thomas (1927, p. 5)], with the obvious meaning for the subscript B, is

$$\begin{bmatrix} x_B \\ y_B \\ z_B \\ ict_B \end{bmatrix} = \begin{bmatrix} 1 + (\gamma - 1)\frac{V_1^2}{V^2}, & (\gamma - 1)\frac{V_1V_2}{V^2}, & (\gamma - 1)\frac{V_1V_3}{V^3}, & i\gamma V_1 \\ (\gamma - 1)\frac{V_2V_1}{V^2}, & 1 + (\gamma - 1)\frac{V_2^2}{V^2}, & (\gamma - 1)\frac{V_2V_3}{V^2}, & i\gamma V_2 \\ (\gamma - 1)\frac{V_3V_1}{V^2}, & (\gamma - 1)\frac{V_3V_2}{V^2}, & 1 + (\gamma - 1)\frac{V_3^2}{V^2}, & i\gamma V_3 \\ -i\gamma V_1, & -i\gamma V_2, & -i\gamma V_3, & \gamma \end{bmatrix} \begin{bmatrix} x_A \\ y_A \\ z_A \\ ict_A \end{bmatrix} \tag{2}$$

which is a linear transformation with *Minkowski–Lorentz matrix* $M(\mathbf{V})$. This matrix can be written more succinctly, in partioned form, as

$$M(\mathbf{V}) = \begin{bmatrix} I_3 + (\gamma - 1)\mathbf{V}\mathbf{V}^T/V^2, & i\gamma\mathbf{V} \\ -i\gamma\mathbf{V}^T, & \gamma \end{bmatrix} \tag{3}$$

which generalizes the ordinary or special transformation matrix

$$\begin{bmatrix} \gamma & 0, & 0, & i\gamma V \\ 0, & 1, & 0, & 0 \\ 0, & 0, & 1, & 0 \\ -i\gamma V, & 0, & 0, & \gamma \end{bmatrix} \tag{4}$$

In the definition (3), I_3 denotes the three-by-three unit matrix. Note that $\mathbf{V}\mathbf{V}^T$ is a symmetric three-by-three matrix of rank one, and that $M(\mathbf{V})$ is Hermitian. It is also pseudo-orthogonal [not unitary, but "complex orthogonal" in the terminology of Mathematical Society of Japan (1977), p. 269J] in the sense that $M(\mathbf{V})M(\mathbf{V})^T = I_4$. We now prove this assertion. Hold in mind that $\mathbf{V}^T\mathbf{V} = V^2$. We have

$$M(\mathbf{V})M(\mathbf{V})^T$$
$$= \begin{bmatrix} I_3 + (\gamma - 1)\frac{\mathbf{V}\mathbf{V}^T}{V^2}, & i\gamma\mathbf{V} \\ -i\gamma\mathbf{V}^T, & \gamma \end{bmatrix} \begin{bmatrix} I_3 + (\gamma - 1)\frac{\mathbf{V}\mathbf{V}^T}{V^2}, & -i\gamma\mathbf{V} \\ i\gamma\mathbf{V}^T, & \gamma \end{bmatrix} \tag{5}$$

The top left-hand three-by-three matrix in this product is

$$I_3 + 2(\gamma - 1)\frac{\mathbf{V}\mathbf{V}^T}{V^2} + (\gamma - 1)^2\frac{\mathbf{V}\mathbf{V}^T\mathbf{V}\mathbf{V}^T}{V^4} - \gamma^2\mathbf{V}\mathbf{V}^T$$

$$= I_3 + \frac{\mathbf{V}\mathbf{V}^T}{V^2}[2(\gamma - 1) + (\gamma - 1)^2 - \gamma^2 V^2] = I_3$$

because the bracketed expression equals $\gamma^2(1 - V^2) - 1 = 0$.

The top right-hand three-by-one matrix (a column 3-vector) in the product (5) is

$$-i\gamma\mathbf{V} - i\gamma(\gamma - 1)\frac{\mathbf{V}\mathbf{V}^T\mathbf{V}}{V^2} + i\gamma^2\mathbf{V} = -i\mathbf{V}[\gamma + \gamma(\gamma - 1) - \gamma^2] = \mathbf{0}$$

Similarly, or because the product of any matrix with its transpose is symmetric, the bottom left-hand horizontal vector is $\mathbf{0}^T$. The bottom right-hand element is

$$-\gamma^2\mathbf{V}^T\mathbf{V} + \gamma^2 = \gamma^2(1 - V^2) = 1$$

Thus

$$M(\mathbf{V})M(\mathbf{V})^T = I_4 \tag{6}$$

so the pseudo-orthogonality of $M(\mathbf{V})$ is established. Note that the top left-hand three-by-three matrix of $M(\mathbf{V})$ is *not* orthogonal. It is symmetric.

Now $M(\mathbf{V})^T = M(-\mathbf{V})$, so

$$M(\mathbf{V})M(-\mathbf{V}) = I_4 \qquad \text{and} \qquad M(\mathbf{V})^T M(\mathbf{V}) = I_4 \tag{7}$$

and therefore

$$[M(\mathbf{V})]^{-1} = M(-\mathbf{V}) \tag{8}$$

4. THE MEANING OF A QUASIPARALLEL TRANSFORMATION

Consider two events $(\xi_A, ict_A)^T$ and $(\xi_A', ict_A)^T$, which are simultaneous in Arthur's NCS, indeed in any of his NCSs. The vector joining them is

$$\begin{bmatrix} \xi_A' - \xi_A \\ 0 \end{bmatrix} = \begin{bmatrix} \Delta\xi_A \\ 0 \end{bmatrix}$$

The corresponding vector in Bertha's NCS, chosen to make $M(\mathbf{V})$ appropriate (see below), is

$$\begin{bmatrix} \Delta\xi_B \\ ic\Delta t_B \end{bmatrix} = M(\mathbf{V})\begin{bmatrix} \Delta\xi_A \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} I_3 + (\gamma - 1)\mathbf{V}\mathbf{V}^T/V^2, & i\gamma\mathbf{V} \\ -i\gamma\mathbf{V}^T, & \gamma \end{bmatrix}\begin{bmatrix} \Delta\xi_A \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} [I_3 + (\gamma - 1)\mathbf{V}\mathbf{V}^T/V^2]\Delta\xi_A \\ -i\gamma\mathbf{V}^T\Delta\xi_A \end{bmatrix}$$

(Thus the two events are not simultaneous in B's chosen NCS unless $\mathbf{V}$ is orthogonal to $\Delta\xi_A$.) The spatial vector $\Delta\xi_A$ is transformed into

$$\Delta\xi_B = [I_3 + (\gamma - 1)\mathbf{V}\mathbf{V}^T/V^2]\Delta\xi_A \tag{9}$$

Now $\Delta\xi_A$ can be expressed in the form $\lambda\mathbf{V} + \mu\mathbf{S}$, where the spatial vector $\mathbf{S}$ is orthogonal to $\mathbf{V}$, and λ and μ are real scalars. Since (9) is a linear transformation, it is sufficient to find its effect on $\mathbf{V}$ and $\mathbf{S}$ separately. We have

$$\begin{aligned}[I_3 + (\gamma - 1)\mathbf{V}\mathbf{V}^T/V^2]\mathbf{V} &= \mathbf{V} + (\gamma - 1)\mathbf{V}\mathbf{V}^T\mathbf{V}/V^2 \\ &= \gamma\mathbf{V} \qquad (\text{because } \mathbf{V}^T\mathbf{V} = V^2)\end{aligned}$$

Again

$$[I_3 + (\gamma - 1)\mathbf{V}\mathbf{V}^T/V^2]\mathbf{S} = \mathbf{S} \qquad (\text{because } \mathbf{V}^T\mathbf{S} = 0)$$

Therefore

$$\Delta\xi_B = \lambda\gamma\mathbf{V} + \mu\mathbf{S} \tag{10}$$

Thus any component of $\Delta\xi_A$ orthogonal to $\mathbf{V}$ is the same in Bertha's NCS, while the component parallel to $\mathbf{V}$ remains parallel to $\mathbf{V}$ but is multiplied by γ. So it makes some sense to say that B's NCS is quasiparallel to A's, but it would be misleading to call it parallel because $\Delta\xi_B$ is parallel to $\Delta\xi_A$ only when $\lambda = 0$ or $\mu = 0$. Another description in the literature is that $M(\mathbf{V})$ effects a transformation *without rotation*, the meaning of which will soon be clarified.

The Minkowski–Lorentz matrix for going from B to A is $[M(\mathbf{V})]^{-1}$ and this is equal to $M(-\mathbf{V})$ [see equation (8)]. This shows that, in Bertha's NCS, Arthur's velocity is $-\mathbf{v}$ and that his NCS is quasiparallel to hers. Thus quasiparallelism is a "mutual" attribute, as we claimed earlier, before giving a rigorous definition.

Before Arthur and Bertha were introduced they might not have known what the velocity $\mathbf{v}$ was and there would have been no reason for their NCSs to be quasiparallel. There might be many other observers that could influence the selections of NCSs by Arthur and Bertha, or they might choose their NCSs in a random manner. So there is nothing sacrosanct about quasiparallelism. Thus a still more general transformation from A to B is

$$\begin{bmatrix} R & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} M(\mathbf{V}) \tag{11}$$

where R represents a spatial nonrotating rotation, that is, a three-by-three proper orthogonal matrix. It depends only on what NCSs are chosen by A and B.

The basic assumption of the theory is that, whatever NCSs are chosen, then the transformation from A's coordinates to B's will necessarily be of the form (11). If A chooses a specific NCS, then there will be an NCS that B could choose that would make $R = I_3$. In that case the transformation from A to B can be called quasiparallel (or mutually quasiparallel).

5. INVARIANCE OF THE RELATIVISTIC "INTERVAL"

Just as an orthogonal transformation preserves lengths, similarly the pseudo-orthogonal transformation $M(\mathbf{V})$ preserves the relativistic squared interval

$$(\Delta x_A)^2 + (\Delta y_A)^2 + (\Delta z_A)^2 - c^2(\Delta t_A)^2$$

that is, the interval squared is the same for B as for A. A proof follows from (7) because the interval squared is

$$\begin{bmatrix} \Delta\xi_B \\ ic\Delta t_B \end{bmatrix}^T \begin{bmatrix} \Delta\xi_B \\ ic\Delta t_B \end{bmatrix} = \begin{bmatrix} \Delta\xi_A \\ ic\Delta t_A \end{bmatrix}^T M(\mathbf{V})^T M(\mathbf{V}) \begin{bmatrix} \Delta\xi_A \\ ic\Delta t_A \end{bmatrix}$$
$$= \begin{bmatrix} \Delta\xi_A \\ ic\Delta t_A \end{bmatrix}^T \begin{bmatrix} \Delta\xi_A \\ ic\Delta t_A \end{bmatrix}$$

The most general transformation matrix (11) also leaves the squared interval unchanged.

6. HOW TO DO THE NUMERICAL CALCULATIONS, AND THE BROTHER OF $M(\mathbf{V})$

In numerical examples, it is convenient, at least in current SAS, to avoid the use of complex numbers by means of the following simple observation. To form the product of two conformably partitioned matrices of the form

$$\begin{bmatrix} M_1 & iM_2 \\ -iM_3 & M_4 \end{bmatrix} \begin{bmatrix} N_1 & iN_2 \\ -iN_3 & N_4 \end{bmatrix} \tag{12}$$

compute the product

$$\begin{bmatrix} M_1, & -M_2 \\ -M_3, & M_4 \end{bmatrix} \begin{bmatrix} N_1, & -N_2 \\ -N_3, & N_4 \end{bmatrix} \tag{13}$$

and then, in the product, replace the minus sign by i and $-i$ in the top right-hand and bottom left-hand parts, respectively. In our applications this converts products of pseudo-orthogonal Hermitian matrices into products of real symmetric matrices. In this way we confirmed numerically, in a short and almost instantaneous SAS program, that $M(\mathbf{V})$, $M(\mathbf{U})$, and $M(\mathbf{W})$ have inverses $M(-\mathbf{V})$, $M(-\mathbf{U})$, and $M(-\mathbf{W})$ (where $\mathbf{W}$ is given below). We used the data

$$\begin{aligned} V_1 &= 0.512, \quad V_2 = 0.253, \quad V_3 = 0.759 \\ U_1 &= 0.091, \quad U_2 = 0.151, \quad U_3 = 0.863 \end{aligned} \tag{14}$$

obtained from the first column of the table of random numbers in Fisher and Yates (1953).

Another readily seen relationship between the "brother" matrices

$$M = \begin{bmatrix} M_1, & iM_2 \\ -iM_3, & M_4 \end{bmatrix} \quad \text{and} \quad M_{\text{real}} = \begin{bmatrix} M_1, & -M_2 \\ -M_3, & M_4 \end{bmatrix} \tag{15}$$

is that, if M_{real} has an eigenvector

$$\begin{bmatrix} \mathbf{a} \\ b \end{bmatrix}$$

with eigenvalue λ, then M has an eigenvector

$$\begin{bmatrix} \mathbf{a} \\ ib \end{bmatrix}$$

with the same eigenvalue λ. (Note that if an eigenvector is multiplied by a nonzero scalar, the result is regarded as the *same* eigenvector.) In particular, the Minkowski–Lorentz matrix $M(\mathbf{V})$ has as its brother the *Lorentz* matrix

$$L(\mathbf{V}) = \begin{bmatrix} I_3 + (\gamma - 1)\mathbf{V}\mathbf{V}^T/V^2, & -\gamma\mathbf{V} \\ -\gamma\mathbf{V}^T, & \gamma \end{bmatrix} \tag{16}$$

and we see that $M(\mathbf{V})$ and $L(\mathbf{V})$ have the same eigenvalues and closely related eigenvectors. We shall see later that the eigenvalues are

$$1, \quad 1, \quad \left(\frac{1+V}{1-V}\right)^{1/2}, \quad \text{and} \quad \left(\frac{1-V}{1+V}\right)^{1/2} \tag{17}$$

The product of the eigenvalues is $+1$, so the determinant is $+1$ and $M(\mathbf{V})$ is a *proper* pseudo-orthogonal matrix. It is interesting that the nonunit eigenvalues are equal to the Doppler factors as given, for example, by Einstein (1923, p. 56). He says (Einstein, 1923, p. 58) that it is "remarkable that the energy and the frequency of a light complex vary with the state of motion of the observer in accordance with the same law." We have now seen the same formulas occurring in yet a third context. This is discussed further in Section 12.

7. THE ROTATION "PARADOX"

If B's chosen NCS is quasiparallel to A's, and C's is quasiparallel to B's, then the transformation from A to C is obtained by transforming from A to B and then from B to C, by the formula

$$\begin{bmatrix} \xi_C \\ ict_C \end{bmatrix} = M(\mathbf{U})\begin{bmatrix} \xi_B \\ ict_B \end{bmatrix} = M(\mathbf{U})M(\mathbf{V})\begin{bmatrix} \xi_A \\ ict_A \end{bmatrix} \tag{18}$$

where ξ_C is a self-explanatory notation. So the matrix of the transformation from A to C, under the assumptions just made, is

$$M(\mathbf{U})M(\mathbf{V}) \tag{19}$$

not $M(\mathbf{V})M(\mathbf{U})$. (Hermitian matrices, even two-by-two, resemble symmetric matrices in that they do not commute in general.)

It might be (incorrectly) conjectured that if C's chosen NCS is quasiparallel to B's (which is quasiparallel to A's), then C's is quasiparallel to A's. But this conjecture would be based on the "tyranny of words" or on a false analogy with Euclidean geometry, especially if we had used the word "parallel" instead of "quasiparallel." It turns out that, in general,

$$M(\mathbf{W}) \neq M(\mathbf{U})M(\mathbf{V}) \tag{20}$$

because, under the assumptions made, the NCSs of A and C turn out not to be quasiparallel unless $\mathbf{U}$ is parallel to $\mathbf{V}$ in Bertha's "opinion." [Actually if two vectors are parallel in one inertial system, then they are parallel in all inertial systems, for the Lorentz transformation is linear and transforms finite points to finite points. For a more explicit proof see Good (1994, Endnote 1).] This at first somewhat surprising fact was used as the basis of his work on the "kinematics of an electron with [spinning about] an axis" by Thomas (1927). He says, on his first page, "The main fact used is that the combination of two 'Lorentz transformations without rotation' in general is not of the same form but is equivalent to a Lorentz transformation with a rotation." He does not prove this, nor cite a reference, nor explain clearly what he means, but he regards it as well known. Let us call (20), or more precise forms of it, the *relativistic nonrotating rotation "paradox,"* or, for short, the *rotation "paradox."* It, and Thomas's seeming deduction from it, surprised even the "cognoscenti of the relativity theory (Einstein included!)" (Uhlenbeck, 1976, p. 48). Thomas's paper is hard to understand and was based on an obsolete model of the electron (superseded by Dirac's model). According to Pais (1982, p. 144) it "took Pauli a few weeks before he grasped Thomas's point." The paper has been thought to explain a missing factor of 2 regarding the Zeeman effect of a magnetic field on an atomic spectrum. [The Zeeman effect is discussed, for example, by Jen (1991). He does not cite Thomas;

nor does Dirac (1958).] That factor of 2 is called the "Thomas effect" or "Thomas factor." What Thomas describes as "well known," Uhlenbeck much later calls a "forgotten relativistic effect."

Just possibly the factor of 2 is related to Good (1991, p. 594), where the circular clock paradox is discussed and where an interesting factor of 2 occurs and presumably can be explained in terms of general relativity. But the present paper is essentially concerned with inertial frames and therefore not with rotating systems such as occur in Thomas' model of a spinning electron and in the circular clock paradox.

Since Chipso is in an inertial system we know it must be moving with some uniform velocity **w** in Arthur's NCS if of course we assume STR. The NCS used by Chipso, quasiparallel to Bertha's, is one of the ∞^3 NCSs available to it, and therefore must be a spatial nonrotating rotation of the one quasiparallel to A's NCS. At this stage of the discussion, this nonrotating rotation might be guessed, incorrectly, to be the null rotation. Thus, we know (if we assume STR), this knowledge being equivalent to Møller's expression "for physical reasons" (Møller, 1972, p. 52), that

$$M(\mathbf{W}) = \begin{bmatrix} I_3 + (\epsilon - 1)\mathbf{W}\mathbf{W}^T/W^2, & i\epsilon\mathbf{W} \\ -i\epsilon\mathbf{W}^T, & \epsilon \end{bmatrix}$$

$$= \begin{bmatrix} R_3, & \mathbf{0} \\ \mathbf{0}^T, & 1 \end{bmatrix} M(\mathbf{U})M(\mathbf{V}) \tag{21}$$

where R_3 is a proper orthogonal three-by-three matrix. The reader should check that this (spatial) rotation has no effect on the bottom right-hand (scalar) element of $M(\mathbf{U})M(\mathbf{V})$ nor on the bottom left-hand one-by-three matrix (horizontal vector). Therefore

$$\epsilon = \gamma\delta(1 + \mathbf{U}^T\mathbf{V}) \tag{22}$$

which agrees with formula (11a) of Pauli (1958) or (2) of Good (1994) (of course $\mathbf{U}^T\mathbf{V} = \mathbf{V}^T\mathbf{U}$ because the transpose of a scalar is itself), and

$$-i\epsilon\mathbf{W}^T = -i\delta\mathbf{U}^T[I_3 + (\gamma - 1)\mathbf{V}\mathbf{V}^T/V^2] - i\delta\gamma\mathbf{V}^T \tag{23}$$

Therefore, by taking the transpose, we see that

$$\mathbf{W} = \frac{\mathbf{U} + \mathbf{V}[\gamma + (\gamma - 1)\mathbf{U}^T\mathbf{V}/V^2]}{(1 + \mathbf{U}^T\mathbf{V})\gamma} \tag{24}$$

(coplanar with **U** and **V**). This agrees with Møller (1972, p. 52), formula (2.59), who obtained it by a very different method (depending on differentiation of spatial coordinates with respect to time). As a further check we can calculate W^2 from

$$W^2 = W_1^2 + W_2^2 + W_3^2 \tag{25}$$

$$= [U^2 + V^2 + 2\mathbf{U}^T\mathbf{V} + (\mathbf{U}^T\mathbf{V})^2 - U^2V^2](1 + \mathbf{U}^T\mathbf{V})^{-2} \tag{26}$$

and from

$$W^2 = 1 - 1/\epsilon^2 \tag{27}$$

To check by high-school algebra that (26) follows from (24), we start from $W^2 = \mathbf{W}^T\mathbf{W}$. The numerator of this product is

$$\left\{\mathbf{U}^T + \mathbf{V}^T\left[\gamma\left(1 + \frac{\mathbf{U}^T\mathbf{V}}{V^2}\right) - \frac{\mathbf{U}^T\mathbf{V}}{V^2}\right]\right\}\left\{\mathbf{U} + \mathbf{V}\left[\gamma\left(1 + \frac{\mathbf{U}^T\mathbf{V}}{V^2}\right) - \frac{\mathbf{U}^T\mathbf{V}}{V^2}\right]\right\}$$

The coefficient of γ is readily found to vanish, and then we can deduce (26) by using the fact that $\gamma^2 = (1 - V^2)^{-1}$ together with five or six lines of manipulation. The equality of (26) and (27) then follows readily from the expression for ϵ in (22). The matrix method is much easier. We have checked the equality of (26) and (27) numerically, by using the data (14).

A formula equivalent to (26) was given by Einstein in his original (1905) paper, based on differentiation [like Møller's derivation of (24)], and was published in the English version (Einstein, 1923, p. 50), but with a misprint. [See also, for example, Good (1994), where the formula was used for refuting another attack on the self-consistency of STR.] To be more precise, Einstein considered a particle moving with velocity **u** in what we have called Bertha's NCS, this particle acting the part of our third observer (Chipso).

8. THE PROOF OF THE ROTATION "PARADOX" BEYOND ANY DOUBT

Knowing the formula for **W**, we can calculate $M(\mathbf{W})$. Let us for the moment drop the assumption (21), which was based on "physical reasons." Let us define a four-by-four matrix $R(\mathbf{U}, \mathbf{V})$ as

$$R(\mathbf{U}, \mathbf{V}) = M(\mathbf{W})[M(\mathbf{V})]^{-1}[M(\mathbf{U})]^{-1} \tag{28}$$

which implies

$$M(\mathbf{W}) = R(\mathbf{U}, \mathbf{V})M(\mathbf{U})M(\mathbf{V}) \tag{29}$$

$M(\mathbf{U})$ and $M(\mathbf{V})$ are of course nonsingular, so $R(\mathbf{U}, \mathbf{V})$ is properly defined. We took the numerical values for $\mathbf{V}$ and $\mathbf{U}$ given by (14), and, by means of a SAS program of negligible running time, we obtained the output

$$R(\mathbf{U}, \mathbf{V}) = \begin{bmatrix} 0.9611333, & 0.0291664, & 0.27454, & 0 \\ -0.050347, & 0.996246, & 0.0704205, & 0 \\ -0.271455, & -0.081506, & 0.9589936, & 0 \\ 0, & 0, & 0, & 1 \end{bmatrix} \tag{30}$$

where each number that is printed here as 0 was, in the computer output, less than 10^{-13} in absolute value. The product of $R(\mathbf{U}, \mathbf{V})$ times its transpose is found to be I_4 correct to at least six decimal places in all 16 entries. This verifies that formula (21) is correct far beyond any shadow of doubt, more convincingly (for anyone who runs the brief SAS program with the same or different input data) than pages of heavy algebra could ever be. *Elegant* algebraic proofs often give more insight, but numerical checks can give more personal confidence.

Thus we now know with total confidence that $R(\mathbf{U}, \mathbf{V})$ is of the form

$$\begin{bmatrix} R_3, & \mathbf{0} \\ \mathbf{0}^T, & 1 \end{bmatrix} \tag{31}$$

and it can be calculated as

$$M(\mathbf{W})M(-\mathbf{V})M(-\mathbf{U}) \tag{32}$$

R_3 cannot be calculated as the product of the three-by-three matrices in the top left-hand corners of $M(-\mathbf{V})$ and $M(-\mathbf{U})$. Unlike R_3, they are not even orthogonal.

To summarize, if C chooses its NCS quasiparallel to A's instead of quasiparallel to B's, then the Minkowski–Lorentz matrix for going from A to C will be given [see (29) and (31)] by

$$M(\mathbf{W}) = \begin{bmatrix} R_3, & \mathbf{0} \\ \mathbf{0}^T, & 1 \end{bmatrix} M(\mathbf{U})M(\mathbf{V}) \tag{33}$$

instead of by $M(\mathbf{U})M(\mathbf{V})$. In other words, its NCS will be the nonrotational spatial rotation R_3 of what might have been naively thought (that is, if it had been assumed that quasiparallelism is "transitive").

If we rewrite (33) as

$$M(\mathbf{U})M(\mathbf{V}) = \begin{bmatrix} R_3^{-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} M(\mathbf{W}) \tag{33a}$$

it shows very clearly what is meant by saying that "the combination of two 'Lorentz transformations without rotation' ... is equivalent to a Lorentz transformation with a rotation." The rotation R_3 is a *nonrotating spatial rotation* of Chipso's reference frame, from (i) the one quasiparallel to Bertha's to (ii) the one quasiparallel to Arthur's. Chipso's choice would affect the details of the calculations as in the very elementary example mentioned in the Introduction. But this choice ought not to affect any physical deductions such as the extent to which Chipso's clock runs slower than Arthur's in Arthur's NCS. Indeed, we have already mentioned, following equation (21), that the spatial rotation R_3 does not affect the expressions for ϵ and **W**.

Phipps (1986, p. 267), says ironically, in relation to Møller's expression "for physical reasons," "I think what he means is *for religious reasons*" (Phipps's italics). We have now seen that Møller's "faith" is miraculously justified in the sense that the word *physical* could legitimately be replaced by *mathematical*, thus yet again confirming the self-consistency of STR. (Miracles can convert heretics.) The numerical check shows that an elementary algebraic proof of (33) must certainly "exist" in the Platonic sense, that is, in the universe of mathematical truths. Readers who begin to write out an elementary proof will see how unwieldy it becomes. I have not attempted such a proof by using Mathematica. This would be of interest, but it is not necessary. Only an elegant or short algebraic proof would be worthwhile.

Note that Chipso does not have to choose its NCS to be quasiparallel to Bertha's. It could equally use the one quasiparallel to Arthur's (and therefore nonrotationally rotated with respect to Bertha's) or any other of the ∞^3 NCSs available to it, such as one quasiparallel to an inertial demon named Diabolus, Master of the Dark Matter. No physical rotation of the laboratory, nor of Chipso's swivel chair, is required.

When talking to Bertha, Chipso might find it convenient and polite to choose its NCS quasiparallel to hers, but when talking to Arthur it might, like a politician, change its mind and switch to one quasiparallel to his. Because Chipso's choice of NCS is subject so much to its whim and to its rapport with Arthur, Bertha, Diabolus, or Elvira, the question arises whether the nonrotating spatial rotation R_3 has any physical significance as distinct from its mathematical or computational convenience. Thomas (1927) believed that it does, but Phipps (1986, pp. 266–267) controversially questions it, his doubts being based partly on his experiment with a rotating disc (Phipps, 1974). Perhaps many physicists have accepted Thomas's argument in accordance with "Blackett's Law." Blackett (1946) remarked seriously that a physicist regards an argument as correct if it reaches the right conclusion. But I think he would have agreed that to explain a missing factor of 2 is less impressive than to explain a more complicated factor, such as $\pi + e$, when the argument has a given level of obscurity.

9. A SEMANTIC CONFUSION CONCERNING ROTATION

Galeczki (see the citations in our Introduction) regarded the rotation "paradox" as contradicting the definition of an inertial frame of reference. He overlooked that the rotation is nonrotating. Perhaps he was misled by the fact that Thomas (1927) discussed nonrotating rotations and rotating rotations (of an electron) almost in the same breath. These remarks are not intended to rule out the possibility that Thomas' paper is correct, but to suggest that someone who is sure "they" understand the paper might explain it to the non-Pauli hoi polloi without relying on Blackett's Law.

10. THE REAL LORENTZ MATRIX $L(\mathbf{V})$

Equation (2) is equivalent to

$$\begin{bmatrix} \xi_B \\ ct_B \end{bmatrix} = L(\mathbf{V}) \begin{bmatrix} \xi_A \\ ct_A \end{bmatrix} \tag{34}$$

where

$$L(\mathbf{V}) = \begin{bmatrix} I_3 + (\gamma - 1)\mathbf{V}\mathbf{V}^T/V^2, & -\gamma\mathbf{V} \\ -\gamma\mathbf{V}^T, & \gamma \end{bmatrix} \tag{35}$$

which we call the (real) Lorentz matrix as in Section 6. This was the matrix we used in our SAS program [more precisely, we happened to use $L(-\mathbf{V})$], but $L(\mathbf{V})$ can also be used for theoretical work instead of $M(\mathbf{V})$. Whereas $M(\mathbf{V})$ is Hermitian and pseudo-orthogonal, $L(\mathbf{V})$ is symmetric and real, but not orthogonal, although its determinant is 1. A merit of $M(\mathbf{V})$ is that it fits in well with the concept of pseudo-Euclidean space in which the invariant relativistic squared "interval" is thought of as $(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 + [\Delta(ict)]^2$, but the use of $L(\mathbf{V})$ is less abstract and easier to program. The relationship between $L(\mathbf{V})$ and $M(\mathbf{V})$ is close, and it is easy to switch from one notation to the other one.

11. THE SPECTRAL ANALYSIS OF $L(\mathbf{V})$

The nonunit eigenvalues of $L(\mathbf{V})$ are, as for $M(\mathbf{V})$, $\gamma(1 \pm V)$, corresponding respectively to the eigenvectors

$$\begin{bmatrix} \mp\mathbf{V} \\ V \end{bmatrix} \quad \text{(where } V > 0\text{)}$$

because

$$L(\mathbf{V}) \begin{bmatrix} \mp\mathbf{V} \\ V \end{bmatrix} = \gamma(1 \pm V) \begin{bmatrix} \mp\mathbf{V} \\ V \end{bmatrix} \tag{36}$$

The proof is simply that

$$\begin{bmatrix} I_3 + (\gamma - 1)\mathbf{V}\mathbf{V}^T/V^2, & -\gamma\mathbf{V} \\ -\gamma\mathbf{V}^T, & \gamma \end{bmatrix}\begin{bmatrix} \mathbf{V} \\ V \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{V} + (\gamma - 1)\mathbf{V} - \gamma V\mathbf{V} \\ -\gamma V^2 + \gamma V \end{bmatrix} = \gamma(1 - V)\begin{bmatrix} \mathbf{V} \\ V \end{bmatrix}$$

etc. Similarly, or from Section 6,

$$M(\mathbf{V})\begin{bmatrix} \pm\mathbf{V} \\ iV \end{bmatrix} = \gamma(1 \mp V)\begin{bmatrix} \pm\mathbf{V} \\ iV \end{bmatrix} \tag{37}$$

The two eigenvectors in (36) are orthogonal in real four-dimensional space with coordinates (x, y, z, ct), that is,

$$\begin{bmatrix} \mathbf{V} \\ V \end{bmatrix}^T\begin{bmatrix} -\mathbf{V} \\ V \end{bmatrix} = 0 \tag{38}$$

The two eigenvalues just mentioned can also be written in the forms

$$\left(\frac{1 + V}{1 - V}\right)^{1/2} \quad \text{and} \quad \left(\frac{1 - V}{1 + V}\right)^{1/2} \tag{39}$$

Again, if **S** is a spatial vector through the origin and orthogonal to **V**, then

$$L(\mathbf{V})\begin{bmatrix} \mathbf{S} \\ 0 \end{bmatrix} = \begin{bmatrix} I_3 + (\gamma - 1)\mathbf{V}\mathbf{V}^T/V^2, & -\gamma\mathbf{V} \\ -\gamma\mathbf{V}^T, & \gamma \end{bmatrix}\begin{bmatrix} \mathbf{S} \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{S} \\ 0 \end{bmatrix}$$

That is,

$$\begin{bmatrix} \mathbf{S} \\ 0 \end{bmatrix}$$

spans an eigenplane with eigenvalue 1. This plane contains the origin. It may be called the *unit eigenplane* (corresponding to **V**). Thus the complete set of eigenvalues of $L(\mathbf{V})$, and equally of $M(\mathbf{V})$, is

$$1, \quad 1, \quad \left(\frac{1 + V}{1 - V}\right)^{1/2}, \quad \left(\frac{1 - V}{1 + V}\right)^{1/2}$$

as promised in Section 6. The check may be noted that the sum of the four eigenvalues is $2 + 2\gamma$, which is the trace of $L(\mathbf{V})$ and of $M(\mathbf{V})$.

On the light cone we must have

$$(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 = (\Delta ct)^2 \tag{40}$$

so an arbitrary ray on the light cone is of the form

$$\begin{bmatrix} \pm\mathbf{Q} \\ Q \end{bmatrix} \quad \text{(where } Q > 0\text{)}$$

if the coordinates used are (x, y, z, ct). The Lorentz transform of

$$\begin{bmatrix} \mathbf{Q} \\ Q \end{bmatrix}$$

is

$$L(\mathbf{V})\begin{bmatrix} \mathbf{Q} \\ Q \end{bmatrix} = \begin{bmatrix} \mathbf{Q} + \mathbf{V}[(\gamma - 1)\mathbf{V}^T\mathbf{Q}/V^2 - \gamma Q] \\ \gamma(Q - \mathbf{V}^T\mathbf{Q}) \end{bmatrix} \tag{41}$$

(Note that $Q - \mathbf{V}^T\mathbf{Q} > 0$.) It can be verified by a page of algebra that (41) is of the form

$$\begin{bmatrix} \mathbf{P} \\ P \end{bmatrix} \quad (P > 0) \tag{42}$$

and this again lies on the light cone (as we knew in advance from Section 5). Similarly, the transform of

$$\begin{bmatrix} -\mathbf{Q} \\ Q \end{bmatrix}$$

is of the form

$$\begin{bmatrix} -\mathbf{P} \\ P \end{bmatrix}$$

where P is again positive. In particular, if $\mathbf{Q} = \pm\mathbf{V}$, we quickly recover the formulas for the eigenvectors on the light cone and for their eigenvalues.

The light cone "as a whole" is the same set of events for A and B (and indeed for all "inertial observers"), but the only light rays whose coordinates are the same for A and B are the two eigenvectors

$$\begin{bmatrix} \pm\mathbf{V} \\ V \end{bmatrix}$$

The part of the plane containing these two eigenvectors, and lying within or on the light cone, is spanned by vectors that can be parametrized by an angle α and can be expressed in two simple ways, thus

$$\cos^2(\alpha/2)\begin{bmatrix}\mathbf{V}\\V\end{bmatrix} + \sin^2(\alpha/2)\begin{bmatrix}-\mathbf{V}\\V\end{bmatrix} = \begin{bmatrix}\mathbf{V}\cos\alpha\\V\end{bmatrix} \tag{43}$$

with $0 \leq \alpha \leq \pi$. Each such vector is, in an obvious sense, pseudo-orthogonal to the unit eigenplane and in the Minkowski plane. Its Lorentz transform is

$$\begin{bmatrix} I_3 + (\gamma - 1)\mathbf{V}\mathbf{V}^T/V^2, & -\gamma\mathbf{V} \\ -\gamma\mathbf{V}^T, & \gamma \end{bmatrix} \begin{bmatrix}\mathbf{V}\cos\alpha\\V\end{bmatrix}$$

$$= \gamma\begin{bmatrix}\mathbf{V}(\cos\alpha - V)\\V(1 - V\cos\alpha)\end{bmatrix} \tag{44}$$

(Note the checks when $\alpha = 0$ and $\alpha = \pi$.)

It is suggestive that the Doppler ratio of frequencies in Einstein (1923, p. 56) is

$$\gamma(1 - V\cos\phi) \tag{45}$$

where his ϕ seems to be the same as our α. Thus the ratio of B's to A's time components in (44) is equal to the Doppler ratio. Einstein gives also a formula for aberration, relating the apparent directions of a very distant "fixed star" from two different frames having relative velocity $\mathbf{v}$. For another proof see Whittaker (1953, p. 40). The formula, in Whittaker's notation, is

$$\cos\psi' = \frac{\cos\psi - V}{1 - V\cos\psi} \tag{46}$$

where ψ and ψ' denote the angles between $\mathbf{V}$ and the beam directions as measured in the two frames. The 4-vector on the right side of (44) has the same direction as

$$\begin{bmatrix}\mathbf{V}\cos\alpha'\\V\end{bmatrix} \tag{47}$$

where

$$\cos\alpha' = \frac{\cos\alpha - V}{1 - V\cos\alpha} \tag{48}$$

The identity of form of (46) and (48) can hardly be coincidental, but I have not found an explanation.

12. THE IMAGINARY MINKOWSKI PLANE

For discussing the imaginary Minkowski rotation (using geometry in the field of complex numbers) it is convenient to use the time variable $\tau =$

ict rather than *ct*, and the *M* matrix rather than the *L*. We are now using the word "rotation" in yet a third sense.

Spacetime can be "factorized" into the "direct product" of two planes. One of them is the unit eigenplane, say $\Pi_1(\mathbf{V})$, and the other, say $\Pi_2(\mathbf{V})$, is pseudo-orthogonal to $\Pi_1(\mathbf{V})$. [The planes $\Pi_1(\mathbf{V})$ and $\Pi_2(\mathbf{V})$ have just one point of intersection, namely the origin.] That is, the relativistic scalar product of a vector in $\Pi_1(\mathbf{V})$ and a vector in $\Pi_2(\mathbf{V})$ is always zero. The plane $\Pi_2(\mathbf{V})$ contains $\mathbf{V}$ and the τ axis, and may be called the *Minkowski plane.* We now consider the effect of applying $M(\mathbf{V})$ to the Minkowski plane. Let $\mathbf{V}^{(1)}$ denote the 4-vector

$$\mathbf{V}^{(1)} = \begin{bmatrix} \mathbf{V} \cos \alpha \\ iV \end{bmatrix} \tag{49}$$

Then

$$M(\mathbf{V})\mathbf{V}^{(1)} = \mathbf{V}^{(2)} \tag{50}$$

where

$$\mathbf{V}^{(2)} = \gamma \begin{bmatrix} \mathbf{V}(\cos \alpha - V) \\ iV(1 - V \cos \alpha) \end{bmatrix} \tag{51}$$

The relativistic squared lengths (squared "intervals") of $\mathbf{V}^{(1)}$ and $\mathbf{V}^{(2)}$ are both $-V^2 \sin^2\alpha$, naturally equal because $M(\mathbf{V})$ is pseudo-orthogonal, but the reader is invited to check it from (49) and (51). Note, too, what happens when $\alpha = 0$ or π. The relativistic scalar product of $\mathbf{V}^{(1)}$ and $\mathbf{V}^{(2)}$ is $-\gamma^2 \sin^2\alpha$. So, if θ is the angle from $\mathbf{V}^{(1)}$ to $\mathbf{V}^{(2)}$, we have

$$\cos \theta = \gamma, \qquad \sin \theta = \pm(1 - \gamma^2)^{1/2}, \qquad \tan \theta = \pm iV \tag{52}$$

Thus, holding in mind that $V < 1$, we have

$$\begin{aligned} \theta = \pm\tan^{-1}(iV) &= \pm\left[iV - \frac{(iV)^3}{3} + \frac{(iV)^5}{5} - \cdots \right] \\ &= \pm i \tanh^{-1} V \end{aligned} \tag{53}$$

Since $\mathbf{V}^{(2)}$ is a continuous function of $\mathbf{V}$ and α, if $0 < \alpha < \pi$, the ambiguous sign must always be plus or always minus. But when $\mathbf{V}^T$ is of the form $(V_1, 0, 0, 0)$ we have the familiar special Minkowski–Lorentz transformation where the sign is positive. Thus, for all $\mathbf{V}$ we have

$$\theta = \tan^{-1}(iV) = i \tanh^{-1} V = \frac{i}{2} \log \frac{1 + V}{1 - V} \qquad (0 < \alpha < \pi) \tag{54}$$

which is independent of α. Therefore the Minkowski–Lorentz transformation rotates the Minkowski plane by an imaginary angle equal to i times the "rapidity," as defined, for example, by Eddington (1930, p. 22 with an acknowledgment to A. A. Robb). The rotation is "about" the plane $\Pi_1(\mathbf{V})$ (in the sense in which one talks of a rotation of a rigid body, in three dimensions, about an axis).

For the special Lorentz transformation one is used to talking about a rotation of the (x, τ) plane, but that is merely because the x axis is chosen parallel to $\mathbf{V}$. It is more general to describe the rotation as of the Minkowski plane.

When $V = \pm 1$, the right side of formula (53) takes the form $\pm i\infty$, which is not strictly meaningful, but it is not too bad because, as $V \to \pm 1$, the imaginary angle tends to $\pm i\infty$. It is interesting to make a comparison with planar projective geometry. The projective geometer Askwith (1921, p. 243), for example, states that the line $y = ix$ makes the same angle $\tan^{-1}(i)$ with every line $y = mx$, where $m \neq i$, because

$$\tan^{-1}\left(\frac{i - m}{1 + im}\right) = \tan^{-1}(i) \tag{55}$$

He does not point out that $\tan^{-1}(i) = i\infty$. The physical interpretation is that, however fast you travel, it remains just as difficult to catch up with a light beam, the "Red Queen effect," so to speak. ("It takes all the running you can do to keep in the same place.") In projective geometry, the pair of lines $y = \pm ix$, which is also a circle of zero radius, is known as the pair of isotropic or absolute lines which pass through the circular points at infinity. [They were introduced by J. V. Poncelet in 1822 according to Kline (1972, p. 845).] Similarly, the sphere $x^2 + y^2 + z^2 + \tau^2 = 0$, of radius zero, is an imaginary cone cutting the hyperplane at infinity in an imaginary sphere (Mathematical Society of Japan, 1977, Section 344E). The name *absolute* is also appropriate in the context of relativity theory because of the invariance of the light cone.

13. THE INTUITIVE RESOLUTION OF THE ROTATION "PARADOX"

For the sake of simplicity consider, for the moment, spacetime in $2 + 1 = 3$ dimensions. It was noticed by Euler (see, for example, Lamb, 1929, pp. 2–3) that a rotation of a solid three-dimensional body followed by another such rotation is equivalent to yet a third rotation, in general unique. (In matrix terminology the product of two proper three-by-three orthogonal matrices is a third proper orthogonal matrix. Each of the three has in general a unique real eigenvector with eigenvalue 1.) We have just seen that a "general" Lorentz transformation "without spatial rotation" is equivalent to an imaginary rotation *about* a spatial plane, or about a spatial line if the spacetime has only

2 + 1 dimensions. From this point of view, two consecutive transformations in 2 + 1 dimensions each without spatial rotation are equivalent to two consecutive imaginary rotations about two *distinct* spatial lines, distinct if, in our previous notation, **u** is not parallel to **v**. By Euler's theorem one would expect the result to be equivalent to a single imaginary rotation, corresponding to $M(\mathbf{W})$. This shows that the "paradox" is only to be expected in 2 + 1 dimensions and therefore should not be surprising in 3 + 1 dimensions. Indeed it would be highly paradoxical if it did *not* occur.

ACKNOWLEDGMENTS

I am much indebted to George Galeczki and Thomas E. Phipps, Jr. for their communications, which were friendly in spite of differences of opinion, and for drawing my attention to (i) what I have called the rotation "paradox" and (ii) Thomas (1927). I also thank Michele Marini for the programming.

REFERENCES

Arzelies, H. (1966). *Relativistic Kinematics,* Pergamon Press, New York.
Askwith, E. H. (1921). *A Course in Pure Geometry.* Cambridge University Press, Cambridge.
Blackett, P. M. S. (1946). Personal communication in Donner House, Manchester, England.
Dirac, P. A. M. (1958). *The Principles of Quantum Mechanics,* Oxford University Press, Oxford.
Eddington, A. S. (1930). *The Mathematical Theory of Relativity,* Cambridge University Press, Cambridge.
Einstein, A. (1923). On the electrodynamics of moving bodies [English translation of *Annalen der Physik,* **17**, 891 (1905)], in H. A. Lorentz, A. Einstein, H. Weyl, and H. Minkowski, *The Principle of Relativity,* Methuen, London [reprinted Dover, New York (1952)].
Fisher, R. A., and Yates, F. (1953). *Statistical Tables for Biological, Agricultural and Medical Research,* 4th ed., Oliver and Boyd, Edinburgh.
Galeczki, G. (1993). *Physics Essays,* **6**, 448–450.
Galeczki, G. (1994a). *Chinese Journal of Systems Engineering and Electronics,* **5**, 77–78.
Galeczki, G. (1994b). *Physics Essays,* **7**, 83–86.
Good, I. J. (1991). *Physics Essays,* **4**, 591–595.
Good, I. J. (1994). The self-consistency of the kinematics of special relativity, *Physics Essays,* 7(4), in press.
Jen, C. K. (1991). Zeeman and Stark effects, in *Encyclopedia of Physics,* Rita G. Lerner and George L. Trigg, eds., pp. 1370–1374.
Kline, M. (1972). *Mathematical Thought from Ancient to Modern Times,* Oxford University Press, Oxford.
Lamb, H. (1929). *Higher Mechanics,* Cambridge University Press, Cambridge.
Mathematical Society of Japan (1977). *Encyclopedic Dictionary of Mathematics, I,* Shôkichi Iyanaga and Yukiyosi Kawada, eds., MIT Press, Cambridge, Massachusetts.
Møller, C. (1972). *The Theory of Relativity,* 2nd ed., Clarendon Press, Oxford.
Pais, A. (1982). *Subtle Is the Lord: The Science and the Life of Albert Einstein,* Clarendon Press, Oxford.

Pauli, W. (1958). *Theory of Relativity,* Pergamon Press, London.
Phipps, T. E., Jr. (1974). *Lettere al Nuovo Cimento,* **9**, 467–470.
Phipps, T. E., Jr. (1986). *Heretical Verities,* Classic Non-Fiction Library, Urbana, Illinois.
Thomas, L. H. (1927). The kinematics of an electron with an axis, *Philosophical Magazine 7th Series,* **3**, 1–22.
Tonnelat, M.-A. (1966). *The Principles of Electromagnetic Theory and of Relativity,* Gordon and Breach, New York.
Uhlenbeck, G. E. (1976). Fifty years of spin. Personal reminiscences, *Physics Today,* **1976**(June), 43–48.
Whittaker, E. T. (1953). *History of the Theories of Aether and Electricity: The Modern Theories, 1900–1926,* Nelson, London.