The Thomas rotation

John P. Costella,[a)] Bruce H. J. McKellar,[b)] and Andrew A. Rawlinson[c)]
School of Physics, The University of Melbourne, Parkville, Victoria 3052, Australia

Gerard J. Stephenson, Jr.[d)]
Department of Physics and Astronomy, University of New Mexico, Albuquerque, New Mexico 87131

(Received 2 August 2000; accepted 14 February 2001)

We review why the *Thomas rotation* is a crucial facet of special relativity, that is just as fundamental, and just as ''unintuitive'' and ''paradoxical,'' as such traditional effects as length contraction, time dilation, and the ambiguity of simultaneity. We show how this phenomenon can be quite naturally introduced and investigated in the context of a typical introductory course on special relativity, in a way that is appropriate for, and completely accessible to, undergraduate students. We also demonstrate, in a more advanced section aimed at the graduate student studying the Dirac equation and relativistic quantum field theory, that careful consideration of the Thomas rotation will become vital as modern experiments in particle physics continue to move from unpolarized to polarized cross sections.
[DOI: 10.1119/1.1371010]

I. INTRODUCTION

Recently, a number of the current authors have reviewed how aspects of relativistic quantum mechanics can be appreciated from the point of view of relativistic *classical* mechanics. In Ref. 1, the Foldy–Wouthuysen transformation was reviewed, where it was emphasized that many of the operators of the Dirac equation become, after transformation, completely recognizable from the point of view of classical physics. In Ref. 2, the Feynman–Stueckelberg formulation of antiparticles was reviewed, entirely within the domain of classical mechanics, and it was emphasized that one can make good sense of antiparticle motion without needing to resort to quantum mechanical arguments.

In extending these ideas to the domain of quantum field theory, we have found that there is a third aspect of classical relativistic mechanics that is of crucial theoretical and practical importance, but which rates scarcely a mention in most textbooks on special relativity: the *Thomas rotation*. (In the case of a continuous evolution of infinitesimal rotations, this effect is usually referred to as the Thomas *precession*; but here we are mainly concerned with the more general case of a single, *finite* rotation.) Historically, the relative obscurity of this effect can, perhaps, be traced to the fact that the special theory of relativity was two decades old before Thomas made his discovery. Pais's summary of events[3] is instructive:

> Twenty years later [after his seminal 1905 paper on special relativity], Einstein heard something about the Lorentz group that greatly surprised him. It happened while he was in Leiden. In October 1925 George Eugene Uhlenbeck and Samuel Goudsmit had discovered the spin of the electron and thereby explained the occurrence of the alkali doublets, but for a brief period it appeared that the magnitude of the doublet splitting did not come out correctly. Then Llewellyn Thomas supplied the missing factor, 2, now known as the Thomas factor. Uhlenbeck told me that he did not understand a word of Thomas's work when it first came out. 'I remember that, when I first heard about it, it seemed unbelievable that a relativistic effect could give a factor of 2 instead of something of order v/c . . . . Even the cognoscenti of the relativity theory (Einstein included!) were quite surprised.' At the heart of the Thomas precession lies the fact that a Lorentz transformation with velocity $\mathbf{v}_1$ followed by a second one with velocity $\mathbf{v}_2$ in a different direction does not lead to the same inertial frame as one single Lorentz transformation with the velocity $\mathbf{v}_1+\mathbf{v}_2$. (It took Pauli a few weeks before he grasped Thomas's point.)

It seems remarkable—but, according to the above account, undeniable—that neither Einstein nor Pauli came across the Thomas rotation before 1925. However, the effect we now call the Thomas rotation was known before Thomas's paper. The early history has been traced by Ungar.[4]

Now, most textbooks on special relativity follow the extraordinarily clear exposition of the theory given by Einstein in his seminal paper. Unfortunately, this has meant that little or no attention has usually been given to the ''Thomas effect,'' which has generally been relegated to a brief mention in textbooks on quantum mechanics and atomic structure. As far as we are aware, the best treatment of the Thomas *precession* in a textbook still in print is arguably that contained in Jackson's book on classical electrodynamics.[5] A similar discussion is given in Goldstein,[6] who emphasizes the complexity of the general calculations. This complexity is inhibiting to both the writers and the readers of the textbooks.

That the Thomas rotation, or precession, still puzzles students and their teachers can be discerned from the pages of this journal. In Question #57, MacKeown[7] asks ''... is said to introduce a velocity independent constant factor. Can any simple, convincing, argument be given for this?'' Ungar[4] and Goedecke[8] have countered the complexity by introducing new formalisms, a ''weakly associative-commutative groupoid'' by Ungar, and the tetrad formalism by Goedecke. While they offer useful insights, and emphasize the Thomas rotation, they are not well suited to the introductory course. Muller[9] (in the Appendix) and Philpott[10] (to introduce the main point of his paper) give derivations of the Thomas precession which are related to the present one. But we believe that the straightforward treatment below, and its emphasis on Thomas *rotation*, offers conceptual and pedagogical advantages which make it suitable to an introductory course.

In this paper, we show how an instructive, elementary, and intriguing discussion of the Thomas rotation can be ‘‘grafted on’’ to any standard introductory course on special relativity. As a prerequisite we assume nothing more than the standard expression for a Lorentz boost along the x axis of a system of coordinates. For simplicity, we also make use of the energy–momentum four-vector, as well as matrix multiplication, although such references could be deleted if thought necessary, albeit at the expense of rendering the algebra a little less transparent. (The widespread availability of calculators and computer programs capable of matrix multiplication means that the complexities of the following calculations can be drastically *minimized* by the use of matrix notation—leaving more time for the contemplation of the physical results.) These preliminaries are covered in Sec. II.

In Sec. III, we show how these simple building blocks can be put together to create a sequence of intriguing and completely counterintuitive ‘‘paradoxes.’’ This material could be presented almost verbatim in any introductory course on special relativity.

Section IV provides full solutions and explanations of these elementary Thomas rotation ‘‘paradoxes,’’ and general expressions are derived for the Thomas rotation in arbitrary cases.

In Sec. V, we provide a further ‘‘paradox’’ in the context of the polarization properties of the scattering of a Dirac particle. This example is more advanced, in that it presumes familiarity with at least an introductory level of quantum field theory; and thus it would not usually be appropriate for an introductory course on special relativity. On the other hand, this example is arguably much more *practically* important than the others, in that it shows that real-life calculations of scattering cross sections can be completely erroneous if due regard is not taken of this subtle facet of relativistic kinematics. Section VI provides a full solution of this ‘‘paradox.’’

Finally, Sec. VII summarizes our conclusions.

II. PRELIMINARIES

In this section we review those features of special relativity that we would assume to have been taught in an introductory course before the discussion of Thomas rotation, to set the scene and to establish our notation.

Throughout this paper, we shall use a ‘‘naturalized’’ set of units, in which $c=1$. To convert any expression to SI units, one need simply replace t by ct,$\mathbf{v}$ by $\mathbf{v}/c$,E by E/c^2, and $\mathbf{p}$ by $\mathbf{p}/c$. (Boldface denotes a three-vector.) In Secs. V and VI we shall also use units in which $\hbar=1$.

The Lorentz transformation from a frame S with coordinates (t,x,y,z), to a frame S', moving with respect to S with a velocity v along the x axis, in which the coordinates are (t',x',y',z'), is

$$t'=\gamma(t-vx), \quad x'=\gamma(x-vt), \quad y'=y, \quad z'=z, \tag{1}$$

where

$$\gamma\equiv\frac{1}{\sqrt{1-v^2}}. \tag{2}$$

This transformation, written in matrix notation, is

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}. \tag{3}$$

We shall denote the matrix that effects this boost by velocity v in the x direction as $B_x(v)$:

$$B_x(v)\equiv\begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4}$$

Clearly, a boost by velocity v in the y direction or in the z direction would be effected by

$$B_y(v)\equiv\begin{pmatrix} \gamma & 0 & -\gamma v & 0 \\ 0 & 1 & 0 & 0 \\ -\gamma v & 0 & \gamma & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$
$$B_z(v)\equiv\begin{pmatrix} \gamma & 0 & 0 & -\gamma v \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma v & 0 & 0 & \gamma \end{pmatrix}. \tag{5}$$

The *energy–momentum four-vector*,

$$p^\mu=\begin{pmatrix} E \\ p^x \\ p^y \\ p^z \end{pmatrix}, \tag{6}$$

will play a key role in our analysis. A particle of mass m, at rest in a system of coordinates, has $E=m$ and $\mathbf{p}=\mathbf{0}$. If we boost our system of coordinates by the velocity $-v$ in the x direction (so that, relative to this new system of coordinates, the particle has velocity $+v$ in the x direction), then the application of $B_x(-v)$ yields

$$\begin{pmatrix} E \\ p^x \\ p^y \\ p^z \end{pmatrix} = \begin{pmatrix} \gamma & \gamma v & 0 & 0 \\ \gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} m \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} m\gamma \\ m\gamma v \\ 0 \\ 0 \end{pmatrix}.$$

This implies that the x velocity of the particle can be ‘‘extracted’’ from the components p^μ of its four-momentum by computing the ratio p^x/E. Since the x direction is arbitrary, the immediate generalization to a particle moving with velocity $\mathbf{v}$ in any direction is

$$\mathbf{v}=\frac{\mathbf{p}}{E}. \tag{7}$$

To obtain the law for the composition of two velocities v_1 and v_2 in the same direction, we may simply apply $B_x(-v_1)$ followed by $B_x(-v_2)$ to a particle at rest:

$$B_x(-v_2)B_x(-v_1)\begin{pmatrix} m\\ 0\\ 0\\ 0\end{pmatrix}=\begin{pmatrix} m\gamma_1\gamma_2(1+v_1v_2)\\ m\gamma_1\gamma_2(v_1+v_2)\\ 0\\ 0\end{pmatrix}. \quad (8)$$

(In this and subsequent equations we shall take it to be understood that the four components of the column vector refer to the components of the four-momentum.) On using (7), Eq. (8) yields

$$v_x=\frac{v_1+v_2}{1+v_1v_2}, \quad (9)$$

namely, the standard result. It will be noted that the same result would have been obtained if we had applied $B_x(-v_2)$ first and $B_x(-v_1)$ second: boosts in the same direction commute.

III. SOME ELEMENTARY THOMAS ROTATION ''PARADOXES''

Let us now apply the results obtained in Sec. II to some hypothetical maneuvers of the *USS Enterprise* under impulse power. In the following, the system of coordinates being considered is that of an observer on board the bridge of the *Enterprise*.

Let us assume that the *Enterprise* begins at rest relative to some particular star. We ignore gravitational effects, so that if the *Enterprise* were to not fire any thrusters, then it would remain at rest relative to the star.

Let us now apply a boost to the *Enterprise* by some velocity v_0 in the x direction, and follow it by a boost by the velocity $-v_0$, again in the x direction. We expect that the net effect on the velocity of the *Enterprise* would be zero: it would move in the x direction during the maneuver (by what distance is of no interest to us here), but at the end of the maneuver it would again be at rest relative to the star. We can confirm this by considering the effect of applying $B_x(v_0)$ and then $B_x(-v_0)$ on, say, the components of the four-momentum of the star, as observed from the *Enterprise*. If the star has the mass m, then a straightforward calculation verifies that

$$B_x(-v_0)B_x(v_0)\begin{pmatrix} m\\ 0\\ 0\\ 0\end{pmatrix}=\begin{pmatrix} m\\ 0\\ 0\\ 0\end{pmatrix},$$

and indeed $B_x(-v_0)B_x(v_0)=I$, where I is the identity matrix. [In performing these calculations, and those that follow, it is useful to replace even powers of v, wherever they occur, by means of the identity

$$v^2\equiv 1-\frac{1}{\gamma^2}, \quad (10)$$

which can be derived from the definition (2).] We can perform the same maneuver in the y direction: namely,

$$B_y(-v_0)B_y(v_0)\begin{pmatrix} m\\ 0\\ 0\\ 0\end{pmatrix}=\begin{pmatrix} m\\ 0\\ 0\\ 0\end{pmatrix},$$

and $B_y(-v_0)B_y(v_0)=I$, again as expected.

Having thus verified the action of our ''thrusters'' in the x and y directions, by means of these four boosts, let us now try another test maneuver, by mixing the order of these boosts. Namely, let us apply $B_x(v_0)$, followed by $B_y(v_0)$, then $B_x(-v_0)$, and finally $B_y(-v_0)$. Again, we expect that the star will be at rest, relative to the *Enterprise*, at the end of the maneuver. However, if we perform the calculations, then (after some algebra) we find

$$B_y(-v_0)B_x(-v_0)B_y(v_0)B_x(v_0)\begin{pmatrix} m\\ 0\\ 0\\ 0\end{pmatrix}$$
$$=\begin{pmatrix} m+m(\gamma_0+1)(\gamma_0-1)^3\\ m\gamma_0^2(\gamma_0-1)v_0\\ -m\gamma_0(\gamma_0-1)(-\gamma_0^2+\gamma_0+1)v_0\\ 0\end{pmatrix}. \quad (11)$$

Something has gone wrong! Instead of ending up with the star at rest, we find that it is now ''drifting.'' What has happened?

One can repeat and check the algebraic calculations above as many times and in as many ways as one wishes; but the result (11) is not a computational error. We can check its self-consistency by noting that, for any four-momentum of a particle of mass m, the identity $p^\mu p_\mu\equiv E^2-\mathbf{p}^2=m^2$ should be satisfied—as it indeed is for the components listed in (11). Moreover, if one simply changes the order of the final two boosts in (11), then one finds

$$B_x(-v_0)B_y(-v_0)B_y(v_0)B_x(v_0)\begin{pmatrix} m\\ 0\\ 0\\ 0\end{pmatrix}=\begin{pmatrix} m\\ 0\\ 0\\ 0\end{pmatrix},$$

which would be unlikely to be true had we made any trivial error in computing any of the boost matrices.

Let us therefore try to find out where our intuition has led us astray in the calculation (11), by breaking it down into smaller parts. We already know what happens to the components of the four-momentum of a particle, originally at rest, when we subject our system of coordinates to a single Lorentz boost, so let us consider instead the effect of the first *two* boosts in (11), namely, $B_x(v_0)$ followed by $B_y(v_0)$. If we stop our calculation at this point, we find

$$B_y(v_0)B_x(v_0)\begin{pmatrix} m\\ 0\\ 0\\ 0\end{pmatrix}=\begin{pmatrix} m\gamma_0^2\\ -m\gamma_0v_0\\ -m\gamma_0^2v_0\\ 0\end{pmatrix}. \quad (12)$$

Now, since we have boosted the *Enterprise* in the positive-x and positive-y directions, we expect that the star will be moving (relative to the *Enterprise*) with a negative velocity in the x and y directions; and this is borne out by the result (12). However, we are surprised to find that the x and y velocities *are not equal*, despite us boosting the *Enterprise* by the same velocity v_0 in each direction! Indeed, making use of relation (7) with the components (12) of p^μ, we find that the components of the three-velocity of the star, relative to the *Enterprise*, are given by

$$v_x = -\frac{v_0}{\gamma_0}, \quad v_y = -v_0. \tag{13}$$

Thus, the second (y) boost has been fully effective—but it has, in the process, reduced the velocity of the first (x) boost.

Let us put this unexpected asymmetry to one side, for the moment, and return to our first perplexing result, namely, the nonzero velocity represented by Eq. (11). We have found that the application of $B_x(v_0)$ and then $B_y(v_0)$ to the *Enterprise* leads to the star having the velocity components (13) (relative to the *Enterprise*). Let us now consider the final two boosts in Eq. (11), namely, $B_x(-v_0)$ followed by $B_y(-v_0)$. Instead of applying them after the first two boosts, let us instead apply them to the *original Enterprise*, which was at rest relative to the star. The effect of these two boosts on the components of the four-momentum of the star, in this modified scenario, would be

$$B_y(-v_0)B_x(-v_0)\begin{pmatrix} m \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} m\gamma_0^2 \\ m\gamma_0 v_0 \\ m\gamma_0^2 v_0 \\ 0 \end{pmatrix}, \tag{14}$$

leading to the velocity components

$$v_x = \frac{v_0}{\gamma_0}, \quad v_y = v_0. \tag{15}$$

Thus, comparing (13) and (15), we find that applying $B_x(-v_0)$ and then $B_y(-v_0)$ results in the exact opposite velocity to that obtained by applying $B_x(v_0)$ and then $B_y(v_0)$. (We would, of course, expect that this *would* be the case—but, given the problems we are having, it is essential to ensure that we do not make intuitive assumptions without testing them mathematically.)

We now find that our original result, Eq. (11), has not been clarified in the least. For we have shown that our sequence of four boosts can be broken down into a boost by the velocity components (13) (let us, for definiteness, refer to this three-velocity as $\mathbf{v}_{xy}$), followed by a boost by the velocity components (15) (namely, $-\mathbf{v}_{xy}$). But surely Einstein's very derivation of the Lorentz transformation guarantees us that a boost by any velocity $\mathbf{v}$, followed by a boost by $-\mathbf{v}$, must return us to the original inertial frame? How, then, can we make any sense of the result (11), which seems to imply that

$$B(-\mathbf{v}_{xy})B(\mathbf{v}_{xy}) \neq I? \tag{16}$$

Let us, again, put this problem to one side, and instead try the following tack: What if we were to perform four boosts, again in the $+x$, $+y$, $-x$, and $-y$ directions, respectively, but now adjusting the magnitude of each boost velocity so as to maintain some sort of control over the resultant overall velocity? Let us again start with a boost by velocity v_0 in the positive-x direction:

$$B_x(v_0)\begin{pmatrix} m \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} m\gamma_0 \\ -m\gamma_0 v_0 \\ 0 \\ 0 \end{pmatrix}.$$

We now apply a boost by some velocity v_1 (not equal to v_0) in the positive-y direction:

$$B_y(v_1)B_x(v_0)\begin{pmatrix} m \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} m\gamma_0\gamma_1 \\ -m\gamma_0 v_0 \\ -m\gamma_0\gamma_1 v_1 \\ 0 \end{pmatrix}.$$

Let us now adjust v_1 so that the overall velocity has equal components in the x and y directions (as was our original intention). We can do this by ensuring that p^x and p^y are equal—namely, by insisting that

$$\gamma_1 v_1 = v_0.$$

After some algebra, one finds that this is satisfied for

$$v_1 = \frac{\gamma_0 v_0}{\sqrt{2\gamma_0^2 - 1}}, \quad \gamma_1 = \frac{\sqrt{2\gamma_0^2 - 1}}{\gamma_0}, \tag{17}$$

so that

$$B_y(v_1)B_x(v_0)\begin{pmatrix} m \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} m\sqrt{2\gamma_0^2 - 1} \\ -m\gamma_0 v_0 \\ -m\gamma_0 v_0 \\ 0 \end{pmatrix}.$$

Let us now apply a boost by velocity v_2 in the negative-x direction:

$$B_x(-v_2)B_y(v_1)B_x(v_0)\begin{pmatrix} m \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} m\gamma_2(\sqrt{2\gamma_0^2 - 1} - \gamma_0 v_0 v_2) \\ -m\gamma_2(\gamma_0 v_0 - v_2\sqrt{2\gamma_0^2 - 1}) \\ -m\gamma_0 v_0 \\ 0 \end{pmatrix}.$$

We can use this third boost to reduce the x component of the velocity to zero by choosing

$$v_2 = v_1 = \frac{\gamma_0 v_0}{\sqrt{2\gamma_0^2 - 1}},$$

yielding

$$B_x(-v_1)B_y(v_1)B_x(v_0)\begin{pmatrix} m \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} m\gamma_0 \\ 0 \\ -m\gamma_0 v_0 \\ 0 \end{pmatrix}.$$

Finally, it is evident that we can apply the fourth boost by the original velocity v_0 in the negative-y direction, resulting in

$$B_y(-v_0)B_x(-v_1)B_y(v_1)B_x(v_0)\begin{pmatrix} m \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} m \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

We finally seem to have found a sequence of four boosts, in the $+x$, $+y$, $-x$, and $-y$ directions, respectively, that returns the *Enterprise* to a state of rest relative to the star at the end of the maneuver: namely, the sequence of boosts

$$B_y(-v_0)B_x(-v_1)B_y(v_1)B_x(v_0) \tag{18}$$

together with the relation (17) between v_1 and v_0.

Let us now go back in time to our original *Enterprise*, at rest relative to the nearby star. The crew of the *Enterprise* had noted that there was a shuttlecraft, of mass m_s, moving with velocity v_s in the positive-x direction; in other words, the components of its four-momentum, relative to the *Enterprise*, were

$$p^{\mu}=\begin{pmatrix} m_s\gamma_s \\ m_s\gamma_s v_s \\ 0 \\ 0 \end{pmatrix}. \tag{19}$$

What happens to the components of the four-momentum of the shuttlecraft after the sequence of boosts (18)? We would *expect* that they—like those of the star—would be unchanged. However, if we perform the calculations, we find that

$$B_y(-v_0)B_x(-v_1)B_y(v_1)B_x(v_0)\begin{pmatrix} m_s\gamma_s \\ m_s\gamma_s v_s \\ 0 \\ 0 \end{pmatrix}$$

$$=\begin{pmatrix} m_s\gamma_s \\ m_s\gamma_s v_s\sqrt{2\gamma_0^2-1}/\gamma_0^2 \\ m_s\gamma_s v_s(1-1/\gamma_0^2) \\ 0 \end{pmatrix}. \tag{20}$$

We can't seem to take a trick! Even though the sequence of boosts (18) has left the velocity of the star unchanged, relative to the *Enterprise*, it has *changed* the velocity of the shuttlecraft.

But how can this be possible?

IV. SOLUTIONS TO THE ELEMENTARY ''PARADOXES''

Let us now discover the fallacies contained in the ''paradoxes'' described above. We shall begin by unraveling our chain of arguments, starting with the final ''paradox,'' and working our way back to the first. By this stage, the reasons for each ''paradox'' will be clear. We shall complete this section by listing general expressions for the Thomas rotation in arbitrary cases.

We begin with the result (20) for the final four-momentum of the shuttlecraft. We were surprised to find that it differed from the four-momentum (19) of the shuttlecraft prior to the sequence of boosts. However, on closer inspection, we find that the result is not total chaos. In particular, the *energy* of the shuttlecraft has not changed. This, in turn, implies that its *speed* is also unchanged—in other words, the final velocity has the same magnitude as the original velocity, but *it has been rotated in space*. We can confirm this by using Pythagoras's theorem to find the resultant of the components v_x and v_y in (20); and we indeed find that its magnitude is simply v_s, the original speed of the shuttlecraft.

This rotation of the spatial axes is what we know as the *Thomas rotation*. It almost always occurs when we apply a sequence of noncollinear boosts that returns us to an inertial frame that is at rest relative to the original frame. This rotation had no net effect on the four-momentum of the star, because the star was at rest—the ''spatial'' components of its four-momentum vanished; in contrast, the motion of the shuttlecraft defines a direction in space (the x direction, in the original inertial frame), that was subject to the rotation.

We can obtain a clearer view of this rotation if we compute not the components of the four-momentum (20), but rather the entire matrix (18) that is applicable to *arbitrary* four-vectors in the original frame:

$$B_y(-v_0)B_x(-v_1)B_y(v_1)B_x(v_0)$$

$$=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \dfrac{\sqrt{2\gamma_0^2-1}}{\gamma_0^2} & -\dfrac{\gamma_0^2-1}{\gamma_0^2} & 0 \\ 0 & \dfrac{\gamma_0^2-1}{\gamma_0^2} & \dfrac{\sqrt{2\gamma_0^2-1}}{\gamma_0^2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{21}$$

The 2×2 matrix in the middle of this result is an orthogonal transformation, resulting in a rotation of the axes of the $x-y$ plane by an angle

$$\theta=-\arctan\left(\frac{\gamma_0^2-1}{\sqrt{2\gamma_0^2-1}}\right). \tag{22}$$

In the nonrelativistic limit, the magnitude of θ approaches v^2 rad (i.e., v^2/c^2 in conventional units), and so is completely negligible for terrestrial applications. (Even if v_0 is set to the Earth's orbital velocity around the Sun, the Thomas rotation angle amounts to a mere 0.004 arcsec.) In the ultrarelativistic limit, on the other hand, θ approaches $-90°$ for this particular sequence of boosts.

Defining the *direction* of the Thomas rotation, however, requires some care. Let us consider the above sequence of boosts from the point of view of an inertial observer, jettisoned from the *Enterprise* before the sequence of boosts commenced, who *remained* at rest relative to the star (and the distant ''fixed stars'') throughout the procedure. Relative to this fixed observer, the *Enterprise's* velocity rotated in the direction $+x\rightarrow +y$. The velocity of the shuttlecraft, *as seen by the Enterprise*, was rotated in this same direction. This means that, relative to the fixed observer, the axes of the *Enterprise*'s coordinate system rotated in the *opposite* direction to its orbital rotation, as indicated by the minus sign in Eq. (22). This is a general feature of the Thomas rotation.

Nonrelativistic physics has conditioned us to assume that Cartesian coordinate systems can be defined in space, in such a way that all inertial observers ''agree'' on the directions of the axes. The Thomas rotation demonstrates that this assumption requires an operational definition, as Einstein showed was necessary to clarify our understanding of the physics of relativity. For example, say that observer A defines a set of Cartesian axes. If observer B is at rest relative to A, then B can align her axes to ''agree'' with those of A. If observer A remains at rest, but observer B is boosted to some finite velocity relative to A, by one boost or by a sequence of boosts, *then the resultant orientation of* B's *axes depends on the particular sequence of boosts used*. If such boosts are at all times in the same direction (relative to A, say), then it is meaningful to say that B's axes are still aligned with A's, in the sense that if we apply any sequence of boosts to B that is at all times collinear with this direction, that returns B to rest with respect to A, then their axes will be found to still point in the same directions. On the other hand, if B is, at any two times, subject to boosts in *different* directions, then a sequence of boosts bringing B back to rest relative to A will, in general, lead to B finding her axes rotated relative to A's (unless the sequence of boosts ''backtracked'' precisely the original sequence).

To discuss the general case we need the expression for a (simple) boost by velocity $|\mathbf{v}|$ in the direction of $\mathbf{v}$. We can obtain it most simply by noting that our original boost transformation along the x axis, $B_x(v)$ of Eq. (4), is an archetypical simple boost. If we rewrite $B_x(v)$ in three-covariant notation (i.e., in terms of three-vectors and three-vector operations, rather than individual components), then we know from vector analysis that the result will be the boost we require for $\mathbf{v}$ in an arbitrary direction. It is now straightforward to verify that[5]

$$t' = \gamma t - \gamma(\mathbf{v}\cdot\mathbf{x}), \quad \mathbf{x}' = \mathbf{x} - \gamma t\mathbf{v} + \frac{\gamma^2}{\gamma+1}(\mathbf{v}\cdot\mathbf{x})\mathbf{v}, \tag{23}$$

is equivalent to (4) for $\mathbf{v}=(v,0,0)$; and thus (23) is the simple boost operation $B(\mathbf{v})$ that we are seeking. (It is a straightforward calculation to confirm that the component of $\mathbf{x}$ in the direction of $\mathbf{v}$ satisfies the usual Lorentz transformation:

$$\frac{\mathbf{v}\cdot\mathbf{x}'}{v} = \frac{\gamma\mathbf{v}\cdot\mathbf{x}}{v} - \gamma v t,$$

and that the components of $\mathbf{x}$ normal to the velocity $\mathbf{v}$ are unchanged:

$$\mathbf{v}\times\mathbf{x}' = \mathbf{v}\times\mathbf{x},$$

since the component of $\mathbf{x}$ or $\mathbf{x}'$ in the direction of $\mathbf{v}$ does not contribute to the cross product.) Written out in matrix form, we have

$$B(\mathbf{v}) = \begin{pmatrix} \gamma & -\gamma v_x & -\gamma v_y & -\gamma v_z \\ -\gamma v_x & 1+\dfrac{\gamma^2 v_x^2}{\gamma+1} & \dfrac{\gamma^2 v_x v_y}{\gamma+1} & \dfrac{\gamma^2 v_x v_z}{\gamma+1} \\ -\gamma v_y & \dfrac{\gamma^2 v_x v_y}{\gamma+1} & 1+\dfrac{\gamma^2 v_y^2}{\gamma+1} & \dfrac{\gamma^2 v_y v_z}{\gamma+1} \\ -\gamma v_z & \dfrac{\gamma^2 v_x v_z}{\gamma+1} & \dfrac{\gamma^2 v_y v_z}{\gamma+1} & 1+\dfrac{\gamma^2 v_z^2}{\gamma+1} \end{pmatrix}. \tag{24}$$

A sequence of two such boosts, which are in different directions, is not a simple boost, but is rather a combination of a rotation and a simple boost. If we consider $B(\mathbf{v}_1)$ followed by $B(\mathbf{v}_2)$, and denote the velocity of the boost implied by the composite transformation as $\mathbf{v}_{12}$, then the mathematical expression of this observation is that

$$B(\mathbf{v}_2)B(\mathbf{v}_1) = B(\mathbf{v}_{12})R(\mathbf{v}_1,\mathbf{v}_2),$$

where $R(\mathbf{v}_1,\mathbf{v}_2)$ is a spatial rotation, depending on the velocities $\mathbf{v}_1$ and $\mathbf{v}_2$.

Let us now return to the ''unexpected asymmetry'' in the result (13), namely, the fact that a boost by the velocity v_0 in the x direction, followed by a boost by v_0 in the y direction, leads to $v_x \neq v_y$ relative to the original frame of reference. We can understand the result $v_y = -v_0$ by the following argument: Imagine that, after the x boost, there is an object that is observed to be at rest. Applying the y boost to ourselves, we have no choice but to observe this object moving with velocity $-v_0$ in the y direction. This same argument must apply to any object in the original frame which had no velocity in the y direction.

The reduction in the x velocity by the y boost seems counterintuitive, but a little thought makes sense of it. We know, from Einstein's ingenious arguments, that lengths of rods perpendicular to a velocity vector are unchanged by the relative motion. But lengths are nothing more than differences in positions; and positions are themselves the spatial components of the four-vector x^μ. Thus, taking into account the universality of Lorentz covariance, Einstein's arguments imply that, for *any* four-vector, the spatial components perpendicular to the boost velocity are unchanged by the boost, as can be verified from Eq. (23). But the *spatial momentum* components are simply the spatial components of the four-vector p^μ; therefore, p^x and p^z must be unaltered by a y boost. And, indeed, the result (12) shows us that the x momentum of the star *was* unchanged by the second boost: it remained $-m\gamma_0 v_0$. Rather, the *energy* of the star increased (due to its new y velocity); and hence, by Eq. (7), its x velocity *decreased*. To have an object maintain its momentum, but lose velocity, is nonrelativistically counterintuitive; but one can make sense of it by remembering that all velocities must remain smaller than that of light, and so for a large enough boost in the y direction, any original velocity in the x direction must be ''quenched'' (although not its momentum!).

This asymmetry tells us that the noncommutativity of two noncollinear boosts is more complicated than is widely appreciated: One not only finds a relative Thomas rotation between the two resulting frames, but furthermore *the resulting frames are not even moving with the same velocity*. This is the source of the result expressed in Eq. (11). There we considered a sequence of four boosts which we naively expected to return us to our initial state of motion. However, the first two boosts do not combine to give a pure Lorentz boost, but rather involve a Thomas rotation. This rotation is not compensated by the later boosts—indeed, there is a *further* rotation in the same direction. Thus, we should not be surprised that the sequence of four boosts gives the counterintuitive result of Eq. (11).

This fundamental asymmetry is ''hidden'' in many introductory accounts of the addition of noncollinear velocities, by means of a judicious mixing of an *active* transformation for one velocity (i.e., the object is considered to be boosted, with we as observers being kept at rest) together with a *passive* transformation for the other (i.e., we as observers are being boosted), rather than two successive passive transformations as used in this paper. This ''trick'' gives the illusion of a greater degree of symmetry than is generally the case. (Einstein's seminal 1905 paper used this ''trick'').

All of these various points must be kept in mind if one wishes to analyze Thomas rotations in full generality. Any ''closed'' sequence of finite boosts (i.e., that returns us to a frame at rest relative to the original frame) will, in general, result in a Thomas rotation. Any such closed sequence may be broken down into a succession of closed sequences, each consisting of three boosts, in the same way that any arbitrary polygon (not necessarily planar) can be broken down into a ''triangular mesh'' by the addition of internal edges. Thus, the basic ''building block'' of a finite Thomas rotation is a sequence of three pure boosts: the first two are arbitrary, and the third must be chosen so as to make the sequence "closed." The first two velocities, then, determine the Thomas rotation for this "building block." Complicating such calculations, however, is the fact that the "sum" $\mathbf{v}_{12}$ of two arbitrary velocities $\mathbf{v}_1$ and $\mathbf{v}_2$ is, in the general case, quite a complicated function of the first two velocities:

$$\mathbf{v}_{12}=\frac{1}{\gamma_2(1+\mathbf{v}_1\cdot\mathbf{v}_2)}\left\{\mathbf{v}_1+\gamma_2\mathbf{v}_2+\frac{\gamma_2^2}{\gamma_2+1}(\mathbf{v}_1\cdot\mathbf{v}_2)\mathbf{v}_2\right\}. \quad (25)$$

[We can clearly see here the asymmetry between the two velocities $\mathbf{v}_1$ and $\mathbf{v}_2$; it is only if $\mathbf{v}_1$ and $\mathbf{v}_2$ are collinear that (25) becomes symmetrical under their interchange, and reproduces the usual formula (9) for the relativistic addition of velocities, as a short calculation shows.] The expression for the Thomas rotation is, in turn, even more complicated. Let us assume that we have an arbitrary three-vector $\mathbf{r}$ in our initial frame. After the sequence of pure boosts $B(\mathbf{v}_1)$, $B(\mathbf{v}_2)$, and $B(-\mathbf{v}_{12})$, the three-vector $\mathbf{r}$ is rotated to

$$\mathbf{r}'=\mathbf{r}+\frac{\gamma_1\gamma_2(\mathbf{v}\times\mathbf{v}_2)\times\mathbf{r}-\mathbf{Q}}{1+\gamma_1\gamma_2(1+\mathbf{v}_1\cdot\mathbf{v}_2)}, \quad (26a)$$

where

$$Q\equiv\frac{\gamma_1^2(\gamma_2^2-1)(\mathbf{v}_1\cdot\mathbf{r})\mathbf{v}_1+\gamma_2^2(\gamma_1^2-1)(\mathbf{v}_2\cdot\mathbf{r})\mathbf{v}_2-2\gamma_1^2\gamma_2^2(\mathbf{v}_1\cdot\mathbf{v}_2)(\mathbf{v}_1\cdot\mathbf{r})\mathbf{v}_2}{(\gamma_1+1)(\gamma_2+1)}. \quad (26b)$$

(Again, $\mathbf{Q}$ has no particular symmetry under the interchange of $\mathbf{v}_1$ and $\mathbf{v}_2$.) It can be verified, after some algebra, that $\mathbf{r}'^2=\mathbf{r}^2$, i.e., that $\mathbf{r}'$ is indeed simply a rotation of $\mathbf{r}$ in three-space. If $\mathbf{v}_2$ is small (but $\mathbf{v}_1$ arbitrarily large), then the expression $\mathbf{Q}$ of Eq. (26b) is of order $\mathbf{v}_2^2$, and hence is negligible in the context of Eq. (26a). If we are considering the continuous Thomas *precession*, then we can set $\mathbf{v}_1=\mathbf{v}$ and $\mathbf{v}_2=\delta\mathbf{v}_2$. Then, to quantities of first order, Eq. (25) yields

$$\delta\mathbf{v}\equiv\mathbf{v}_{12}-\mathbf{v}_1=\delta\mathbf{v}_2-(\mathbf{v}\cdot\delta\mathbf{v}_2)\mathbf{v}.$$

Thus, if $\delta\mathbf{v}_2$ is perpendicular to the velocity $\mathbf{v}$, then $\delta\mathbf{v}=\delta\mathbf{v}_2$; but if $\delta\mathbf{v}_2$ is parallel to $\mathbf{v}$, then one must take into account the fact that the velocity must remain smaller than that of light. On the other hand, in all cases we have $\mathbf{v}\times\delta\mathbf{v}=\mathbf{v}\times\delta\mathbf{v}_2$, so that Eq. (26a) yields, to first order,

$$\delta\mathbf{r}\equiv\mathbf{r}'-\mathbf{r}=\frac{\gamma}{\gamma+1}(\mathbf{v}\times\delta\mathbf{v})\times\mathbf{r}, \quad (27)$$

which is the standard expression for the Thomas precession.[5] [To compare with Jackson's result following his (11.117), note that our $\delta\mathbf{v}$ is his $\Delta\boldsymbol{\beta}$, and that our $\mathbf{v}\times\delta\mathbf{v}$ is, in his notation, $\boldsymbol{\beta}\times\Delta\boldsymbol{\beta}=\gamma\boldsymbol{\beta}\times\delta\boldsymbol{\beta}$.]

If we now consider the ultrarelativistic limit of Eq. (27), then we find something remarkable. This limit may be taken to be defined by the relations

$$\gamma\to\infty,\quad \mathbf{v}^2\to 1,\quad \mathbf{v}\cdot\delta\mathbf{v}\to 0, \quad (28)$$

the latter two of which simply reflect the fact that the velocity is at all times almost the speed of light, and (hence) that any changes $\delta\mathbf{v}$ to the velocity $\mathbf{v}$ must be perpendicular to $\mathbf{v}$. In this limit, Eq. (27) becomes

$$\delta\mathbf{r}\to(\mathbf{v}\times\delta\mathbf{v})\times\mathbf{r}. \quad (29)$$

Consider, now, the expression $(\mathbf{v}\times\delta\mathbf{v})\times\mathbf{v}$. By a standard three-vector identity, we have

$$(\mathbf{v}\times\delta\mathbf{v})\times\mathbf{v}\equiv\mathbf{v}^2\delta\mathbf{v}-(\mathbf{v}\cdot\delta\mathbf{v})\mathbf{v},$$

which, on account of the relations (28), tells us that, in the ultrarelativistic limit,

$$\delta\mathbf{v}\to(\mathbf{v}\times\delta\mathbf{v})\times\mathbf{v}. \quad (30)$$

Comparing (29) and (30), we thus find that $\mathbf{r}$ and $\mathbf{v}$ are being rotated by the same amount about the same axis. Recalling our discussion above that the direction of the Thomas rotation of the *axes* is opposite to the rotation of $\mathbf{r}$ *relative* to these axes, we therefore find that we have proved the following remarkable theorem: *For any ultrarelativistic object, the Thomas rotation is equal and opposite to the orbital rotation*.

This theorem explains why we obtained a rotation angle of 90° for our sequence of four boosts in the ultrarelativistic limit. For we can think of any *finite* boost as simply a sequence of infinitesimal boosts in the same direction. For our first $(+x)$ boost, we simply boosted the *Enterprise's* velocity to ultrarelativistic speeds. The second $(+y)$ boost was designed to bring the *Enterprise*'s velocity around to a 45° angle between the $+x$ and $+y$ directions; and the third $(-x)$ boost to bring it around another 45° to the $+y$ direction. The final $(-y)$ boost was antiparallel to this velocity, and simply brought the *Enterprise* back to rest. Thus, the velocity of the *Enterprise*, relative to a fixed observer, was rotated by 90° at ultrarelativistic speeds; and hence, by the above theorem, the Thomas rotation is just 90°, which is what we found by elementary means above.

V. AN ADVANCED "PARADOX": POLARIZATION PROPERTIES OF SCATTERING EVENTS

Let us now consider a more advanced situation: the calculation of a polarized cross section in quantum field theory. For simplicity, let us consider the scattering of a Dirac electron by the (idealized) fixed Coulomb field of an infinitely heavy, pointlike nucleus. For definiteness, we shall follow the notation and conventions employed in the introductory textbook by Mandl and Shaw.[11] In any frame for which the scattered electron momentum $\mathbf{p}'$ has the same magnitude as the incident momentum $\mathbf{p}$ (i.e., for which the electron's energy is unchanged by the scattering), the fully polarized cross section is given by[11]

$$\left(\frac{d\sigma}{d\Omega'}\right)_{rs}=\left(\frac{me}{2\pi}\right)^2|\mathcal{M}_{rs}|^2$$
$$=\left(\frac{me}{2\pi}\right)^2|\overline{u_s}(\mathbf{p}')A\!\!\!/_e(\mathbf{q})u_r(\mathbf{p})|^2, \quad (31)$$

where m is the mass and $-e$ the charge of the electron, $\mathcal{M}_{rs}$ is the Feynman amplitude for the process, $\mathbf{q} \equiv \mathbf{p}' - \mathbf{p}$ is the momentum transfer, $A_e(\mathbf{q})$ is the "external" electromagnetic field (i.e., the Coulomb field of the nucleus) in momentum space, and $\not{A}_e \equiv A_e^\mu \gamma_\mu$, where γ_μ are the Dirac gamma matrices [not to be confused with the factor γ defined in Eq. (2)]. The indices r and $s(=1,2)$ label the two possible spin states of the incident and scattered electron, respectively.

Let us first calculate all of the polarized cross sections for the following scenario. We choose an inertial frame in which the nucleus is at rest, so that[11]

$$A_e^\mu(\mathbf{x}) = \left(\frac{Ze}{4\pi|\mathbf{x}|}, 0,0,0 \right),$$

which under a Fourier transform yields

$$A_e^\mu(\mathbf{q}) = \left(\frac{Ze}{|\mathbf{q}|^2}, 0,0,0 \right),$$

where Ze is the charge of the nucleus. We then have[11]

$$\left(\frac{d\sigma}{d\Omega'} \right)_{rs} = \frac{(2m\alpha Z)^2}{|\mathbf{q}|^4} |\overline{u_s}(\mathbf{p}')\gamma^0 u_r(\mathbf{p})|^2, \tag{32}$$

where $\alpha \equiv e^2/4\pi$ is the fine-structure constant. Let us consider the case when the incident electron has velocity components

$$u_x = -\tfrac{2}{3}, \quad v_y = 0, \quad v_z = +\tfrac{2}{3},$$

and the scattered electron has velocity components

$$v_x = +\tfrac{2}{3} \quad v_y = 0, \quad v_z = +\tfrac{2}{3},$$

so that the electron is being scattered by 90° in the $z-x$ plane. From Eq. (2) we find that $\gamma = 3$, so that the incident and scattered four-momenta have the components

$$p^\mu = \begin{pmatrix} 3m \\ -2m \\ 0 \\ 2m \end{pmatrix}, \quad p'^\mu = \begin{pmatrix} 3m \\ 2m \\ 0 \\ 2m \end{pmatrix},$$

and $|\mathbf{q}| = 4m$. Now, the positive-energy spin–momentum eigenstates are, in the Dirac–Pauli representation of the Dirac matrices, given by[11]

$$u_1(\mathbf{p}) = c_1 \begin{pmatrix} 1 \\ 0 \\ c_2 p^z \\ c_2(p^x + ip^y) \end{pmatrix}, \quad u_2(\mathbf{p}) = c_1 \begin{pmatrix} 0 \\ 1 \\ c_2(p^x - ip^y) \\ -c_2 p^z \end{pmatrix}, \tag{33}$$

where

$$c_1 \equiv \sqrt{\frac{E+m}{2m}}, \quad c_2 \equiv \frac{1}{E+m}, \tag{34}$$

where $u_1(u_2)$ is the spin-up (spin-down) eigenstate relative to the z direction. The conjugate bispinor eigenstates, in this representation, are consequently

$$\overline{u_1}(\mathbf{p}) = c_1 \begin{pmatrix} 1 \\ 0 \\ -c_2 p^z \\ c_2(-p^x + ip^y) \end{pmatrix}^T,$$

$$\overline{u_2}(\mathbf{p}) = c_1 \begin{pmatrix} 0 \\ 1 \\ -c_2(p^x + ip^y) \\ c_2 p^z \end{pmatrix}^T. \tag{35}$$

For our particular case, $E = 3m$, so $c_1 = \sqrt{2}$ and $c_2 = 1/4m$. For the incident electron we therefore have

$$u_1(\mathbf{p}) = \frac{1}{\sqrt{2}} \begin{pmatrix} 2 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \quad u_2(\mathbf{p}) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 2 \\ -1 \\ -1 \end{pmatrix},$$

and for the scattered electron we have

$$\overline{u_1}(\mathbf{p}') = \frac{1}{\sqrt{2}} \begin{pmatrix} 2 \\ 0 \\ -1 \\ -1 \end{pmatrix}^T, \quad \overline{u_2}(\mathbf{p}') = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 2 \\ -1 \\ 1 \end{pmatrix}^T.$$

Let us now compute the cross-section (32). The quantity $\overline{u_s}(\mathbf{p}')\gamma^0 u_r(\mathbf{p})$ is equal to 2 for no spin flip in the z direction (i.e., for $\overline{u_1}$ with u_1, or for $\overline{u_2}$ with u_2), and is equal to ± 1 for spin flip in the z direction (i.e., for $\overline{u_1}$ with u_2, or for $\overline{u_2}$ with u_1). We thus find that

$$\frac{d\sigma}{d\Omega'}(\text{no spin flip}) = \frac{\alpha^2 Z^2}{16m^2}, \quad \frac{d\sigma}{d\Omega'}(\text{spin flip}) = \frac{\alpha^2 Z^2}{64m^2}. \tag{36}$$

To obtain the unpolarized cross section, we average over the initial spin states and sum over the final spin states in the standard way. This results in

$$\begin{aligned} \frac{d\sigma}{d\Omega'}(\text{unpolarized}) &\equiv \frac{1}{2} \sum_{r=1}^{2} \sum_{s=1}^{2} \left(\frac{d\sigma}{d\Omega'} \right)_{rs} \\ &= \frac{d\sigma}{d\Omega'}(\text{no spin flip}) + \frac{d\sigma}{d\Omega'}(\text{spin flip}) \\ &= \frac{5\alpha^2 Z^2}{64m^2}. \end{aligned}$$

We can compare this result with the standard Mott scattering formula,[11]

$$\frac{d\sigma}{d\Omega'}(\text{Mott}) = \frac{(\alpha Z)^2}{4E^2 v^4 \sin^4(\theta/2)} [1 - v^2 \sin^2(\theta/2)],$$

by noting that, for our case, $\theta = 90°$ so $\theta/2 = 45°$ and hence $\sin^2(\theta/2) = 1/2$; $v^2 = 8/9$; and $E = 3m$; which yields precisely the same result:

$$\frac{d\sigma}{d\Omega'}(\text{Mott}) = \frac{5\alpha^2 Z^2}{64m^2}.$$

We may therefore be confident that we have not made any elementary mistakes in calculating the polarized cross-sections (36).

Let us now compute these cross sections from the point of view of a different inertial frame. Specifically, let us view the process from an inertial frame which moves along the positive z axis with velocity 2/3 relative to the inertial frame used above. Applying $B_z(2/3)$ to the components of p^μ and p'^μ, we find

$$B_z(2/3)\begin{pmatrix} 3m \\ -2m \\ 0 \\ 2m \end{pmatrix} = \begin{pmatrix} m\sqrt{5} \\ -2m \\ 0 \\ 0 \end{pmatrix},$$

$$B_z(2/3)\begin{pmatrix} 3m \\ 2m \\ 0 \\ 2m \end{pmatrix} = \begin{pmatrix} m\sqrt{5} \\ 2m \\ 0 \\ 0 \end{pmatrix}, \tag{37}$$

so that, from the point of view of this new frame, the electron travels in the negative-x direction with energy $E = m\sqrt{5}$ and speed $2/\sqrt{5}$, and is then reflected elastically to travel in the positive-x direction with the same energy and speed. We also need to boost the components of the four-potential A_e^μ:

$$B_z(2/3)A_e^\mu(\mathbf{q}) = \frac{Ze}{\sqrt{5}|\mathbf{q}|^2}(3,0,0,-2),$$

so that the equivalent expression to (32) for the polarized cross section is

$$\left(\frac{d\sigma}{d\Omega'}\right)_{rs} = \frac{(2m\alpha Z)^2}{5|\mathbf{q}|^4}|3\overline{u_s}(\mathbf{p}')\gamma^0 u_r(\mathbf{p}) + 2\overline{u_s}(\mathbf{p}')\gamma^z u_r(\mathbf{p})|^2. \tag{38}$$

(We would, in general, need to transform the argument $\mathbf{q}$ as well as the components A_e^μ under a Lorentz transformation. However, if we define $q^\mu \equiv p'^\mu - p^\mu$, then in the original frame $|\mathbf{q}|^2 = -q^\mu q_\mu$ because $q^0 = 0$, i.e., the electron energy is conserved. Since $q^\mu q_\mu$ is a Lorentz scalar, then we find that $|\mathbf{q}|^2$ is invariant in any frame in which the electron energy is conserved—as is the case in the frame we have defined above.) Finally, from Eq. (34) we find that, using the boosted momentum values (37), the constants c_1 and c_2 are given by

$$c_1 = \sqrt{\frac{1+\sqrt{5}}{2}}, \quad c_2 = \frac{1}{m(1+\sqrt{5})},$$

so that for the incident electron we have

$$u_1(\mathbf{p}) = \frac{1}{\sqrt{2(1+\sqrt{5})}}\begin{pmatrix} 1+\sqrt{5} \\ 0 \\ 0 \\ -2 \end{pmatrix},$$

$$u_2(\mathbf{p}) = \frac{1}{\sqrt{2(1+\sqrt{5})}}\begin{pmatrix} 0 \\ 1+\sqrt{5} \\ -2 \\ 0 \end{pmatrix},$$

and for the scattered electron we have

$$\overline{u_1}(\mathbf{p}') = \frac{1}{\sqrt{2(1+\sqrt{5})}}\begin{pmatrix} 1+\sqrt{5} \\ 0 \\ 0 \\ -2 \end{pmatrix}^T,$$

$$u_2(\mathbf{p}') = \frac{1}{\sqrt{2(1+\sqrt{5})}}\begin{pmatrix} 0 \\ 1+\sqrt{5} \\ -2 \\ 0 \end{pmatrix}^T,$$

We now find that the quantity $\overline{u_s}(\mathbf{p}')\gamma^0 u_r(\mathbf{p})$ is unity for no spin flip in the z direction, but vanishes for spin flip. The quantity $\overline{u_s}(\mathbf{p}')\gamma^z u_r(\mathbf{p})$, on the other hand, vanishes for no spin flip, but has the value ± 2 for spin flip. Inserting these values into expression (38), we find that

$$\frac{d\sigma}{d\Omega'}(\text{no spin flip}) = \frac{9\alpha^2 Z^2}{320m^2},$$

$$\frac{d\sigma}{d\Omega'}(\text{spin flip}) = \frac{\alpha^2 Z^2}{20m^2}. \tag{39}$$

We've struck another disaster! The coefficients 9/320 and 1/20 in (39) look nothing at all like the values 1/16 and 1/64 that we found in (36). But we have merely performed the *same* calculation in two different inertial frames! How on Earth could the value of the cross section—which can be directly related to the number of particles that would be expected to be measured in an appropriately configured experiment—depend on an arbitrary choice of theoretical viewpoint? For example, if we prepare a beam of incident electrons so that they are completely polarized in the z direction, and filter the scattered electrons so that only those polarized in the z direction are detected, then what would the cross section be: $\alpha^2 Z^2/16m^2$ or $9\alpha^2 Z^2/320m^2$? There cannot be two different answers!

One might, at first glance, suspect that some trivial mistake or oversight has been made. However, the calculations above can be checked; they do not contain any arithmetical errors. Failing this, one might then suspect that we have not taken into account the transformation of the solid angle differential $d\Omega'$ under a Lorentz boost. However, if one checks the derivation[11] of the first of the relations (31), then one finds that it holds true in *any* elastic scattering of a single particle from an ''external'' field—essentially, the other kinematical factors happen to ''cancel out'' in this special class of scattering events.

There is, of course, a simple way to confirm or refute any suspicion one might have about the veracity of the results (39): One need simply combine them to find the *unpolarized* cross section. Surely, any trivial errors made in obtaining the results (39) would (in all but the most contrived of situations) render the unpolarized combination similarly erroneous. But we are now flabbergasted to find that

$$\tfrac{9}{320} + \tfrac{1}{20} = \tfrac{1}{16} + \tfrac{1}{64} = \tfrac{5}{64}.$$

Thus, even though we have obtained two sets of irreconcilably contradictory polarized cross sections, we find that their unpolarized combinations agree completely (and agree with the standard Mott formula)!

What is going on?

VI. SOLUTION TO THE POLARIZATION ''PARADOX''

Let us now use the general discussion of Sec. IV to understand the polarization ''paradox'' of the previous section. The key flaw in the arguments presented above is the description ‘‘polarized in the z direction.’’ *We have not specified whose z direction is being used!* The second calculation is simpler, in this regard, because the electron's final velocity is collinear with its initial velocity (i.e., it is in the same direction, but has the opposite sense). Thus, it is consistent for us to define ‘‘the z direction’’ to be our z axis, since all boosts to the electron's frames of reference are collinear. We can, say, prepare an electron polarized in the z direction, and measure only those scattered electrons polarized in the z direction, without ambiguity.

The first calculation, on the other hand, is more subtle. In using the standard expressions (33) and (35), we are (implicitly) applying one single Lorentz boost from our frame of reference to the initial electron's frame, and another single Lorentz boost from our frame to the scattered electron's frame. These two boosts, however, are not collinear; and so *our* description of events is different from how the *electron* would describe matters. (By giving the electron apparently human powers, we are of course imagining an observer traveling along with the electron.) In effect, the electron's very rest frame is *Thomas rotated* by the scattering event, relative to us. For example, imagine that the electron state does not get spin-flipped, as determined by the electron itself. From *our* point of view, however, the direction of polarization of the electron has changed!

The lesson of this example is clear. If one has need to calculate relativistic polarized cross sections explicitly, and if the incident and scattered momenta of the particles involved are not absolutely collinear (and in most practical experiments they are not), then one must be extremely cautious about how one defines the spins or polarizations of the particles involved. In particular, kinematical and semiclassical arguments must be examined in fine detail, to ensure that the nonrelativistic concept of universality of orientation has not been inappropriately applied.

Finally, we may use expressions (25) and (26) to reanalyze these polarized cross-section calculations *quantitatively*. If we set $\mathbf{v}_1$ to be the initial electron velocity, namely, $(-2/3,0,2/3)$, then it is straightforward to verify that a boost by $\mathbf{v}_2=(12/13,0,0)$ results in the correct final electron velocity of $\mathbf{v}_{12}=(2/3,0,2/3)$. If one sets $\mathbf{r}$ to be, say, (0, 0, 1), then, after some calculation, one finds that $\mathbf{r}'=(4/5,0,3/5)$. Thus, the electron's rest frame has been Thomas-rotated by an angle $\theta_T=\arctan(4/3)\approx 53°$ in the $z-x$ plane. If we now list the matrix elements corresponding to the polarized Feynman *amplitudes* found in Sec. V (rather than the cross sections), then for the first and second frames of reference we found, respectively,

$$\mathcal{M}_{rs}^{(1)}=\frac{2m\alpha Z}{|\mathbf{q}|^2}\begin{pmatrix}2 & -1\\ 1 & 2\end{pmatrix},\quad \mathcal{M}_{rs}^{(2)}=\frac{2m\alpha Z}{\sqrt{5}|\mathbf{q}|^2}\begin{pmatrix}3 & -4\\ 4 & 3\end{pmatrix},$$

where the rows in these matrices represent the z component of the initial spin, and the columns the z component of the final spin. To reconcile these Feynman amplitudes, we need simply apply the Thomas rotation or its inverse to either the initial or the final spin state in the first frame of reference. Remembering that spinors transform under rotations by half-angles, and noting that $\cos(\theta_T/2)=2/\sqrt{5}$ and $\sin(\theta_T/2)=1/\sqrt{5}$, we finally obtain

$$\begin{pmatrix}\cos\dfrac{\theta_T}{2} & -\sin\dfrac{\theta_T}{2}\\ \sin\dfrac{\theta_T}{2} & \cos\dfrac{\theta_T}{2}\end{pmatrix}\mathcal{M}_{rs}^{(1)}=\mathcal{M}_{rs}^{(2)}.$$

VII. CONCLUSIONS

We have shown how the Thomas rotation of relativistic mechanics can be introduced, and its ‘‘paradoxical’’ nature discussed, at quite an introductory level; that resolving such ‘‘paradoxes’’ is not overly difficult; and that a general expression for arbitrary Thomas rotations can be obtained without excessive effort. We have also shown how this general result connects up with standard textbook accounts of the infinitesimal Thomas precession. We have endeavored to show that the ramifications of such effects are deep, and fundamental, and that they may also be of immediate practical importance in the analysis and interpretation of relativistic polarized scattering experiments.

In the interests of keeping this discussion at an introductory level, we have refrained from using more advanced theoretical concepts to explain or analyze the Thomas rotation more elegantly or concisely. For example, group-theoretical methods are hinted at in the above derivations, but are not made explicit. (See, for example, Ref. 5 for a thorough treatment in these terms.) Boosts can be viewed as simply ‘‘rotations’’ between space and time; and since two rotations about different spatial axes do not, in general, commute, then one would (rightly) presume that two boosts in different directions do not commute either; this is another path to the Thomas rotation. Alternatively, one may make use of the concept of parallel transport—more familiar in the general theory of relativity, but equally applicable to boosts or accelerations in flat space–time—to arrive at the Thomas rotation by yet another path.[12] We believe that all of these more abstract views of the Thomas rotation do, in fact, augment, rather than detract from, the elementary nature and beauty of the effect as described here.

ACKNOWLEDGMENTS

This work was supported in part by the Australian Research Council. Helpful discussions with Brian J. Morphett are gratefully acknowledged. This paper is dedicated to the memory of A. J. Drinan.

[a]Current address: Mentone Grammar, 63 Venice Street, Mentone, Victoria 3194, Australia; electronic mail: jpc@physics.unimelb.edu.au; jpcostella@hotmail.com; www.ph.unimelb.edu.au/~jpc
[b]Electronic mail: mckellar@physics.unimelb.edu.au
[c]Electronic mail: arawlins@physics.unimelb.edu.au
[d]Electronic mail: gjs@baryon.phys.unm.edu; gjs@swcp.com

[1]J. P. Costella and B. H. J. McKellar, ‘‘The Foldy–Wouthuysen transformation,’’ Am. J. Phys. **63**, 1119–1121 (1995).
[2]J. P. Costella, B. H. J. McKellar, and A. A. Rawlinson, ‘‘Classical antiparticles,’’ Am. J. Phys. **65**, 835–841 (1997).
[3]A. Pais, *Subtle is the Lord* (Oxford U. P., Oxford, 1982), p. 143.
[4]A. A. Ungar, ‘‘Thomas precession and its associated grouplike structure,’’ Am. J. Phys. **59**, 824–834 (1991).
[5]J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1999), 3rd ed., Sec. 11.8.
[6]H. Goldstein, *Classical Mechanics* (Addison–Wesley, Reading, MA, 1980), 2nd ed., p. 285.

[7]P. K. MacKeown, ''Question #57. Thomas precession,'' Am. J. Phys. **65**, 105 (1997).
[8]G. H. Goedecke, ''Geometry of the Thomas precession,'' Am. J. Phys. **46**, 1055–1056 (1978).
[9]R. A. Muller, ''Thomas precession: Where is the torque?,'' Am. J. Phys. **60**, 313–317 (1992).
[10]R. J. Philpott, ''Thomas precession and the Liénard–Wiechert field,'' Am. J. Phys. **64**, 552–556 (1996).
[11]F. Mandl and G. Shaw, *Quantum Field Theory* (Wiley, Chichester, 1984), Secs. 8.7 and A.8.
[12]C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, New York, 1970).

CONNECTIONS

Most scientists glance at a mathematical formula and see awesome complexity, and marvel at the brain that first derived it. Like priests, some then seek to retain the mysteries. The awe is well aimed, but nevertheless not quite the correct response. There should be awe reserved for the original discoverer, for few have the power to discover new continents, or even islands, of knowledge. There should be delight, not awe, for the reconfirmation that the human brain is such a brilliant instrument that it can make light of darkness. But most important of all, there should be realization that a connection and a simplicity have been exposed. The connection is the formula, which bundles several knowns together, and shows that they account for another known. The simplicity is the reduction of the concepts that the new relation implies, although this is often interpreted as a complexity.

P. W. Atkins, ''The Limitless Power of Science,'' in *Nature's Imagination—The Frontiers of Scientific Vision*, edited by John Cornwell (Oxford University Press, New York, 1995).