

Proving the Relativistic Rotation Paradox

Christian P. H. Salas

Faculty of Mathematics & Statistics, The Open University, Milton Keynes
and Department of Mathematics, Waltham Forest College, London
e-mail: c.p.h.salas@open.ac.uk

Abstract

An apparent paradox in Einstein's Special Theory of Relativity, known as a Thomas precession rotation in atomic physics, has been verified experimentally in a number of ways. However, somewhat surprisingly, it has not yet been demonstrated algebraically in a straightforward manner using Lorentz-matrix-algebra. Authors in the past have resorted instead to computer verifications, or to overly-complicated derivations, leaving undergraduate students in particular with the impression that this is a mysterious and mathematically inaccessible phenomenon. This is surprising because, as shown in the present note, it is possible to use a basic property of orthogonal Lorentz matrices and a judicious choice for the configuration of the relevant inertial frames to give a very transparent algebraic proof. It is pedagogically useful for physics students particularly at undergraduate level to explore this. It not only clarifies the nature of the paradox at an accessible mathematical level and sheds additional light on some mathematical properties of Lorentz matrices and relatively-moving frames. It also illustrates the satisfaction that a clear mathematical understanding of a physics problem can bring, compared to uninspired computations or tortured derivations.

Keywords: *Special relativity, Rotation paradox, Thomas precession*

Mathematics Subject Classification: *83A02*

1 Introduction

Multiplying two Lorentz boosts whose velocity vectors are collinear gives a third Lorentz boost whose velocity can be calculated from the first two using the Einstein velocity addition law. If the two original velocities are not collinear, however, we do not get a pure Lorentz boost as the product, but

rather a Lorentz boost multiplied by a certain 4×4 matrix whose columns and rows are orthonormal. This orthogonal matrix has the effect of rotating the spatial components of vectors in spacetime, while leaving their temporal component unaffected.

An interesting review paper written by the eminent British mathematician I. J. Good [1] discusses this relativistic rotation paradox, but Good clearly struggles to provide a straightforward algebraic proof in the 3+1 case. Mathematically, it is necessary to show in a four-dimensional Minkowski spacetime that a certain matrix product involving two Lorentz boosts with linearly independent velocity vectors is generally equivalent to an orthogonal Lorentz matrix. In section 8 of [1], which seeks to prove the paradox ‘beyond any doubt’, Good admits to being unable to do this algebraically and instead resorts to providing numerical confirmations. Such computational confirmations are easy to carry out, so he suggests that ‘only a short and elegant algebraic proof would be worthwhile’. One of the motivations for the present note is that a proof like this still seems to be lacking in the literature.

What we have at present in the way of mathematical demonstrations are either elaborate approximations involving power series and extraneous assumptions such as infinitesimally small relative velocities (see, e.g., section 7.3 in [3]), or otherwise lengthy and overly-sophisticated expositions usually not easily accessible to, say, undergraduate physics students. Some expositions are intended to be more accessible but still seem rather involved and/or do not make use of Lorentz-matrix-algebra, e.g., [4] and [5].

Due to the lack of a short and transparent algebraic treatment, this interesting relativistic phenomenon is simply left unmentioned and unexplored in almost all undergraduate texts, which seems a pity. The following argument could be used shortly after introducing Lorentz transformation matrices and their properties to budding relativists.

2 Proof of the rotation paradox

Let $G = \text{diag}(1, -1, -1, -1)$ be the metric tensor in a four-dimensional Minkowski manifold with events specified by a time coordinate $x^0 = ct$ and rectangular spatial coordinates x^1, x^2, x^3 . A 4×4 Lorentz matrix Λ preserves the quadratic form $x^T G x$ in the sense that if $y = \Lambda x$ then $y^T G y = x^T G x$, so

$$\Lambda^T G \Lambda = G \tag{1}$$

The set of all Lorentz matrices thus defined constitutes a group under matrix multiplication, so inverses and products of Lorentz matrices are also Lorentz.

Let O, \bar{O} and $\overline{\bar{O}}$ be three inertial frames with collinear axes and with their origins initially coinciding. Let $\beta = (\beta_i) = \left(\frac{v_i}{c}\right)$, $i = 1, 2, 3$, be the 3×1 velocity

vector of \bar{O} relative to O with corresponding Lorentz factor $\gamma = \frac{1}{\sqrt{1-\beta^2}}$, where $\beta^2 \equiv \beta^T \beta$. Similarly, let a vector $\bar{\beta} = (\bar{\beta}_i)$, which is not collinear with β , be the velocity vector of $\bar{\bar{O}}$ relative to \bar{O} with corresponding Lorentz factor $\bar{\gamma}$. Using a standard formula, e.g., formula (24) in [1] or formula (2.59) in [2], the velocity vector of $\bar{\bar{O}}$ relative to O is given by

$$\bar{\bar{\beta}} = \frac{\bar{\beta} + \beta[\gamma + (\gamma - 1)(\bar{\beta}^T \beta)/\beta^2]}{(1 + \bar{\beta}^T \beta)\gamma} \quad (2)$$

with corresponding Lorentz factor $\bar{\bar{\gamma}}$. Using a simplification similar to one described in section 7.3 of [3], we can let the plane defined by the vectors β and $\bar{\beta}$ be the $\bar{x}^1 \bar{x}^2$ -plane of \bar{O} so that $\bar{\beta}_3 = 0$, and we can arrange the frames O and \bar{O} so that the vector β is along the x^1 axis of O , implying $\beta_2 = \beta_3 = 0$. We can do this for any given pair of velocity vectors which are not collinear, so there is no loss of generality here. Then (2) gives

$$\bar{\bar{\beta}}_1 = \frac{\bar{\beta}_1 + \beta_1}{1 + \bar{\beta}_1 \beta_1} \quad (3)$$

$$\bar{\bar{\beta}}_2 = \frac{\bar{\beta}_2}{(1 + \bar{\beta}_1 \beta_1)\gamma} \quad (4)$$

$$\bar{\bar{\beta}}_3 = 0 \quad (5)$$

and a standard formula, e.g., formula (7.11) in [3], allows us to write Lorentz transformations L and \bar{L} from O to \bar{O} and from \bar{O} to $\bar{\bar{O}}$ respectively as

$$L = \begin{pmatrix} \gamma & -\gamma\beta_1 & 0 & 0 \\ -\gamma\beta_1 & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (6)$$

and

$$\bar{L} = \begin{pmatrix} \bar{\gamma} & -\bar{\gamma}\bar{\beta}_1 & -\bar{\gamma}\bar{\beta}_2 & 0 \\ -\bar{\gamma}\bar{\beta}_1 & 1 + (\bar{\gamma} - 1)\frac{\bar{\beta}_1^2}{\bar{\beta}^2} & (\bar{\gamma} - 1)\frac{\bar{\beta}_1\bar{\beta}_2}{\bar{\beta}^2} & 0 \\ -\bar{\gamma}\bar{\beta}_2 & (\bar{\gamma} - 1)\frac{\bar{\beta}_1\bar{\beta}_2}{\bar{\beta}^2} & 1 + (\bar{\gamma} - 1)\frac{\bar{\beta}_2^2}{\bar{\beta}^2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (7)$$

A Lorentz boost $\bar{\bar{L}}$ from O to $\bar{\bar{O}}$ with velocity vector $\bar{\bar{\beta}}$ would be a matrix like (7), but with $\bar{\bar{\gamma}}$, $\bar{\bar{\beta}}_1$, $\bar{\bar{\beta}}_2$ and $\bar{\bar{\beta}}$ replacing $\bar{\gamma}$, $\bar{\beta}_1$, $\bar{\beta}_2$ and $\bar{\beta}$ respectively.

The relativistic rotation paradox is that, in general, $\bar{\bar{L}} \neq \bar{L} \times L$, but rather

$$\bar{\bar{L}} = R \times \bar{L} \times L \quad (8)$$

or equivalently

$$R = \overline{\overline{L}} \times L^{-1} \times \overline{L}^{-1} \quad (9)$$

where (9) is equation (28) in [1]. Numerical evidence in [1] suggests that

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & r_1 & s_1 & t_1 \\ 0 & r_2 & s_2 & t_2 \\ 0 & r_3 & s_3 & t_3 \end{pmatrix} \quad (10)$$

where the 3×3 submatrix in (10) is orthogonal. An approximation to (10) is also provided in equation (7.21) of [3] under the assumptions that the components of $\overline{\beta}$ are small and only need to be retained to first order, that $\overline{\gamma} \approx 1$, and that the distinction among γ , $\overline{\gamma}$ and $\overline{\overline{\gamma}}$ can be ignored to first order.

However, it is straightforward to obtain an exact algebraic proof that R in (9) is indeed an orthogonal matrix of the type given in (10) by observing that R must be Lorentz, since it is a product of Lorentz matrices. Therefore all that is required to prove the rotation paradox is to show that the 00-element of $\overline{\overline{L}} \times L^{-1} \times \overline{L}^{-1}$ is equal to 1, and that all the remaining elements in the first row are equal to zero, since any Lorentz matrix with a first row of this form must necessarily be an orthogonal matrix of the type given in (10). This assertion can easily be verified by substituting a generic 4×4 matrix with first row of the form $(1 \ 0 \ 0 \ 0)$ into the left-hand side of (1), setting the result equal to G on the right-hand side, and then comparing corresponding elements.

Note that L^{-1} and \overline{L}^{-1} are immediately obtained from (6) and (7) simply by removing the negative signs in the first row and first column. To prove that the 00-element of $\overline{\overline{L}} \times L^{-1} \times \overline{L}^{-1}$ equals 1, multiply the first row of $\overline{\overline{L}}$ by each of the columns of L^{-1} to get the 1×4 row vector

$$\left(\overline{\overline{\gamma}}\gamma(1 - \overline{\overline{\beta}}_1\beta_1) \quad \overline{\overline{\gamma}}\gamma(\beta_1 - \overline{\overline{\beta}}_1) \quad -\overline{\overline{\gamma}}\overline{\overline{\beta}}_2 \quad 0 \right) \quad (11)$$

and then multiply this row vector by the first column of the matrix \overline{L}^{-1} to get

$$\begin{aligned} & \gamma \overline{\overline{\gamma}} \overline{\overline{\gamma}}(1 - \overline{\overline{\beta}}_1\beta_1 + \overline{\beta}_1\beta_1 - \overline{\beta}_1\overline{\overline{\beta}}_1) - \overline{\overline{\gamma}} \overline{\overline{\gamma}} \overline{\overline{\beta}}_2\overline{\overline{\beta}}_2 \\ &= \frac{(1 + \beta_1\overline{\beta}_1)^2 - (\beta_1 + \overline{\beta}_1)^2 - \overline{\beta}_2^2(1 - \beta_1^2)}{\sqrt{1 - \beta_1^2}\sqrt{1 - \overline{\beta}_1^2 - \overline{\beta}_2^2}\sqrt{(1 + \beta_1\overline{\beta}_1)^2 - (\beta_1 + \overline{\beta}_1)^2 - \overline{\beta}_2^2(1 - \beta_1^2)}} = 1 \end{aligned}$$

as required. To prove that the 01-element of $\overline{\overline{L}} \times L^{-1} \times \overline{L}^{-1}$ equals 0, multiply the row vector in (11) by the second column of \overline{L}^{-1} to get

$$\gamma \overline{\overline{\gamma}} \overline{\overline{\gamma}}(\overline{\beta}_1 - \beta_1\overline{\beta}_1\overline{\overline{\beta}}_1) + \left[1 + (\overline{\overline{\gamma}} - 1)\frac{\overline{\beta}_1^2}{\overline{\beta}^2} \right] \gamma \overline{\overline{\gamma}} (\beta_1 - \overline{\beta}_1) - \overline{\overline{\gamma}} \overline{\overline{\beta}}_2 \left[(\overline{\overline{\gamma}} - 1)\frac{\overline{\beta}_1\overline{\beta}_2}{\overline{\beta}^2} \right]$$

$$= \frac{\bar{\beta}_1(1 - \beta_1^2)}{(1 - \beta_1^2)(1 - \bar{\beta}_1^2 - \bar{\beta}_2^2)} - \frac{\bar{\beta}_1(\bar{\beta}_1^2 + \bar{\beta}_2^2)}{(\bar{\beta}_1^2 + \bar{\beta}_2^2)(1 - \bar{\beta}_1^2 - \bar{\beta}_2^2)} = 0$$

as required. To prove that the 02-element of $\bar{L} \times L^{-1} \times \bar{L}^{-1}$ equals 0, multiply the row vector in (11) by the third column of \bar{L}^{-1} to get

$$\begin{aligned} & \gamma \bar{\gamma} \bar{\gamma} (\bar{\beta}_2 - \beta_1 \bar{\beta}_2 \bar{\beta}_1) + \left[(\bar{\gamma} - 1) \frac{\bar{\beta}_1 \bar{\beta}_2}{\bar{\beta}^2} \right] \gamma \bar{\gamma} (\beta_1 - \bar{\beta}_1) - \bar{\gamma} \bar{\beta}_2 \left[1 + (\bar{\gamma} - 1) \frac{\bar{\beta}_2^2}{\bar{\beta}^2} \right] \\ &= \frac{\bar{\beta}_2(1 - \beta_1^2)}{(1 - \beta_1^2)(1 - \bar{\beta}_1^2 - \bar{\beta}_2^2)} - \frac{\bar{\beta}_2(\bar{\beta}_1^2 + \bar{\beta}_2^2)}{(\bar{\beta}_1^2 + \bar{\beta}_2^2)(1 - \bar{\beta}_1^2 - \bar{\beta}_2^2)} = 0 \end{aligned}$$

as required. Finally, to prove that the 03-element of $\bar{L} \times L^{-1} \times \bar{L}^{-1}$ equals 0, multiply the row vector in (11) by the fourth column of \bar{L}^{-1} . This equals 0 by inspection, so the relativistic rotation paradox is proved.

References

- [1] Good, I. J., Lorentz Matrices: A Review, *Int. J. Theor. Phys.*, Vol. 34, pp. 779-799 (1995).
- [2] Mller, C., *The Theory of Relativity*, 2nd ed., Clarendon Press, Oxford (1972).
- [3] Goldstein, H., Poole, C., and Safko, J., *Classical Mechanics*, 3rd ed., Addison-Wesley, London (2002).
- [4] Ferraro, R., Thibeault, M., Generic composition of boosts: an elementary derivation of the Wigner rotation, *Eur. J. Phys.*, Vol. 20, p. 143 (1999).
- [5] O'Donnell, K., Visser, M., Elementary analysis of the special relativistic combination of velocities, Wigner rotation and Thomas precession, *Eur. J. Phys.*, Vol. 32, p. 143 (2011).