LINEAR ALGEBRA AS A PROOF OF SPECIAL RELATIVITY

NICOLE GORDON

Contents

1. Introduction

This paper discusses the application of linear transformations to special relativity. In particular, I will prove the Lorentz transformation is a linear transformation over Minkowski Space.

In 1905, Albert Einstein published "On the Electrodynamics of Moving Bodies." This paper introduced his special theory of relativity which solved a number of problems that had arisen in physics in the 19th century. The motivations for Einstein's special relativity came from the asymmetries in Maxwell's equations for bodies in motion and from the lack of proof that Earth moves through an "ether" [1].

Maxwell's equations describe the theory of electromagnetism. The equations hold in one reference frame, but if they are transformed to another reference frame using the Galilean transformations, they take a different form. The equivalent to this in mathematics is that each reference frame is a

Date: June 2, 2023.

vector space and the Galilean transformations between them are not linear transformations. For ref-

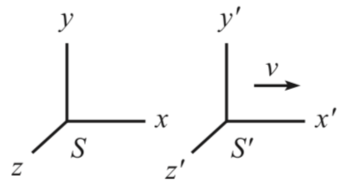

FIGURE 1. Reference frame S' moving with velocity v in the x direction with respect to frame S.

erence frames S and S' moving relative to each other in the x direction, the Galilean transformations are [2]

$$\Delta x = \Delta x' + v\Delta t' \tag{1}$$

$$\Delta t = \Delta t'$$

These transformations hold for slow speeds, but the asymmetries occur when bodies are moving near the speed of light. Contrary to this, it had already been proved that Maxwell's equations hold when switching between reference frames via the Lorentz transformations which take into account bodies moving near speed of light. Einstein's 1905 paper showed that the Lorentz transformation reduce to the Galilean transformations when the speeds involved are much less than the speed of light. Using linear algebra, it will be shown in this paper that the Lorentz transformations are in fact linear transformations.

Since light is an electromagnetic wave, it was known that light moves at about 3×10^8m/s. Like sound waves in air, it was proposed that light waves travel in a medium called the ether. In the 19th-century, to be a physicist meant to study the ether. The ether was believed to be luminiferous, all-pervasive, and unifying. In an 1887 experiment, Michelson and Morely attempted to measure the speed of the Earth relative to the ether. Michelson and Morely used an interferometer, which measures the interference between two light waves. The device was rotated to determine which angle produced the maximum interference due to the orientation of the ether. However, after many experiments performed over the course of a few months, no interference was found. Assuming the ether exists, this was a problematic result. This led to the second revolution that came from Einstein's paper, which was that the ether does not exist and light propagates without a medium.

2. Definitions

Definition 1. [3] A *linear transformation* between two vector spaces V, W is a function $f : V \to W$ such that:

A) $f(w + w) = f(v) + f(w) \quad \forall\ v, w \in V$
B) $f(\alpha v) = \alpha f(v) \quad \forall\ \alpha \in k, v \in V$

Definition 2. A function $f : V \to W$ is *bijective* if:

(1) it is injective:
$\forall\ x, y \in V \quad f(x) = f(y) \Rightarrow x = y$
(2) and it is surjective:
$\forall\ w \in W\ \exists\ v \in V$ such that $f(v) = w$

Definition 3. [3] A *linear isomorphism* $f : V \to W$ is a linear transformation which is bijective.

Definition 4. [3] The *dot product* on the vector space $\mathbb{R}^n$ over $\mathbb{R}$ is a function $\cdot : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ such that

$$(x_1, \ldots, x_n) \cdot (y_1, \ldots, y_n) = x_1 y_1 + \ldots x_n y_n$$

3. The Two Postulates

Motivated by these two large inconsistencies in physics, Einstein put forth the theory of special relativity. His theory included two postulates:

- **Postulate 1:** All inertial reference frames are considered equivalent.

  Known as the relativity postulate, this means there is no preferred reference frame as long as the frames are not accelerating. It also means that for every reference frame in which the laws of mechanics hold, the laws of electrodynamics also hold.
- **Postulate 2:** In empty space, light always propagates with speed c ($\sim 3 \times 10^8$ m/s) in any inertial frame.

  This means that the speed of light is independent of the emitting body. Unlike normal objects whose relative velocities add, the speed of light is constant, even when emitted from a moving body. This was suggested by James Clerk Maxwell in 1864.

4. Minkowski Space

Hermann Minkowski was a German mathematician who created a geometrical representation of space and time, called Minkowski Space, where space is represented on the x-axis and time is represented on the y-axis. The geometrical representation is known as the Minkowski Diagram, as seen in Figure 2. This idea did not become widely popular and accepted until Einstein showed

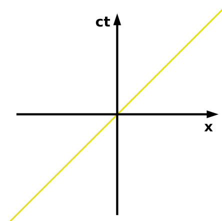

FIGURE 2. The path of a photon in Minkowski space. At $t = 0$ the photon is at point $x = 0$.

that this "spacetime" was a direct result of special relativity. In Minkowski Space the three spatial dimensions are x, y, and z and the time dimension is t. The time dimension is multiplied by the speed of light, c, so that all four dimensions have the same units. Thus, a vector in Minkowski Space is represented by

$$m = \begin{bmatrix} x \\ y \\ z \\ ct \end{bmatrix}$$

Minkowski Space is a four-dimensional vector space over $\mathbb{R}$.

5. A Moving Frame of Reference

Let there be a map $f : M \to M$ such that the speed of light remains and homogeneity of space is preserved. The homogeneity of space means that if the laws of physics hold in one point in space, they also hold in every other. From the postulate on the homogeneity of space, it follows that the intervals between any two events is invariant across reference frames [4]. It can be assumed from this that the "amount" of space-time does not change from one reference frame to another, even though a transformation may stretch or compress the space-time. If no space-time is lost when transforming from one vector space to another, that means the transformation produces no null-space. Thus, if the homogeneity of space is preserved, the transformation $f : M \to M$ is an isomorphism. The main goal of this paper is to prove that f is a linear transformation.

In physics, a frame of reference consists of a coordinate system that is fixed by reference points, allowing the positions and velocities of objects in the frame to be measured using that system. Let

S and S' be reference frames. S' is the reference frame S moving at speed v in the x direction. Let $S \subseteq M$ and $S' \subseteq M$ be subspaces. The reference point in S is the vector

$$O = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Now, I form an orthonormal basis for S. Let the vectors of the four dimensions – one time dimension and three spatial dimensions – be

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, e_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

where e_4 is the time dimension e_1, e_2, and e_3 are the x, y, and z dimensions, respectively.

Since S' is only moving in the x direction, if $s \in S$ and $s' \in S'$, then they take the form

$$s = \begin{bmatrix} x \\ y \\ z \\ ct \end{bmatrix}, \quad s' = \begin{bmatrix} x' \\ y \\ z \\ ct' \end{bmatrix}$$

We will use the following four axioms:

(1) The speed of light is the same in S and S'.
(2) $f : M \to M$ is an isomorphism.
(3) The y and z coordinates are the same in S and S'.
(4) The x and t coordinates are independent of the y and z coordinates.

6. The Matrix of the Lorentz Transformation

Let T be the matrix representation of f in the standard basis.

Theorem 5. *The matrix T is of the form:*

$$T = \begin{bmatrix} * & 0 & 0 & * \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ * & 0 & 0 & * \end{bmatrix}$$

Proof. In order to satisfy Axiom 3, it must be true that $T(e_2) = e_2$ and $T(e_3) = e_3$. Then by Axiom 4, $T(e_1)$, and $T(e_2)$ also have $x = 0$ and $t = 0$. By Axiom 3, the y and z coordinates are the same for any transformation T. Thus, the x, y, z, and t coordinates are the same for e_1 and $T(e_1)$, and for e_2 and $T(e_2)$. So, $T(e_2) = e_2$ and $T(e_3) = e_3$ as desired. The only way to preserve e_2 and e_3 is if the second and third columns of T are e_2 and e_3, respectively.

Next, I claim that $\text{Span}(e_1, e_4)$ is T-invariant. This means that for $w \in \text{Span}(e_1, e_4)$, $T(w)$ is also in $\text{Span}(e_1, e_4)$. If $w \in \text{Span}(e_1, e_4)$, it must take the form $w = (a, 0, 0, b)$ for some $a, b \in \mathbb{R}$ By Axiom 3, the y and z coordinates do not change under the transformation, so $T(w) = (c, 0, 0, d)$ for some $c, d \in \mathbb{R}$. This is also in $\text{Span}(e_1, e_4)$, so $\text{Span}(e_1, e_4)$ is T-invariant. The only way for this to be true are if the second and third entries in the first and fourth columns of T are all 0.

From these two claims, we get that T is of the desired form. In the following sections, we will determine what the remaining entries (*) in T are by adding further constraints. □

7. Inner Product Space

Consider an event that occurs at some point in S. The image of the event travels away at speed c in all directions. Thus, after time t, the event forms a sphere of radius ct. Any point on the sphere is described by the spatial coordinates (x, y, z). We get the equation

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

This can be re-arranged to give E, the set of all events on the sphere, where

$$E = \{x, y, z, t \mid x^2 + y^2 + z^2 - (ct)^2 = 0, t \geq 0\} \tag{3}$$

Now, for vectors w, u, let the dot product be

$$\langle w, u \rangle = \sum_{i=1}^{3} w_i u_i \tag{4}$$

Define the matrix A to be

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Let $w \in M$. Then

$$\langle A(w), w \rangle = \sum_{i=1}^{3} w_i w_i - w_4 w_4 \tag{5}$$

Equation 5 is called the ”interval” in special relativity.

If $w \in M$, it can be verified that $w \in E$ if $A(w)$ is orthogonal to w, i.e., the dot product $\langle A(w), w\rangle = 0$. If $\langle A(w), w\rangle = 0$, then

$$\sum_{i=1}^{3} w_i w_i - w_4 w_4 = 0 \tag{6}$$

Vectors in M are represented by (x, y, z, ct). So, using Equation 6 gives $x^2 + y^2 + z^2 - (ct)^2 = 0$, which is exactly a vector in E.

So, assume that $w \in E$. Let w be a vector in the reference frame, or vector space, S. If $w \in E$, then w lies on the sphere that is defined by the light traveling away from an event. This also means that in S,

$$\langle A(w), w\rangle = 0 \tag{7}$$

But by Axiom 1, the light traveling from an event must form a sphere in *any* reference frame, or any vector space. Thus, it does not matter whether the event is viewed from S or S'. Mathematically, this means that

$$\langle AT(w), T(w)\rangle = 0 \tag{8}$$

This is true because the vector $T(w)$ is the vector w transformed from the vector space S to the vector space S'. The relation in Equation 8 is the same relation as in Equation 7; the difference is that Equation 7 is for the vector $w \in S$, whereas Equation 8 is for the vector $T(w) \in S'$.

8. Determining the Form of T

We can now use the fact that $\langle AT(w), T(w)\rangle = 0$ to further specify the form of T.

Theorem 6. *Let $B = T^t AT$ where the exponent t denotes the transpose. Then $B = A$.*

Proof. Consider $w_1, w_2 \in M$ such that $w_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$ and $w_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$.

First, we want to show that

(1) $B(w_1) = \alpha w_2$
(2) $B(w_2) = \beta w_1$

for some $\alpha, \beta \neq 0$.

Proof of (1).

$\{w_1, w_2\}$ forms an orthogonal basis for $\mathrm{Span}(e_1, e_4)$. Also note that $w_1 \in E$. By Equation 8, and the properties of the dot product we get

$$\langle AT(w_1), T(w_1)\rangle = \langle T^t AT(w_1), w_1\rangle = 0 \tag{9}$$

The second dot product being equal to 0 means that $B(w_1)$ is orthogonal to w_1. From Theorem 5 and from the definition of A, $\mathrm{Span}(e_1, e_4)$ must also be B-invariant. So, $B(w_1) \in \mathrm{Span}(e_1, e_4)$. Since $B(w_1)$ is orthogonal to w_1, this must mean that $B(w_1)$ is a scalar multiple of w_2 such that $B(w_1) = \alpha w_2$. Since B is invertible, α must be non-zero, as desired.

Proof of (2).

$\{w_1, w_2\}$ forms an orthogonal basis for $\mathrm{Span}(e_1, e_4)$. Also note that $w_2 \in E$. By Equation 8, and the properties of the dot product we get

$$\langle AT(w_2), T(w_2)\rangle = \langle T^t AT(w_2), w_2\rangle = 0 \tag{10}$$

The second dot product being equal to 0 means that $B(w_2)$ is orthogonal to w_2. From Theorem 5 and from the definition of A, $\mathrm{Span}(e_1, e_4)$ must also be B-invariant. So, $B(w_2) \in \mathrm{Span}(e_1, e_4)$. Since $B(w_2)$ is orthogonal to w_2, this must mean that $B(w_2)$ is a scalar multiple of w_1 such that $B(w_2) = \beta w_1$. Since B is invertible, β must be non-zero, as desired.

Now, we construct a matrix for B. We can re-write e_1 and e_4 in terms of w_1 and w_2. Specifically, $e_1 = \frac{w_1+w_2}{2}$ and $e_4 = \frac{w_1-w_2}{2}$. Since B is linear, $B(e_1) = \frac{B(w_1)+B(w_2)}{2}$ and $B(e_4) = \frac{B(w_1)-B(w_2)}{2}$. We can replace $B(w_1)$ and $B(w_2)$ with the result found above. We than have $B(e_1) = \frac{\alpha w_2+\beta w_1}{2}$ and $B(e_4) = \frac{\alpha w_2-\beta w_1}{2}$ for some $\alpha, \beta \neq 0$. By Theorem 5 and the form of the matrix A, $B(e_2) = e_2$ and $B(e_3) = e_3$. Now, let $p = \frac{\alpha+\beta}{2}$ and $q = \frac{\alpha-\beta}{2}$. Thus,

$$B = \begin{bmatrix} p & 0 & 0 & q \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -q & 0 & 0 & -p \end{bmatrix}$$

It was assumed that $B = T^t AT$. Thus, it is also true that $B = T^t A^t T$ because A is symmetric. We know that the transpose of a product is the product of the transposes in the reverse order, so $B = (T^t AT)^t$. The product in the parentheses is B by assumption, so $B = B^t$. Thus, q must be 0. So, we have

$$B = \begin{bmatrix} p & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -p \end{bmatrix}$$

Now, consider $w_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \in E$. By Equation 8, $\langle AT(w_3), T(w_3)\rangle = 0$. By the properties of the dot product, $\langle AT(w_3), T(w_3)\rangle = \langle T^t AT(w_3), w_3\rangle$. By the definition of B, this equals $\langle B(w_3), w_3\rangle$. Using the form of B we know, $B(w_3) = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -p \end{bmatrix}$. Dotting this with w_3 gives $\langle B(w_3), w_3\rangle = 1 - p$. Since this must equal 0, we have that $p = 1$. Finally, we get that

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Thus, $B = A$, as desired. □

We can now use Theorem 6 to find the remaining entries in the matrix T. We do this by transforming an event in two equivalent ways. Consider an event in S' that happens at $s' = \begin{bmatrix} 0 \\ 0 \\ 0 \\ ct' \end{bmatrix}$ for some $t' > 0$. S' is moving at speed v away from S, so the coordinates in S are $s = \begin{bmatrix} vt \\ 0 \\ 0 \\ ct \end{bmatrix}$. The time is arbitrary, so let $t = \frac{1}{c}$. Thus, by re-scaling we may assume $s = \begin{bmatrix} \frac{v}{c} \\ 0 \\ 0 \\ 1 \end{bmatrix}$

Lemma 7. $T \begin{bmatrix} \frac{v}{c} \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \sqrt{1 - \left(\frac{v}{c}\right)^2} \end{bmatrix}$

Proof. From the dot product, we have

$$\langle T^t AT(s), s \rangle = \langle AT(s), T(s) \rangle$$

Since s' is s under the transformation T, $T(s) = s'$. So $\langle AT(s), T(s) \rangle = \langle A(s'), s' \rangle$ We defined A earlier, so by matrix multiplication and the dot product, $\langle A(s'), s' \rangle = -(ct')^2$ Thus,

$$\langle T^t AT(s), s \rangle = -(ct')^2 \tag{11}$$

By Theorem 6, we also have that $\langle T^t AT(s), s \rangle = \langle A(s), s \rangle$. Again, we already defined A, so by matrix multiplication $\langle A(s), s \rangle = \left(\frac{v}{c}\right)^2 - 1$. Thus,

$$\langle T^t AT(s), s \rangle = \left(\frac{v}{c}\right)^2 - 1 \tag{12}$$

Combining Equation 11 and Equation 12, we get that

$$ct' = \sqrt{1 - \left(\frac{v}{c}\right)^2} \tag{13}$$

□

Lemma 8. $T \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{v}{c\sqrt{1-\left(\frac{v}{c}\right)^2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}} \end{bmatrix}$

Proof. Let $a = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$ be in S. From the dot product, we have

$$\langle T^t AT(a), a \rangle = \langle AT(a), T(a) \rangle$$

. To get a in the S' frame of reference, we transform is using $T(a)$. S' is moving away from S at speed v, so the x coordinate in S' is $-vt''$ for some $t'' > 0$. The time coordinate, t'', is multiplied by c so that the units are consistent. So, $T(a) = \begin{bmatrix} -vt'' \\ 0 \\ 0 \\ ct'' \end{bmatrix}$. Let this vector be a'

So $\langle AT(a), T(a)\rangle = \langle A(a'), a'\rangle$ We defined A earlier, so by matrix multiplication and the dot product, $\langle A(s'), s'\rangle = v^2(t'')^2 - c^2(t'')^2$ Thus,

$$\langle T^t AT(a), a\rangle = v^2(t'')^2 - c^2(t'')^2 \tag{14}$$

By Theorem 6, we also have that $\langle T^t AT(a), a\rangle = \langle A(a), a\rangle$. Again, we already defined A, so by matrix multiplication $\langle A(s), s\rangle = -1$. Thus,

$$\langle T^t AT(a), a\rangle = -1 \tag{15}$$

Combining Equation 14 and Equation 15, we get that

$$t'' = \frac{1}{\sqrt{c^2 - v^2}} \tag{16}$$

□

We can now use Lemma 7 and Lemma 8 to express the matrix for T.

Theorem 9. $T = \begin{bmatrix} \frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}} & 0 & 0 & -\frac{v}{c\sqrt{1-\left(\frac{v}{c}\right)^2}} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{v}{c\sqrt{1-\left(\frac{v}{c}\right)^2}} & 0 & 0 & \frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}} \end{bmatrix}$

Proof. The second and third columns were established in Theorem 5. The fourth column comes from Lemma 8. So, we only need to show that $T(e_1) = e_1$ in order to complete the matrix T.

Case 1:

Let $v = 0$, meaning S' is stationary. If $v = 0$, the first and last entries in the fourth column of T become 0 and 1, respectively. We still do not know what the first and last entries of the first column of T are. We have that

$$T = \begin{bmatrix} j & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ k & 0 & 0 & 1 \end{bmatrix}$$

for some j, k.

If S' is stationary, the transformation from S to S' using T would yield the original vector. More precisely, $T \begin{bmatrix} x \\ y \\ z \\ ct \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \\ ct \end{bmatrix}$. This must mean that T is the identity matrix and that $j = 1$ and $k = 0$. So, for $v = 0$, we have that

$$T(e_1) = e_1 = \begin{bmatrix} \frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}} \\ 0 \\ 0 \\ -\frac{v}{c\sqrt{1-\left(\frac{v}{c}\right)^2}} \end{bmatrix}$$

Case 2:

Let $v > 0$. Then $T(e_1) = T\left(\left(\begin{bmatrix} v \\ 0 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right) \cdot \frac{c}{v}\right)$. T is linear, and using Lemma 7 and Lemma 8, we get $T(e_1) = \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ \sqrt{1-\left(\frac{v}{c}\right)^2} \end{bmatrix} - \begin{bmatrix} -\frac{v}{c\sqrt{1-\left(\frac{v}{c}\right)^2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}} \end{bmatrix}\right) \cdot \frac{c}{v}$. Simplifying this gives

$$T(e_1) = \begin{bmatrix} \frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}} \\ 0 \\ 0 \\ -\frac{v}{c\sqrt{1-\left(\frac{v}{c}\right)^2}} \end{bmatrix}$$

which is the same result as in Case 1. Thus, the theorem is proved. $\square$

9. Conclusion

In this paper I applied concepts in linear algebra to special relativity. Using linear transformations and the dot product, I proved that the Lorentz transformation is a linear transformation. The linear transformation is $f : M \to M$ where M is Minkowski Space. The transformation goes from a reference frame, S, to another reference frame, S', where S' is moving with speed v with respect to S. T is the matrix representation of f. The linear transformation from S to S' is given by

$$T = \begin{bmatrix} \frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}} & 0 & 0 & -\frac{v}{c\sqrt{1-\left(\frac{v}{c}\right)^2}} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{v}{c\sqrt{1-\left(\frac{v}{c}\right)^2}} & 0 & 0 & \frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}} \end{bmatrix}$$

where c is the speed of light.

References

[1] Albert Einstein. On the electrodynamics of moving bodies. 1923.

[2] David J Morin. *Special Relativity: For the Enthusiastic Beginner*. CreateSpace Independent Publishing Platform, 2017.

[3] Sheldon Axler. *Linear Algebra Done Right*. Springer, 2015.

[4] Lorentz transformation, Nov 2019.