The Linear Algebra Present in Lorentz Transformations

Andrew Valentini[1]

[1]Carthage College, Kenosha, Wisconsin

December 2022

Abstract:

In this paper, the Lorentz transformation is explained within the context of linear algebra. To lead to a strong understanding of a concept that is typically very non-intuitive, this paper first explains the basic history of Hendrik Lorentz and the historical context of physics at the time he began to rise through the field. The paper then covers linear transformations more generally because this is the primary tool used in the Lorentz transformation, as the name implies. After providing this mathematical background, the paper then presents the Lorentz transformation as a simple linear transformation. After getting a physical and mathematical intuition behind the Lorentz transformation, the paper closes with a specific example of its so the reader can see how the transformation is actually carried out. The Lorentz transformation is a difficult concept to understand so this paper attempts to reveal its beauty by expressing it in the terms of a linear transformation, a concept learned in an introductory course on linear algebra.

Keywords— Lorentz transformation, Special relativity, Linear transformation

Introduction

In this paper, the Lorentz transformation will be expressed as a linear transformation. Too often in physics classes are concepts such as these given to students without a firm understanding of the intuition behind the idea and the actual math that resulted in the concept. Because the Lorentz transformation is a difficult topic to understand, expressing it in terms of linear algebra will give readers a sense of why the formula for the transformation is given in the way that it is. To accomplish this goal, this paper will first detail the historical conditions that Hendrik Lorentz was born into and what this era meant for theoretical physics and the discoveries to come. Like all work in the field of physics, the Lorentz transformation is ultimately a response to older physics concepts which were beginning to result in contradictions and paradoxes as physicists began imagining more extreme scenarios (such as when an object's velocity approaches the speed of light or when an objects mass begins to get extremely high). In the context of the Lorentz transformation, the concept that it reevaluated was the Galilean transformation. Briefly understanding why the traditional Galilean transformation resulted in paradoxes will give the reader a solid understanding of why Lorentz came up with his own method of transformation.

After this and before the paper moves on to the actual form of the Lorentz transformation, it will first briefly detail the math necessary to understand where this transformation came from. A full course in linear algebra and special relativity would obviously lead to a more full understanding of how Lorentz came about his transformation but, for the scope of this project, only linear transformations will be briefly discussed since the Lorentz transformation can be expressed as one. The reasoning for this will also be expressed.

From a brief overview of what a linear transformation is, the Lorentz transformation will be built up. The Lorentz transformation can be expressed as a linear transformation of some vector space into another through multiplication by an A-matrix that will contain the actual factors of the Lorentz transformation. A full derivation of where the constants in the A-matrix come from is beyond the scope of this project so these values will be given to show how the linear transformation between two reference frames is formed. The derivation for the actual terms in the A-matrix will be provided in the references, however.

After an explanation of what the Lorentz transformation actually looks like with the background knowledge of linear transformations in mind, the paper will then walk the readers through an example of the Lorentz transformation being carried out so the ideas become more apparent.

1 History of Hendrik Lorentz and the Lorentz Transformation

Hendrik Lorentz was born in Arnhem, Netherlands on July 18th, 1853. In high school, Lorentz excelled in the physical sciences and mathematics as well as many classical languages. He was admitted to the University of Leiden in 1870 to study physics and math and graduated from it only a year later in 1971. After earning his bachelors degrees, Lorentz began teaching night classes for his previous high school while preparing for his doctoral thesis "On the Theory of Reflection and Refraction of Light" [3]. He earned his PhD in 1875, at 22 years old, and was awarded the newly-created chair of theoretical physics at the University of Leiden two years later in 1877.

During his first twenty years at the University of Leiden, Lorentz was interested in the electromagnetic theory of electricity, magnetism, and light, continuing in the field that he had done his PhD dissertation on. After this period, his focus shifted to the wider field of theoretical physics that was growing at the time as contradictions in the classical paradigm began to arise. Before the time of Lorentz, Galilean transformations were used to translate the coordinates of two reference frames, operating under the framework of Newtonian Mechanics. For example, if one observer O_1 witnesses another observer O_2, who is travelling at some velocity v_{O_2} perpendicular to the view of O_1, throw a ball at v_{ball} along the direction of motion, O_1 will see that ball travel at a total speed $v_{O_2} + v_{ball}$ while O_2 will only experience the ball travel with a speed v_{ball}. While this transformation is useful for "classical" cases (cases where objects have relatively low mass and are moving at relatively small velocities), these transformations encounter some contradictions if one begins thinking about objects moving at the speed of light. As we now know today, the speed of light is a universal constant ($c \approx 3 \times 10^8$). Before Lorentz, there had been experiments conducted that began suggesting this speed must be constant in any reference frame. James Clerk Maxwell even showed that electromagnetic waves must move at a constant speed in 1864 [4]. This fact cannot be reconciled with the classic Galilean transformations, for if a stationary observer sees a another person moving at half the speed of light and this person emits a beam of photons (which must move at the speed of light), the stationary observer will see the photons moving at $\frac{3}{2}c$ in the Galilean framework, which is not possible. Because of contradictions such as this were being realized around the time Lorentz was doing work, a new method of transforming between reference frames was required. This realization that new physics was required to begin describing "extreme" physics was what later allowed Albert Einstein to form his theories of special and general relativity (which rely on concepts such as the Lorentz Transformation).

While working in theoretical physics, Lorentz began thinking about the propagation of light as observed by various reference frames that all move relative to what was believed to be the ether of space. In doing this, he discovered that the transformation between reference frames in this context would simplified if a new time variable was introduced that depended on some universal time and the location that was being considered. Lorentz developed his own method of a transformation between reference frames, largely based on the work that had been done in 1887 by another physicist named Woldermar Voight [5]. Lorentz's new transformation relied on the fact that the speed of light must be a constant value in any reference frame and so developed a system that would not result in contradictions such as the one described above. Einstein would later use this and many of Lorentz's other concepts developed in theoretical physics in his theory of special relativity that was published in 1905, showing the great importance Lorentz's new transformation had on the realm of theoretical physics.

2 The Math Necessary for the Lorentz Transformation

As the name implies, the primary tool covered in an introductory course on linear algebra that is needed to understand the Lorentz transformation are linear transformations. A linear transformation acts as a function that converts one vector space to another, thereby preserving the essential qualifications of a vector space; vector addition and multiplication by a scalar. A linear transformation take one vector space, the *domain* V, and applies an operation across the entire space to result in a new vector space, called the *co-domain* W. If the domain V is said to contain the vectors $\langle v_0, v_1, ..., v_n \rangle$ and the codomain contains the vectors $\langle w_0, w_1, ..., w_n \rangle$, the linear transformation between those two spaces is expressed in the following way as described by [2]:

$$T : V \rightarrow W \text{ where } T(\langle v_0, v_1, ..., v_n \rangle) = \langle w_0, w_1, ..., w_n \rangle \quad (1)$$

If the co domain of a linear transformation does not fulfill vector addition and multiplication by a scalar, the transformation cannot be a linear one, for violating one of these properties for any specific case within the co-domain would mean that the transformation converted a linear system to a non-linear one (which is not allowed by its definition). The definition of a linear transformation can be used in this case because both observers of an event must be existing within the same dimensions of space (represented by the vector space $\mathbb{R}^4$, meaning that both are existing in spacetime). If the components of the domain and codomain are both vectors, a general matrix form can expressed as shown below:

$$T(\vec{v}) = A\vec{v} = \begin{bmatrix} a_1, 1 & a_1, 2 & ... & a_1, n \\ a_2, 1 & a_2, 2 & ... & a_2, n \\ \vdots & \vdots & \vdots & \vdots \\ a_m, 1 & a_m, 2 & ... & a_m, n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \quad (2)$$

Meaning that there exist some constants in the A-matrix that transform the vectors in the domain into the vectors of the codomain. This should already begin resembling what the Lorentz transformation attempts to do between frames of reference.

3 Linear Algebra Applied to the Lorentz Transformation

As explained in section 1, a Lorentz transformation compares the observation of a single event between two separate reference frames that does not allow for the inherent contradictions with the speed of light as was found to be the case in using Galilean transformations. As explained in equation 1, we can write this transform between the two observers as a multiplication of some A matrix by a vector $\vec{v}$, representing the reference frame of the first observer, and set this equal to another vector $\vec{w}$, representing the second observer's reference frame. The vectors $\vec{v}$ and $\vec{w}$ will be four-dimensional since spacetime is represented in terms of three spatial dimensions and one dimension of time. The intuition behind why the essential properties of vector spaces cannot be violated through a proper linear transformation is important to reflect on here. In the context of the Lorentz transformation, the vector spaces the transformation is occurring in is $\mathbb{R}^4$. If the transformation were to be carried out and a property of vector spaces were to be violated by the codomain, this would physically mean that the transformation took us from a reference frame in $\mathbb{R}^4$ into another that is not in this space. For the Lorentz transformation to work, the dimension that both observers are present in must be conserved across it. Because we are converting one 4 vector into another and we are using matrix multiplication (by the A-matrix) to complete this, the A-matrix must be a 4×4 matrix. This transformation is then written out as:

$$T(\vec{v}) = A\vec{v} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} \end{bmatrix} \begin{bmatrix} \vec{v_x} \\ \vec{v_y} \\ \vec{v_z} \\ \vec{v_t} \end{bmatrix} = \begin{bmatrix} \vec{w_x} \\ \vec{w_y} \\ \vec{w_z} \\ \vec{w_t} \end{bmatrix} \tag{3}$$

The form for the Lorentz transformation is then found by determining the terms for all of the values within the A-matrix. For simplicity, we can assume that the only dimension changing between the two observers is the x and time dimension (which will be represented at ct; or the distance travelled by light in one second). The full Lorentz transformation matrix can be seen in [1] but assuming that the y and z directions do not change with time is often a practical assumption to make. We can then remove the y and z terms from the A-matrix and the $\vec{v}$ and $\vec{w}$ vectors so that the general transformation is in the form:

$$T(\vec{v}) = A\vec{v} = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \begin{bmatrix} \vec{v_x} \\ \vec{v_t} \end{bmatrix} = \begin{bmatrix} \vec{w_x} \\ \vec{w_t} \end{bmatrix} \tag{4}$$

In order to solve for the four terms of the A-matrix, some assumptions about the speed of light being constant in any reference frame and how the origin of the system being described is expressed must be made. This is, again, beyond the scope of this project but the full derivation for these terms can be found in reference [6]. After solving for the terms within the A-matrix, the Lorentz transformation looks as follows:

$$\begin{bmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{bmatrix} \begin{pmatrix} x \\ ct \end{pmatrix} = \begin{pmatrix} x^{'} \\ ct^{'} \end{pmatrix} \tag{5}$$

These β and γ terms are constants that are dependent upon the system being described. They are expressed as:

$$\beta = \frac{v}{c} \ , \ \gamma = \frac{1}{\sqrt{1-(\frac{v}{c})^2}} \tag{6}$$

Meaning that the full form for the Lorentz transformation is written as:

$$\begin{bmatrix} \frac{1}{\sqrt{1-(\frac{v}{c})^2}} & -\frac{v}{c\sqrt{1-(\frac{v}{c})^2}} \\ -\frac{v}{c\sqrt{1-(\frac{v}{c})^2}} & \frac{1}{\sqrt{1-(\frac{v}{c})^2}} \end{bmatrix} \begin{pmatrix} x \\ ct \end{pmatrix} = \begin{pmatrix} x^{'} \\ ct^{'} \end{pmatrix} \tag{7}$$

Using matrix multiplication, we can write equation 5 out in terms of $x^{'}$ and $ct^{'}$ to get a better understanding of how the individual terms are expressed:

$$-\gamma\beta ct + \gamma x = x^{'} \implies x^{'} = \gamma(x - \beta) \tag{8}$$

$$\gamma ct - \gamma\beta x = ct^{'} \implies ct^{'} = \gamma(ct - \beta) \tag{9}$$

This can then be expressed in its full form given the forms for β and γ like so:

$$x^{'} = \frac{1}{\sqrt{1-(\frac{v}{c})^2}}(x - \frac{v}{c}) \tag{10}$$

$$ct^{'} = \frac{1}{\sqrt{1-(\frac{v}{c})^2}}(ct - \frac{v}{c}) \tag{11}$$

4 Example of and Solution to a Lorentz Transformation

As an example to illustrate the process of using the Lorentz transformation, consider an observer, denoted O_1, and a second observer, O_2, flying perpendicular to the direction of 0_1's path of observation (in the positive x-direction) in a spaceship that is moving at half the speed of light ($.5c$). The first observer turns a flashlight on in the direction of O_2's flight path at $t = 0$ and the photons released by this flashlight eventually reaches a target at some time and in some position of space. For 0_1, consider that they observe the photons hitting the target at the point $(x, ct) = (2, 2)$, meaning that they observe it hitting the target at two meters away in the x-direction and this occurs at two light meters (meaning the time it took for light to travel two meters). Because we are dealing with speeds of light, the traditional Galilean transformation will not work here so we must consider the Lorentz transformation to understand at what time and at what point in space this beam of photons will reach the target for O_2's frame of reference since we know these quantities for O_1.

We can numerically compute the Lorentz transformation to give us an exact sense for when in spacetime the event of the photons hitting the target will occur in O_2's frame of reference given that we know when it will occur for 0_1. In order to begin, we must first compute the γ and β values for this example:

$$\beta = \frac{v}{c} = .5$$

$$\gamma = \frac{1}{\sqrt{1 - (\frac{v}{c})^2)}} \approx 1.15$$

With these values, we can carry out the computation required for the Lorentz transformation as written in equations 8 and 9:

$$x^{'} = \gamma(x - \beta) \approx 1.73$$

$$ct^{'} = \gamma(ct - \beta) \approx 1.73$$

With these numbers, we can interpret how O_2 observes this event of photons reaching a target in their own frame of reference (which is moving at $.5c$ relative to O_1's frame of reference). With the factors $x^{'}, ct^{'} \approx 1.73$, calculated from the Lorentz transformation, we know that O_2 will observe the photons from the flashlight hitting the target 1.73 meters in the x-direction from the point declared as the origin (when the flashlight was first turned on). The second observer will also experience the photons hitting the target 1.73 light seconds after the flashlight is turned on. This situation can then be represented by the linear transformation:

$$\begin{bmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{bmatrix} \begin{pmatrix} x \\ ct \end{pmatrix} = \begin{pmatrix} x^{'} \\ ct^{'} \end{pmatrix} \implies \begin{bmatrix} 1.15 & -.575 \\ -.575 & 1.15 \end{bmatrix} \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 1.73 \\ 1.73 \end{pmatrix} \tag{12}$$

It's shown that O_1 and O_2 will both experience the event of the photons striking the target happening in different locations in space as well as in time, which is a result that could not stem from the Galilean transformation and is removed from our daily experience of how we move through space and time. The result of this example emphasizes the importance of the Lorentz transformation because it allows us to gain an understanding of how reality would work at extremely high speeds.

5 Conclusion

In this paper, the Lorentz transformation has been described through a linear transformation. The Lorentz transformation is a very important concept in theoretical physics and was a foundational piece for Einstein's theory of special relativity. The Lorentz transformation and is not an intuitive concept, considering that we have never experienced motion near the speed of light that would result in the paradoxes that Hendrick Lorentz used to build his own transformation between frames of reference. Because of this, it is important to understand the underlying math that describes the transformation because it helps one understand a phenomenon that will likely not be physically experienced by any person for the next few generations. The Lorentz transformation shows how theoretical physics uses mathematical concepts to explain physical phenomena that is often abstracted from our limited perspective of the universe.

References

[1] Keith Fratus. "Physics 103-The Lorentz Transformation". In: (2015), pp. 6–7. URL: https://web.physics.ucsb.edu/~fratus/phys103/LN/SR2.pdf.

[2] Ron Larson. *Elementary Linear Algebra*. Boston, MA: Cengage Learning, 2017.

[3] Hendrik Lorentz. "On the Theory of Reflection and Refraction of Light". PhD thesis. Leiden, Netherlands: Leiden University, June 1875.

[4] OpenStax. "Maxwell's Equations-Electromagentic Waves Predicted and Observed". 2022. URL: https://phys.libretexts.org/Bookshelves/College_Physics/Book%3A_College_Physics_1e_(OpenStax)/24%3A_Electromagnetic_Waves/24.01%3A_Maxwells_Equations-_Electromagnetic_Waves_Predicted_and_Observed.

[5] Woldemar Voigt. "On the Principle of Doppler". In: *Physikalische Zeitschrift* 16 (1887), pp. 381–385.

[6] Victor M. Yakovenko. "Derivation of the Lorentz Transformation". 2019. URL: http://www2.physics.umd.edu/~yakovenk/teaching/Lorentz.pdf.