Special Relativity and Linear Algebra

Corey Adams

May 7, 2010

1 Introduction

Before Einstein's publication in 1905 of his theory of special relativity, the mathematical manipulations that were a product of his theory were in fact already known. The so called Lorentz transformations were tricks that had been found that allowed the speed of light to propagate in all directions at the same speed, which accounted for its strange behavior when traveling through the now infamous "ether." Though Lorentz found the transformations well before him, it was not until Einstein decided to discard the ether and postulate the constancy of the speed of light as a physical law did the theory of relativity truly blossom.

For a complete treatment of relativistic space and time, Einstein's General Theory of Relativity is needed. His special theory makes one assumption that is shown to be false in the more general case: space is linear. In fact, many fascinating effects arise from the non-linearity of space, such as black holes and the curvature of the Universe. But, particularly in the absence of enormous masses, space can be closely approximated by linear means. This motivates a linear algebraic approach to special relativity. To proceed, I will put forward the postulates of special relativity, define the mathematical terminology necessary, motivated certain mathematical assumptions with physical arguments, and use the power of linear algebra to deduce the Lorentz transformations. This paper is aimed at an audience familiar with linear algebra.

2 Postulates of Special Relativity

There are a few basic assumptions involved with the theory of special relativity.

P1: All references frames are equivalent, or that no single reference frame is in any way special.

P2: The speed of light, measured in any reference frame and in any direction, is c.

P3: Equally spaced increments of space and time in one reference frame correspond to equally spaced increments of space and time in any other reference frame.

The first postulate is famously known as the principle of relativity. The second is a theoretical result of electrodynamics, put forward by James Clerk Maxwell in 1864 - before Einstein was even born. The final postulate is the homogeneity of space. It is this postulate that breaks down in the theory of general relativity. However, it holds to excellent approximation in the limit of weak or no gravitational fields.

3 Units and Representations

To work with the physics of relativity in a linear algebra format, it is necessary to define a vector space involving the space and time coordinates of our perceived reality. For physical reasons, each component of these vectors should have the same units. Given that in normal 3 dimensional space the units of any point are units of distance, it is logical to extend the time component to a distance unit by multiplying by the speed of light. That is, in the space-time vector space, the time like coordinate is ct, which does have the correct units of distance. Thus, a point in space-time with the space coordinates of x, y, z and time coordinate of t may be represented as:

$$\mathbf{w} = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

As such, the set of all $\mathbf{w}$ clearly constitutes a vector space $\mathbf{M}$ (commonly called Minkowski Space). What we are interested in is finding a linear transformation from $\mathbf{M}$ to itself that preserves the speed of light and homogeneity of space. This linear transformation will physically represent the conversion between two different reference frames, and mathematically will be an isomorphism from $\mathbf{M}$ to $\mathbf{M}$.

I must clarify what is meant by "reference frame." The intuitive concept is obvious; it was tacitly assumed in **P1** that the reader would understand what was meant. Mathematically, we define a reference frame to be a point $\mathbf{o}$ in space-time with the coordinates

$$\mathbf{o} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

and we orient our three space axes such that the unit vectors in those directions, along with the time component, are defined in the obvious way:

$$\mathbf{e_0} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \mathbf{e_1} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{e_2} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{e_3} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

The zeroth basis vector, $\mathbf{e_0}$, is considered time, and the other three coordinates are the traditional set. I will call such a reference frame S. We will denote any other reference frame moving with speed v with respect to S as S'. S' has all the same properties of S (as ensured by **P1**), and it is the conversion from S to S' that is the concern of this paper.

4 General Properties of the Transformation

I have already made the claim in passing that the transformation from one reference frame to another is linear, so let's prove it. Consider constant motion in the frame S. Then for any x_i in S, we have that $\frac{dx_i}{dt}$ is constant, and $\frac{d^2x_i}{dt^2} = 0$. What about the same motion in the frame S'? By the chain rule,

$$\begin{aligned}
\frac{dx'_i}{dt} &= \sum_j \frac{\partial x_i}{\partial j}\frac{dx_j}{dt} \\
\frac{d^2x_i}{dt^2} &= 0 = \sum_j \frac{d}{dt}\left(\frac{\partial x_i}{\partial j}\frac{dx_j}{dt}\right) \\
&= 0 = \sum_j \frac{\partial x_i}{\partial x_j}\frac{d^2x_j}{dt^2} + \sum_j\sum_k \frac{\partial^2 x_i}{\partial x_j \partial x_k}\frac{dx_j}{dt}\frac{dx_k}{dt}
\end{aligned}$$

Here summation over the full range of j and k is implied. In the last line, the first sum is zero because all $\frac{d^2x_j}{dt^2}$ are zero, and so the second sum must be zero also. For this to hold in general, because $\frac{dx_i}{dt}$ is any constant, all of the $\frac{\partial^2 x_i}{\partial x_j \partial x_k}$ must be zero. This is exactly the condition needed to prove that the transformation between references frames is linear.

Considering **P3**, that space-time is homogeneous, it is reasonable to assume that no matter what reference frame one is using, the "amount" of space-time is unchanged. When transforming from one frame to another, the transformation we seek will stretch or compress space-time, but will not ever be able to eliminate it.

That is, there is no null space of the transformation we seek. Any linear transformation from a finite dimensional vector space to itself whose null space is empty is an isomorphism. Denote this transformation as $T_v : \mathbf{M} \to \mathbf{M}$, where T is a function of v, the velocity between two reference frames.

Some simplifying assumptions make the math tidier, without losing generality. First, assume that the velocity vector is in the $\mathbf{e_1}$ direction, so that one reference frame is moving along the x axis of the other. Because we are not working with accelerations and the velocity between two frames never changes, this constraint does not affect the physical reality at all. What this assumption means explicitly is:

$$\text{If} \quad T_v \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} \quad \text{then } y = y' \text{and } z = z'.$$

Next, I assert that T_v is independent of the y and z components, or that only the x and t components matter. This is because the motion is only in the x direction, so the only effects noticed will be in the x and t components.

From the above assumptions, a few useful points can be proven. First, note that since $T_v(0) = 0$, and the transformation is independent of y and z components, the following must be true:

$$T_v \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

Similarly, $T_v(\mathbf{e_3}) = \mathbf{e_3}$. Clearly, this implies that the span of $\{\mathbf{e_1}, \mathbf{e_2}\}$ under T_v is invariant. So, the span of the other two basis vectors must also be T_v invariant. In a parallel fashion, it may also be shown that these properties all apply to the adjoint T_v^* of T_v.

5 Pseudo Metric and Inner Product Space

Consider an event happening at a point in the reference frame S. Intuitively, news travels out from this event at the speed of light in all directions. So, we can relate the distance of a point from an event to the time it takes for news of that event to reach the point with the equation of a sphere of radius ct: $(ct)^2 = (x^2+y^2+z^2)$. Rearranging, we see that $x^2 + y^2 + z^2 - (ct)^2 = 0$. This is very similar to the familiar notion of the dot product, except for that minus sign on the time component. In fact, if we define

$$\langle \mathbf{w}, \mathbf{u} \rangle = \sum_{i=0}^{3} w_i u_i \tag{1}$$

as the usual dot product, and the matrix η as

$$\eta = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

then we recover the equation of the sphere as

$$\langle \mathbf{w}, \eta(\mathbf{w}) \rangle = \langle \eta(\mathbf{w}), \mathbf{w} \rangle = \sum_{i=1}^{3} w_i w_i - w_0 w_0, \tag{2}$$

where $\mathbf{w}, \mathbf{w}$ are in $\mathbf{M}$. This modified version of a dot product is generally called the “interval” in relativity. Now, using **P2**, we see that light traveling out from an event forms a circle in any reference frame. That is, if $\langle \eta(\mathbf{w}), \mathbf{w} \rangle = 0$ then the relationship is true in any reference frame. So, we can operate on $\mathbf{w}$ with T_v to find $T_v(\mathbf{w}) = \mathbf{w'}$and it still must be true that

$$\langle \eta(\mathbf{w'}), \mathbf{w'} \rangle = 0$$
$$\langle \eta T_v(\mathbf{w}), T_v(\mathbf{w}) \rangle = 0.$$

So, using T^* as the usual notation for the adjoint of T,

$$\langle \eta(\mathbf{w'}), \mathbf{w'} \rangle = \langle \eta T_v(\mathbf{w}), T_v(\mathbf{w}) \rangle = \langle T_v^* \eta T_v(\mathbf{w}), \mathbf{w} \rangle.$$

Or more concisely, if $\langle \eta(\mathbf{w}), \mathbf{w} \rangle = 0$, then $\langle T_v^* \eta T_v(\mathbf{w}), \mathbf{w} \rangle = 0$.

6 Determining the form of T_v

We now have enough information to deduce some useful properties of the transformation from one reference frame to another. We'll consider two vectors in $\mathbf{M}$: [1]

$$w_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, w_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Notice immediately that these vectors are non zero under a normal inner product, but are zero when we use the modified inner product:

$$\langle \eta(\mathbf{w_1}), \mathbf{w_1} \rangle = \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = 0$$

$$\langle \eta(\mathbf{w_2}), \mathbf{w_2} \rangle = \begin{pmatrix} -1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = 0.$$

Note that $\mathbf{w_1}$ and $\mathbf{w_2}$ form an orthogonal basis for $\{\mathbf{e_1}, \mathbf{e_2}\}$. So, the span of $\{\mathbf{w_1}, \mathbf{w_2}\}$ must be both T_v and T_v^* invariant. Clearly they are also invariant under η, and so the following must be true:

$$T_v^* \eta T_v(Span\{\mathbf{w_1}, \mathbf{w_2}\}) = Span\{\mathbf{w_1}, \mathbf{w_2}\}$$

I showed earlier that if $\langle \eta(\mathbf{w}), \mathbf{w} \rangle = 0$, then $\langle T_v^* \eta T_v(\mathbf{w}), \mathbf{w} \rangle = 0$. Thus, since $\langle T_v^* \eta T_v(\mathbf{w_1}), \mathbf{w_1} \rangle = 0$, it follows that $T_v^* \eta T_v(\mathbf{w_1}) = a\mathbf{w_2}$, and similarly $T_v^* \eta T_v(\mathbf{w_2}) = b\mathbf{w_1}$. Hence:

$$T_v{}^* \eta T_v\left(\mathbf{w_1}\right) = T_v{}^* \eta T_v\left(\mathbf{e_0} + \mathbf{e_1}\right) = T_v{}^* \eta T_v\left(\mathbf{e_0}\right) + \ T_v{}^* \eta T_v\left(\mathbf{e_1}\right) = a\mathbf{w_2}$$
$$T_v{}^* \eta T_v\left(\mathbf{w_1}\right) = T_v{}^* \eta T_v\left(-\mathbf{e_0} + \mathbf{e_1}\right) = -T_v{}^* \eta T_v\left(\mathbf{e_0}\right) + \ T_v{}^* \eta T_v\left(\mathbf{e_1}\right) = b\mathbf{w_1}.$$

Adding and subtracting these, the effect of $T_v{}^* \eta T_v$ on the basis vectors is seen:

$$T_v{}^* \eta T_v\left(\mathbf{e_1}\right) = \frac{a\mathbf{w_2} + b\mathbf{w_1}}{2} = \begin{pmatrix} \frac{b-a}{2} \\ \frac{a+b}{2} \\ 0 \\ 0 \end{pmatrix}, T_v{}^* \eta T_v\left(\mathbf{e_2}\right) = \frac{a\mathbf{w_2} - b\mathbf{w_1}}{2} = \begin{pmatrix} \frac{-(a+b)}{2} \\ \frac{a-b}{2} \\ 0 \\ 0 \end{pmatrix}.$$

[1] As it turns out, these two vectors are eigenvectors of the transformation itself.

Obviously, $T_v{}^*\eta T_v(\mathbf{e_2}) = \mathbf{e_2}$, and similarly for $\mathbf{e_3}$. So, we can write the matrix from of $T_v{}^*\eta T_v$ with respect to the standard basis:

$$T_v{}^*\eta T_v = \begin{pmatrix} -\frac{a+b}{2} & -\frac{a-b}{2} & 0 & 0 \\ \frac{a-b}{2} & \frac{a+b}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -p & -q & 0 & 0 \\ q & p & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

with the substitution of $p = (a+b)/2$ and $q = (a-b)/2$ in the second matrix. Note that $(T_v{}^*\eta T_v)^* = T_v{}^*\eta T_v$, or that the above matrix must be self adjoint. So, q must be zero. Next, look at the following vector:

$$\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \text{ with } \langle T_v^*\eta T_v(\mathbf{v}), \mathbf{v}\rangle = 0.$$

So,

$$\langle T_v^*\eta T_v(\mathbf{v}), \mathbf{v}\rangle = 0 \text{ so that } -p+1 = 0.$$

Thus, $q = 0$, $p = 1$, and $T_v^*\eta T_v = \eta$. Now it is a simple matter of testing different vectors to infer the form of the transformation between two reference frames.

Imagine an event happening at the origin of reference frame S, so that its space time coordinates are

$$\mathbf{u} = \begin{pmatrix} ct \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

In the frame S', which is moving past S at speed v, the coordinates must be

$$\mathbf{u}' = \begin{pmatrix} ct' \\ -vt' \\ 0 \\ 0 \end{pmatrix}$$

So, $T_v(\mathbf{u}) = u'$, and by examining the "length" of these two events we see:

$$\langle \eta(\mathbf{u}), \mathbf{u}\rangle = \langle T_v{}^*\eta T_v(\mathbf{u}), \mathbf{u}\rangle = \langle \eta T_v(\mathbf{u}), T_v(\mathbf{u})\rangle = \langle \eta(\mathbf{u}'), \mathbf{u}'\rangle$$

From this, it is clear that $-(ct)^2 = -(ct')^2 + (-vt')^2$. Since the time t is arbitrary, let it equal[2] $1/c$ so that $1 = ct'\sqrt{1-(\frac{v}{c})^2}$, or $t' = \frac{1}{c}\frac{1}{\sqrt{1-(\frac{v}{c})^2}}$. Thus:

$$T_v \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1-(\frac{v}{c})^2}} \\ -\frac{v}{c\sqrt{1-(\frac{v}{c})^2}} \\ 0 \\ 0 \end{pmatrix}$$

Making the conventional substitutions of $\beta = \frac{v}{c}$, and $\gamma = 1/\sqrt{1-(\frac{v}{c})^2}$, this becomes tidier:

$$T_v \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma \\ -\beta\gamma \\ 0 \\ 0 \end{pmatrix}$$

[2]Note that this value of $1/c$ must actually be multiplied by 1 meter, so that it has units of time and not inverse velocity.

Now we consider a very similar situation, with a stationary event at the origin of S', which moves at a speed v so that

$$\mathbf{u} = \begin{pmatrix} ct \\ vt \\ 0 \\ 0 \end{pmatrix}, \ \mathbf{u'} = \begin{pmatrix} ct' \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Using similar methods from above, we find that

$$T_v \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -\beta\gamma \\ \gamma \\ 0 \\ 0 \end{pmatrix}.$$

Noting that the transformation does not affect the basis vectors $\mathbf{e_2}, \mathbf{e_3}$, we can finally write down the transformation matrix with respect to its standard basis in S, using the β and γ notation:

$$T_v = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

This famous matrix is known as the Lorentz transformation. It acts on the vectors in Minkowski space to transform from one reference frame to another.

7 The inverse transformation

Obviously, we are interested in knowing not just the transformation from S to S', but also from S' to S. Because T_v is an isomorphism from $\mathbf{M}$ to itself, we know that an inverse transformation must exist. Intuitively, we expect that the inverse will be the same in form but with v becoming $-v$. This causes no change in γ, since γ depends only on v^2, but β becomes $-\beta$.

Let's actually take the inverse of the Lorentz transformation. After some algebra,

$$(T_v)^{-1} = \begin{pmatrix} \frac{\gamma}{\gamma^2-\beta^2\gamma^2} & \frac{\beta\gamma}{\gamma^2-\beta^2\gamma^2} & 0 & 0 \\ \frac{\beta\gamma}{\gamma^2-\beta^2\gamma^2} & \frac{\gamma}{\gamma^2-\beta^2\gamma^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Notice that γ^2 can be factored out from the denominator, and that $1-\beta^2 = 1-(\frac{v}{c})^2 = (\sqrt{1-(\frac{v}{c})^2})^2 = \frac{1}{\gamma^2}$ This cancels the factor of γ^2 that was factored out, leaving just the following:

$$(T_v)^{-1} = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = T_{-v}.$$

This is exactly as we expected, with $(T_v)^{-1} = T_{-v}$.

8 Eigenvalues and Eigenvectors

Using the characteristic equation, $\det(T_v - \lambda I) = 0$, we can find the eigenvalues of the Lorentz transformation (ignoring the two eigenvalues of value 1 that correspond to y and z directions):

$$\det(T_v - \lambda I) = (\gamma - \lambda)^2 - \beta^2\gamma^2 = 0$$
$$\gamma - \lambda = \pm\beta\gamma$$
$$\lambda_\pm = \gamma(1 \pm \beta)$$
$$\lambda_\pm = \frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}}\left(1 \pm \frac{v}{c}\right) = \frac{1}{\sqrt{\left(1-\frac{v}{c}\right)\left(1+\frac{v}{c}\right)}}\left(1 \pm \frac{v}{c}\right) = \sqrt{\frac{1\pm\frac{v}{c}}{1\mp\frac{v}{c}}}$$

Clearly $\lambda_+ = (\lambda_-)^{-1}$. Note that, for $|x|$ less than one, $\tanh^{-1}(x) = \frac{1}{2}\log\frac{1+x}{1-x}$. So, since v/c is always less than one, we can use this definition to rewrite the eigenvalues as $\lambda_\pm = e^{\tanh^{-1}(\frac{v}{c})}$. We can define a quantity called rapidity to make notation easier here: $\phi = \tanh^{-1}(\frac{v}{c})$, so that $\lambda_\pm = e^{\pm\phi}$. As it turns out, rapidity has some interesting physical properties in its own right and is proportional to velocity at low speeds, but has not risen to prominence due to the fact that velocity itself is easier to measure. The eigenvectors (the two that aren't obvious) will be easy to see after applying the rapidity.

9 Rewrite the Lorentz transformation

In this section some interesting properties of rapidity are investigated, to show that the Lorentz transformation can be written as a type of rotation. First we'll prove some relations with $\gamma, \beta, \text{and}\phi$.

$$\tanh\phi = \frac{\sinh\phi}{\cosh\phi} = \frac{v}{c} \Rightarrow \sinh\phi^2 = (\tfrac{v}{c})^2\cosh\phi^2$$
$$\cosh\phi^2 - \sinh\phi^2 = 1 = \cosh\phi^2 - (\tfrac{v}{c})^2\cosh\phi^2 = \cosh\phi^2(1-(\tfrac{v}{c})^2)$$
$$\Rightarrow \cosh\phi = \frac{1}{\sqrt{1-(\frac{v}{c})^2}} = \gamma$$

From the definition of ϕ, we can easily deduce the other hyperbolic trigonometric function's dependance on γ and $\beta : \sinh\phi = \frac{v}{c}\gamma = \beta\gamma$. Thus, using the rapidity, the Lorentz transformation becomes:

$$T_v = \begin{pmatrix} \cosh\phi & -\sinh\phi & 0 & 0 \\ -\sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \text{ where } \tanh\phi = \frac{v}{c}$$

With this representation, two things are important. First, the eigenvectors of the Lorentz transformation are clearly $\mathbf{w_1}$ and $\mathbf{w_2}$ from before, and the eigenvalues are $\cosh\phi \pm \sinh\phi = e^{\pm\phi}$, as expected. Second, notice that the Lorentz transformation, in this form, looks very similar to a rotation matrix. Applying identities to convert hyperbolic trigonometric functions to normal trigonometric functions, we see that

$$T_v = \begin{pmatrix} \cos i\phi & i\sin i\phi & 0 & 0 \\ i\sin i\phi & \cos i\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Written as such, the Lorentz transformation seems like a rotation through the angle $i\phi$, in the coordinates x and ict. As a concept of space and time, imaginary time perhaps makes even less sense than equating time and distance, and rotating imaginary time and space truly exemplifies how special relativity bends our perception of reality.

10 An Example: Time Dilation

Much fame is given to the fact that special relativity disregards the concept of absolute time, and it is simple to show using the Lorentz transformation derived herein that this must be true. Consider two events occurring, and without loss of any generality place them both at the origin of reference frame S, such that the first happens at time $t = 0$ and the second at time $t = T$. Call these events A and B, respectively. In our vector notation, A and B must be:

$$A = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}_S , B = \begin{pmatrix} cT \\ 0 \\ 0 \\ 0 \end{pmatrix}_S$$

I have included subscripts on these vectors to remind us what reference frame these vectors refer to. It is a simple matter to find the coordinates of these events in S' - just apply the Lorentz transformation:

$$T_v(A) = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}_S = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}_{S'}$$

$$T_v(B) = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} cT \\ 0 \\ 0 \\ 0 \end{pmatrix}_S = \begin{pmatrix} \gamma cT \\ -\beta\gamma cT \\ 0 \\ 0 \end{pmatrix}_{S'}$$

The time elapsed between the events A and B in either reference frame is just the difference in the first coordinate of A and B, divided by the speed of light. It is just T in reference frame S. In reference frame S', it is equal to γT. That is, the time between events in S' is longer than that in S by a factor of γ. Since γ is always greater than or equal to one, time actually slows down in the moving frame. In a similar way, it can be shown that length decreases in all references frames except for one at which an object is at rest.

A physical explanation for this phenomenon is that the time it takes information to travel to an observer depends on how that observer is moving in relation to the event. In the frame S, the time it takes information to reach a stationary observer is the same for each event. However, in the frame S' information about event B takes longer to reach a stationary observer than it does for event A, because to that observer event B is farther away than event A. This physical argument in fact provides exactly the same factor of dilation, and the Lorentz transformation in general. It produces all the same results as the analysis used in this paper.

For all everyday events, $\frac{v}{c}$ is of the order of 10^{-7} at most, corresponding to a γ that does not deviate noticeably from unity. It is not until three orders of magnitude faster (over 20,000 miles per hour) that γ becomes at all noticeable: $\gamma \approx 1.00005$. This is why the non absoluteness of time is not relevent to our perception. However, at speeds approaching the speed of light, γ can grow enormously large - diverging completely at a velocity equal to the speed of light. This deviation from classical ideas of absoluteness plays a large role in many modern physical theories, and is of the utmost relevance in physics research.

11 Transformation of Velocities

A velocity of an object in a coordinate frame is defined as the time derivative of position with respect to time. As we saw above, *both* position and time are subject to change under the Lorentz transformation - this makes defining the velocity transformation slightly more difficult but possible. Given a very small displacement vector in S, we can compute its velocity in the usual way:

$$dA = \begin{pmatrix} cdt \\ dx \\ dy \\ dy \end{pmatrix}_S \Rightarrow u = \frac{dx}{dt}$$

We can transform dA and compute the ratio $\frac{dx'}{dt'}$ to find u':

$$T_v \begin{pmatrix} cdt \\ dx \\ dy \\ dz \end{pmatrix}_S = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} cdt \\ dx \\ dy \\ dz \end{pmatrix} = \begin{pmatrix} \gamma cdt - \gamma\beta dx \\ -\gamma c\beta dt + \gamma dx \\ dy \\ dz \end{pmatrix}_{S'}$$

This implies the association $dt' = \gamma dt - \frac{\gamma\beta dx}{c}$, and $dx' = -\gamma c\beta dt + \gamma dx$. Thus,

$$u' = \frac{dx - c\beta dt}{dt - \frac{\beta dx}{c}}$$

or, dividing above and below by dt and plugging in β:

$$u' = \frac{u-v}{1-\frac{vu}{c^2}}$$

This is the relativistic transformation of velocities. It allows us to verify that the theory is consistent: if an object travels at c in S, it must travel at c in S' by one of the postulates of special relativity. So:

$$c' = \frac{c-v}{1-\frac{vc}{c^2}} = \frac{c(c-v)}{c-v} = c$$

This confirms that the speed of light is in fact an invariant under the Lorentz transformation.

12 Conclusion

This paper has used the power of linear algebra, combined with a few basic physical principles, to derive the Lorentz transformation, representing the coordinate transformation between moving reference frames. A few physical postulates had to be made to proceed, and many physical concepts were translated well into mathematical concepts, such as the fact that the Lorentz transformation is an isomorphism. A brief study of the transformation and its eigenvalues led to the introduction of the concept of rapidity, which allowed the transformation to be seen as very similar in form to a rotation of coordinates. An example of relativistic effects was seen in time dilation; the matrix form of the Lorentz transformation made the calculation very easy to perform, as opposed to the somewhat tricky physical arguments often given.

The Lorentz transformation allows very easy conversion of space time coordinates from one reference frame moving at constant velocity to another. As such, the relativity of things such as time and length are simple to see. The conversion of velocities from one frame to another also follows from the Lorentz transformation, with the strange result that velocities do not necessarily add, and the speed of light is constant in all reference frames.

13 Acknowledgements

Thank you to Professor Doug Haessig for supervising this project. Much of the derivation follows a similar path as one found in [1], and the discussion is inspired in part by both [3] and [2].

References

[1] Spence Friedberg, Insel. *Linear Algebra, 4/E.* Pearson, 2003.

[2] Wolfgang Rindler. *Introduction to Special Relativity.* Oxford University Press, 1982.

[3] Bernard Schutz. *A First Course in General Relativity.* Cambridge University Press, 2009.