Space and Time
Version A: Full Math

W. Blaine Dowler

April 21, 2012

Contents

1 Space and Time Before Relativity

Prior to Einstein's work, the importance of both space and time to physics was clearly recognized. They were, however, treated as completely distinct quantities.

On one hand, they had the spatial dimensions. These came in three fundamental directions. The scientist had the ability to choose these directions to some degree. On the surface of the Earth, one could choose North as one direction, East as another, and Up as the third. If one chooses North then one does *not* choose South; mathematically, they describe the same line, since you can "undo" all travel in the direction North by traveling in the direction South. Directions are only "different" if they cannot be completely canceled by motion in the previously defined directions. One could choose Northeast as a direction, since the "East" portion can't be canceled by North, but choosing directions that aren't at right angles to each other usually makes things more complicated than they need to be when working on the math. The directions are connected through rotations: if you rotate through a right angle, the person who was facing North can now be facing East or Up. It is easy to see how they relate to one another, and that close relationship is one of the reasons we have the ability and inclination to define whatever directions are convenient for us. (For example, when working on a ramp, it is more convenient to pick directions parallel and perpendicular to the surface of the ramp than simply North, East and Up.)

On the other hand, there was (and is) time. No clear connection between space and time was apparent. Sure, they could plot graphs using time as one axis and space on another, but that was a technique used to see how systems evolve, and was not considered an indication of any deeper, fundamental connection between the quantities. Time was absolute and the same (invariant) for all observers, while distances were different for observers in motion (though not as different as they are under Einstein; the distance a car traveled was variable under Newtonian mechanics, but the length of the car was not, since the front and back could be measured "simultaneously.")

2 Space and Time After Relativity

With Einstein's relativity, we learn that time is not invariant, and that perceived time intervals depend greatly upon the motion of the observer. Moreover, the manner in which one measures distances depends on time measurements and vice versa. This implies a much stronger connection between the two quantities than was previously believed. This begs a number of questions: if time is that closely linked to space, can it be treated as another direction? If so, why does that direction behave so much differently than space? Why can we not navigate back and forth through time as easily as we do in space? Why can't we simply rotate from a spatial direction to a time direction as easily as we rotate from North to East? These questions opened some very productive floodgates. Physicists were forced to examine their beliefs about geometry, which led to some startling revelations when they explored the abstract discoveries made by mathematicians in the recent centuries.

Mathematicians can always be depended upon to indulge their curiosity, regardless of whether or not their discoveries apply to the "real world." One such avenue was the exploration of the square roots of negative numbers. For centuries, it was believed that negative numbers could have no square roots. In fact, they can, although it was difficult for even mathematicians to accept that fact when they were first proposed. This resistance is what led to the square roots of negative numbers being named "imaginary numbers:" even the mathematicians at the time had a hard time accepting them.[1] Despite their name, imaginary numbers are just as valid as any other numbers. Moreover, mathematicians had found that one can include directions describing imaginary numbers in geometry and arrive at precisely the type of system physicists needed. If *time* were measured along an imaginary direction, then it would be a direction tightly linked to the spatial directions, and yet would still be distinct and unique in precisely the manner required to satisfy intuition. No simple[2] rotation can point someone in an imaginary direction, yet the direction is an integral part of the geometry of reality. This distinction is enough to explain why we cannot control our motion along the time direction: it is a fundamentally different type of direction. Still, this would have been incredibly difficult to accept if not for the discovery of invariants.

3 Invariants

The realization that this mathematical framework described the time direction went hand in hand with another accidental discovery that had yet to be explained, related to the lengths of certain mathematical objects.

In Newtonian terms, moving observers disagreed on the positions of objects, but they always agreed on the lengths of those objects. Relativity threw that convenience out the window, but did so in a way that pulled time in as another direction to be considered in the geometry. Physicists began working with var-

[1] Most high school students have heard of the real numbers. They were named to contrast the imaginary numbers in an act of professional mockery. "You can study those imaginary numbers if you'd like, but *I* work with *real* numbers.

[2] This word is surprisingly important, as we'll see in section 4 on the following page.

ious ways to combine the space coordinates with the time coordinate to see if this idea of an "invariant" length could be preserved. They had discovered a means to make this happen that was remarkably counterintuitive.

If an object is turned to lie along one of the directions chosen for our coordinate system, then it is relatively easy to measure its length: find the coordinates of the two ends, and subtract the smaller number from the bigger number. Since the days of Pythagoras, there was a known method for measuring the length of an object that didn't lie along one of these directions: measure the coordinates of the extreme ends of the object along each of the chosen directions, square them, add the squares, and then take the square root of the answer. The calculation may be somewhat cumbersome without a calculator, but it is useful to know how to do this when one cannot simply turn the object or the measuring device to measure something in the simple, intuitive manner.

In trying to do this with relativity, physicists noticed something very strange. Adding in the square of the time coordinate didn't resolve the issue: both time and distance were smaller numbers in a moving reference frame. The strange and startling discovery was that these reductions could cancel each other out if subtracted! In other words, if one added the squares of the lengths of the object along the space dimensions, but then *subtracted* the square of the length along the time direction, the answer one got was the same for *all* observers. Quantities that are the same for all observers are known as *invariant* quantities.

This idea was shocking and counterintuitive, but completely consistent with treating time as an imaginary number. Recall from the previous section that imaginary numbers were discovered as the square roots of negative numbers: if one squares an imaginary number, the answer is negative. Adding a negative number is mathematically identical to subtracting a positive number. The implications of relativity get crazier and crazier, and yet somehow remain entirely consistent with each other and with experimental results. Slowly but surely, Einstein's seemingly insane notions were gaining more and more support.

4 Minkowski

Hermann Minkowski laid much of the foundation for this combination of space and time when he addressed the attendees of the 80th Assembly of German Natural Scientists and Physicians[3] on September 21, 1908.[4] His speech began

[3]No, the term "physicist" was not in use yet. Should you travel back in time by more than a century and develop some sort of medical ailment, be sure you know the true expertise of any physician you encounter.

[4]Minkowski was 44 years old at the time of his address, and would die of appendicitis in January 1909.

as follows (after translation):

> The views of space and time which I wish to lay before you have sprung from the soil of experimental physics, and therein lies their strength. They are radical. Henceforth space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality.

Minkowski's address then continued to lay the foundations for the work done earlier on in this lesson, though it should be noted that the usefulness of imaginary numbers had yet to be realized; Minkowski subtracted the square of time, not because he treated time as imaginary, but because it was already known that this was an invariant quantity. Still, he laid a geometric foundation that is incredibly useful today: the space-time continuum.

The difficulty in graphing problems in relativity is that time can be treated as an imaginary fourth direction. We can graph two dimensions easily, and can approximate graphing three dimensions by "projecting" them onto a two dimensional page. Graphing four dimensions cannot be done on two dimensional paper with pen or pencil.[5] Minkowski found a means to graph space and time together without losing the unique qualities of time. He realized that we could not, as is traditional, keep all of our axes at right angles. These so-called Minkowski diagrams formed a picture to go along with the concepts, and opened the gates to allow physicists to envision the situations in a useful way. They could graph situations as measured by two different observers on a single graph, making their work far more efficient. This also added another tool to the toolkit.

With space and time this closely connected, and with multiple observers graphed on the same diagram, an even more startling discovery was made: there *is* a rotation that allows one to rotate from a spatial direction into a time direction! This rotation manifests itself physically in an unusual fashion: accelerating to a new speed is mathematically identical to rotating with respect to the time access. The "rotation" exists, and we do have the ability to implement it, but the imaginary nature of the time direction means it "looks" completely different than a spatial rotation to us.[6] We still cannot travel backwards in time, but Minkowski's work laid foundations to explain that as well.

[5] Approximations can be made by creating multiple three dimensional graphs, each at a different time, but this is cumbersome and makes the time evolution of the system more difficult to see from the graph alone.

[6] It has been proposed that there are life forms which perceive time no differently than they perceive space, though no concrete evidence exists. It has also been proposed that said life forms live inside an artificial wormhole and are worshipped by the inhabitants of a nearby world that look identical to humans aside from their noses, so the science involved is highly questionable.

4.1 Causality

It is an axiom[7] of science that cause precedes effect. Relativity seems to support this concept.

Imagine two events in the space-time continuum. One happens in a given place and time, and the other happens in a different place at a different time. One can calculate the invariant distance between these two events, in the manner described above. These calculations involve both positive and negative quantities, and the results can be positive, negative or zero. We say that the events are *time-like separated* in one case, *space-like separated* in another, and *light-like separated* in the third. Time-like separated events are those which can be connected by cause and effect, while space-like separated events cannot. Light-like separated events form the border between these two categories.

Let us examine explicit examples. Imagine one drops a rubber ball on the floor, and you consider the two events in question to be the first time it hits the floor, and the second time it hits the floor (after bouncing.) These events are time-like separated because they happen at the same point in space; it is only time that separates them. The cause-effect relationship is possible and established by observation. Any two time-like separated events find that the square of the distance between them in the time direction is greater than the sum of those in the spatial directions. As a result, information about one event can travel at the speed of light at reach the spatial coordinates of the second event *before* the second event takes place.

Now imagine two people are bouncing balls, one on planet Earth, and the other on the rock formerly considered a planet which is named Pluto. If an observer in a spacecraft half way between them believes they are bouncing their balls simultaneously, then this observer will say that there is a separation in space, but not in time. These events are space-like separated; neither person can have any effect on the other person's ball bouncing activity. There is a difference in space, but not in time. Information traveling at the speed of light can leave one event the instant it happens, but by the time it reaches the spatial coordinates of the second event, the event has already happened. There can be no cause-effect relationship.

In the case of light-like separated events, the cause-effect relationship is possible, but the calculation of the four dimensional distance between the two events works out to exactly zero.

These ideas were not terribly new. What was new is the mathematical implications that seemed obvious in one of Minkowski's graphs: events that are

[7] An axiom is an idea that is accepted as true, but which cannot be formally proven. They typically haven't been formally disproved, either, as that would result in the loss of the axiom.

time-like separated from the perspective of one observer are time-like separated from the perspectives of *all* possible observers. Similarly, space-like separations are always space-like, and light-like separations are always light-like. The implication supports the idea that cause and effect relationships must be preserved in space and time, and that "time travel"[8] is impossible. This idea will be revisited when gravity comes into play in a later lesson.

Now, for the nitty gritty mathematical details.

5 Coordinate Axes in Euclidean Space

Newton and his contemporaries worked in "Euclidean space," meaning that all geometry could be defined in the context of neat, perpendicular directions, as described in near totality by the ancient Greek mathematician named Euclid. Initially, we work only in Euclidean space. If your exposure to geometry is limited to what you learned in a public school system, then odds are exceptional that this is the only type of space and geometry you've ever been exposed to.

The definitions of "directions" above amounts to different coordinate axes, each of which has a basis vector associated with it. The discussion of "complete cancelation" refers to the linear independence of the different vectors involved. The typical basis vectors are defined as follows:

$$\begin{aligned}
\mathbf{e}_x &= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\
\mathbf{e}_y &= \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \\
\mathbf{e}_z &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}
\end{aligned}$$

These are certainly not the only definitions. In linear algebra terms, a set of vectors are basis vectors if they *span* the space, meaning that an arbitrary vector $\mathbf{v}$ can be expressed as a sum of the multiples of these basis vectors. (i.e. you can "build" any vector using only the basis vectors.) Thus, they don't actually need to be orthogonal to each other, provided they don't overlap completely.

[8] Here the term "time travel" is used to mean traveling from an arbitrary point in time to another, equally arbitrary point in time, with little or no perceived time lapse between them.

For example,

$$
\begin{aligned}
\mathbf{e}_1 &= \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \\
\mathbf{e}_2 &= \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \\
\mathbf{e}_3 &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}
\end{aligned}
$$

is a valid set of basis vectors but

$$
\begin{aligned}
\mathbf{e}_1 &= \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \\
\mathbf{e}_2 &= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\
\mathbf{e}_3 &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}
\end{aligned}
$$

is not (since the middle vector is a combination of the other two.) The most convenient basis vectors to work with are those that are *orthonormal* to each other, meaning that

$$
\mathbf{e}_i \cdot \mathbf{e}_j = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}
$$

5.1 Rotations

The length of a vector v can be calculated by $|\mathbf{v}| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$, also called the norm of a vector. (This is where the "norm" part came from in "orthonormal.") A matrix R represents a rotation if, and only if, it changes neither the lengths of vectors it acts on nor the angles between vectors it acts on. These two conditions can be written mathematically as follows:

1. $|\mathbf{v}| = |R\mathbf{v}|$ for every possible v.
2. $\mathbf{u} \cdot \mathbf{v} = (R\mathbf{u}) \cdot (R\mathbf{v})$ for every possible u and v.

The second condition works because the dot product between two vectors depends solely upon the lengths of those vectors and the angle between them. If we know the lengths haven't changed and the dot product hasn't changed, then the angle between the two vectors couldn't possibly have changed. To further explore this second condition, we note that

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v}$$

where the superscript T indicates that we are taking the transpose of that vector. So, if

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

then

$$\mathbf{u}^T = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}$$

With this in mind, our second condition reduces to the following:

$$(R\mathbf{u}) \cdot (R\mathbf{v}) = (R\mathbf{u})^T (R\mathbf{v}) = \mathbf{u}^T R^T R \mathbf{v} = \mathbf{u}^T \mathbf{v}$$

or

$$R^T R = I \tag{1}$$

which means

$$R^T = R^{-1}$$

It can be shown (but won't be shown here) that this single condition is equivalent to both of the above conditions. If $R^T = R^{-1}$, then R represents a rotation in Euclidean space.

6 Coordinate Axes in Non-Euclidean Space

The revelation that time is best measured with imaginary numbers has led to another startling discovery: the geometry of the world we live in is *not* Euclidean. Sure, the spatial dimensions are, but once you introduce time, that goes out the window.

When dealing with a non-Euclidean space, one requires new definitions of vectors and the dot product. Rather than use the traditional vectors and matrices, we branch out into *tensors*. Tensors differ from the familiar objects in a very basic way: they are organized collections of variables, rather than organized collections of numbers. So, the component of a position tensor which may be treated as the time coordinate in one reference frame transforms into the time coordinate in all reference frames. The *value* of this coordinate may change drastically from one reference frame to another, but the *meaning* of that coordinate remains unchanged.

Before moving into explicit examples, notation will be confirmed: all objects in Euclidean space will be surrounded with square brackets [], while tensor objects in non-Euclidean space will be surrounded with round brackets (). Euclidean vectors will appear in bold face such as $\mathbf{v}$, while their four dimensional tensor counterparts will appear with a vector arrow above as $\vec{v}$. Components of Euclidean objects are marked with Latin indices such as u_i, while non-Euclidean tensors are marked with Greek indices such as u_α. Euclidean indices can take on the values 1, 2 or 3, while non-Euclidean tensors use 1, 2 and 3 for spatial coordinates and 0 for time.

6.1 Dot Products and the Metric

In non-Euclidean space, it is frequently impossible to use an orthonormal basis of vectors. Thus, the dot products taken in non-Euclidean space must be adapted in some way. In special relativity, there are a few options, and all depend upon the definition of the metric tensor. To begin with, instead of the "length" of a vector, we discuss the "length" of the interval. This is identical to the above definition of a Euclidean norm, save for the fact that it remains squared to avoid the introduction of imaginary numbers.

The metric tensor $g_{\mu\nu}$ is the mathematical object that describes the shape of the geometry being used, represented by interactions between different basis vectors.[9] Its components are formed by taking the dot products of the basis vectors of the geometry. There are four conventions to the metric and tensors in special relativity.

6.1.1 Convention One: Imaginary time, Euclidean metric

One convention (preferred by Stephen Hawking, for example) is to use the identity matrix as the metric, just as is done in Euclidean geometry. Doing so means that the four components of a tensor are different types of components; the time component is strictly imaginary, while other components are strictly real. Thus,

[9] These, technically, exist in Euclidean geometry as well, but using metrics in Euclidean geometry amounts to multiplying by the identity matrix in most cases, which makes utterly no difference at the end of the day.

for a position tensor $\vec{x}$, we have

$$\begin{aligned}
\vec{x}\cdot\vec{x} &= \begin{pmatrix} ict & x & y & z \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ict \\ x \\ y \\ z \end{pmatrix} \\
&= \begin{pmatrix} ict & x & y & z \end{pmatrix} \begin{pmatrix} ict \\ x \\ y \\ z \end{pmatrix} \\
&= (ict)^2 + x^2 + y^2 + z^2 \\
&= -c^2t^2 + x^2 + y^2 + z^2
\end{aligned}$$

Why, do you ask, is the time component *ict* and not just *it*? This is because the units must be consistent for all components of a tensor (or vector). To transform the units of time into units of space, we multiply by the speed of light. This is related to the unpopularity of this convention; people expect consistency from one component to another. Mathematically, they need to have the same units and be in the same set of numbers. While this convention puts all components in the set of complex numbers, some people still find it aesthetically displeasing to find that manifesting itself as a purely imaginary component and three purely real components.

6.1.2 Convention Two: Real tensor components, negative time intervals

The second convention (preferred by Albert Einstein and this author) is to use strictly real values for all components of tensors, and encode the imaginary nature of time in the $g_0 0$ component of the metric. This is done by defining the metric tensor as

$$g_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

such that the dot product of the position vector becomes

$$\begin{aligned}
\vec{x}\cdot\vec{x} &= \begin{pmatrix} ct & x & y & z \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \\
&= \begin{pmatrix} -ct & x & y & z \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \\
&= -c^2t^2 + x^2 + y^2 + z^2
\end{aligned}$$

In this case, events that satisfy the rules of causality (i.e. information from event 1 can reach event 2 by traveling at or below the speed of light) have *negative* lengths, or norms. This is identical to the convention using imaginary time, but allows all tensor components to be real numbers at the cost of having a nontrivial metric.

6.1.3 Convention Three: Real tensor components, positive time intervals

The third convention, which may be the most popular convention, is to define the metric so that causal events are separated by positive intervals. Doing this requires a sign change in the entire metric:

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

resulting in a sign change all the way through the dot product:

$$\begin{aligned}
\vec{x}\cdot\vec{x} &= \begin{pmatrix} ct & x & y & z \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \\
&= \begin{pmatrix} ct & -x & -y & -z \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \\
&= c^2t^2 - x^2 - y^2 - z^2
\end{aligned}$$

This means that most tensors of interest have positive intervals, but the spatial components take on negative values when squared. This is a perfectly valid convention, but one that is often found to be counterintuitive in this manner. Still, most people seem to have an easier time with negative space components than negative intervals, and so it is very popular.

6.1.4 Convention Four: Quaternions

This fourth and least popular convention is a blend of conventions one and three. In this one, tensors are represented with quaternions, which are essentially complex numbers with three imaginary components. Thus, the metric remains Euclidean, and the mix of real and imaginary numbers feels more natural as it comes in an established mathematical form. The dot product looks like this:

$$\begin{aligned}
\vec{x}\cdot\vec{x} &= \begin{pmatrix} ct & ix & jy & kz \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ ix \\ jy \\ kz \end{pmatrix} \\
&= \begin{pmatrix} ct & ix & jy & kz \end{pmatrix} \begin{pmatrix} ct \\ ix \\ jy \\ kz \end{pmatrix} \\
&= (ct)^2 + (ix)^2 + (jy)^2 + (kz)^2 \\
&= c^2t^2 - x^2 - y^2 - z^2
\end{aligned}$$

This text will use convention two, using entirely real components to a tensor, but with a non-Euclidean metric, due in part to the fact that non-Euclidean metrics will be inescapable in lessons 7 through 9.

6.2 Basic Operations with Matrices and Tensors

6.2.1 Matrices

Matrices can be added, subtracted, multiplied, and (after a fashion) divided, subject to certain rules. They are rectangular in shape, with numbers arranged in rows and columns. When multiplying matrices, order matters. In other words, if the two matrices are A and B, then $AB \neq BA$ in most cases. In fact, it may not even be possible to calculate AB when BA is perfectly well defined. For example, let

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}$$

and

$$B = \begin{bmatrix} 4 & 7 \\ 5 & 8 \\ 6 & 9 \end{bmatrix}$$

To calculate the number in row i and column j of AB, we multiply each entry in row i of matrix A by each entry in column j of matrix B and add them up. Thus,

$$\begin{aligned} AB &= \begin{bmatrix} 1 \cdot 4 + 1 \cdot 5 + 1 \cdot 6 & 1 \cdot 7 + 1 \cdot 8 + 1 \cdot 9 \\ 2 \cdot 4 + 2 \cdot 5 + 2 \cdot 6 & 2 \cdot 7 + 2 \cdot 8 + 2 \cdot 9 \\ 3 \cdot 4 + 3 \cdot 5 + 3 \cdot 6 & 3 \cdot 7 + 3 \cdot 8 + 3 \cdot 9 \end{bmatrix} \\ &= \begin{bmatrix} 15 & 24 \\ 30 & 48 \\ 45 & 72 \end{bmatrix} \end{aligned}$$

We can do this because the number of columns in A is the same as the number of rows in B. We cannot calculate BA: there are only two entries in the first row of B, but there are three entries in the first column of A. A has no entries to pair with the bottom row of B. When multiplication of two arbitrary matrices A and B is possible, the resultant matrix has the same number of rows as A and the same number of columns as B. In general, if A has n rows and m columns ("A is $n \times m$"), and B has r rows and s columns ("B is $r \times s$"), then AB is defined if and only if $m = r$, and matrix AB will have n rows and s columns ("AB is $n \times s$"). The entry in row i and column j of matrix A is denoted A_{ij}.

Matrices of all sizes have an algebraic equivalent of zero: every entry in the matrix is a zero. Not all sizes of matrix have an algebraic equivalent of one. The only matrices that behave algebraically like the number one are the identity matrices we call I. I behaves as $AI = IA = A$ no matter what A is. This must be the same I when multiplied by either the left or the right, so (by the rules of multiplication) it must be a square ($n \times n$ for some n) matrix. If I is the identity matrix, then $I_{ij} = 1$ if $i = j$ and $I_{ij} = 0$ if $i \neq j$. Visually, this is a matrix that has the number 1 appear in every entry of the diagonal from the top left corner to the bottom right corner, with the number 0 everywhere else.

Matrices cannot be easily divided. Square matrices can approximate this in certain cases. Some square matrices (but not all) have an inverse; matrix C is the inverse of matrix B if $CB = BC = I$. C is usually denoted B^{-1} if it exists, and it will have the same size as B. Details about which matrices have inverses and which do not will not be provided here; search online for "determinant" to find the answer to that question.

Matrices can be added and subtracted, as well, provided the two matrices involved are of exactly the same size, whatever size that is. If A is $n \times m$, then A can only be added to B if B is also $n \times m$.

The basic operations with matrices, if defined, can be generally expressed as follows:

$$\begin{aligned}
(A+B)_{ij} &= A_{ij} + B_{ij} \\
(A-B)_{ij} &= A_{ij} - B_{ij} \\
(AB)_{ij} &= \sum_k A_{ik} \cdot B_{kj} \\
(A \div B)_{ij} &= AB^{-1} = \sum_k A_{ik} \cdot \left(B^{-1}\right)_{kj}
\end{aligned}$$

The dot product of two vectors **u** and **v** can then be written as

$$\sum_i \sum_j g_{ij} u_i v_j = \sum_i u_i v_i$$

where we have used the Euclidean metric

$$g_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}$$

6.2.2 Tensors

There are comparable situations for tensors, though the indices vary somewhat. When a tensor fills the role of a four dimensional vector, it is called a *four-vector*, and its index is raised, as with $\vec{x}^\mu$ for the position four-vector. When written out in matrix form within curved brackets, they perform in the same manner as Euclidean vectors. The differences lie in the fact that indices can be raised or lowered due to the non-Euclidean metric. The dot product of two four-vectors $\vec{u}$ and $\vec{v}$ is given by

$$\sum_\mu \sum_\nu g_{\mu\nu} u^\mu v^\nu = \sum_\nu u_\nu v^\nu$$

Notice that the indices move when summed over an index with the metric tensor. This process transforms a four-vector into a *one-form*, which has the lowered index. In its explicit form, this amounts to taking the transpose of $\vec{u}$ and multiplying it through the metric from the left as follows:

$$\begin{aligned}
\sum_\mu g_{\mu\nu} u^\mu &= \begin{pmatrix} u_0 & u_1 & u_2 & u_3 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} -u_0 & u_1 & u_2 & u_3 \end{pmatrix}
\end{aligned}$$

When the indices are not explicitly written, there must be some way to distinguish between four-vectors and one-forms. Just as $u^\mu = \vec{u}$ is the four-vector notation, $u_\mu = \tilde{u}$ is the one-form notation.

There will be a lot of summations over indices in the upcoming lessons, so we adopt the *Einstein summation convention*: if the same letter appears as an index that is both raised and lowered in a term, then the sum over that index is implied. This reduces the writing in our dot product from

$$\sum_\mu \sum_\nu g_{\mu\nu} u^\mu v^\nu = \sum_\nu u_\nu v^\nu$$

to

$$g_{\mu\nu} u^\mu v^\nu = u_\nu v^\nu$$

which saves a lot of writing and typing. Similarly, we can write $\vec{u} \cdot \vec{v} = \tilde{u}\vec{v}$, in which the order of the terms matters; $\vec{v}\tilde{u}$ is a very different quantity. (In matrix form, the former becomes a single scalar quantity, while the latter becomes a 4×4 matrix of its own.) In the case of the metric of special relativity itself, as long as both indices are together, we have the above metric: $g_{\mu\nu} = g^{\mu\nu}$. In the case of mixed indices, such as $g^\mu{}_\nu$, then we have the identity matrix.

Tensors are most interesting when one transforms them from one reference frame to another.

6.3 The Lorentz Boost Tensor

The Lorentz boost tensor $\Lambda^{\mu'}{}_\nu$ which transforms the position four-vector

$$x^\nu = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

as measured in reference frame S into the position four-vector

$$x^{\mu'} = \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma\left(ct - \frac{v}{c}x\right) \\ \gamma\left(x - vt\right) \\ y \\ z \end{pmatrix}$$

as measured in reference frame S' is given by

$$\Lambda^{\mu'}{}_\nu = \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where we continue to use the standard

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Thankfully, it can (but won't) be shown that this is the same tensor that transforms all tensors from frame S to frame S', where the $'$ symbol is placed on the indices to indicate which are and are not in frame S'.

6.3.1 Lorentz Boost as a Rotation

In Euclidean space, a matrix R represents a rotation if it satisfies equation 1 on page 9:

$$R^T R = I$$

In non-Euclidean space, the condition is slightly different:

$$R^T g R = g$$

In other words, the metric must be preserved. We can show that this is the case for the Lorentz boost by multiplying out the matrix forms:

$$\begin{aligned}
\Lambda^T g \Lambda &= \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} -\gamma & -\frac{v}{c}\gamma & 0 & 0 \\ \frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} -\gamma^2\left(1 - \frac{v^2}{c^2}\right) & 0 & 0 & 0 \\ 0 & \gamma^2\left(1 - \frac{v^2}{c^2}\right) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= g
\end{aligned}$$

6.4 Invariants

The invariant quantities between reference frames are those whose magnitude does not change. In other words, a quantity is invariant if and only if

$$\vec{u} \cdot \vec{w} = \vec{u}' \cdot \vec{w}'$$

or

$$u_\nu w^\nu = g_{\mu\nu} u^\mu w^\nu = g_{\alpha'\beta'} u^{\alpha'} w^{\beta'} = u_{\beta'} w^{\beta'}$$

Well, the right hand side can be rewritten as follows:

$$\begin{aligned}
g_{\alpha'\beta'} u^{\alpha'} w^{\beta'} &= \Lambda^\mu{}_{\alpha'} \Lambda^\nu{}_{\beta'} g_{\mu\nu} u^{\alpha'} w^{\beta'} \\
&= \Lambda^\nu{}_{\beta'} g_{\mu\nu} u^\mu w^{\beta'} \\
&= g_{\mu\nu} u^\mu w^\nu \\
&= u_\nu w^\nu
\end{aligned}$$

which now matches the left hand side. Thus, all tensors are automatically invariant. In fact, the formal definition of a tensor which distinguishes it from Euclidean objects is that it *does* maintain an invariant interval after a Lorentz transformation.

To see this is all its tedious, matrix form glory, we have:

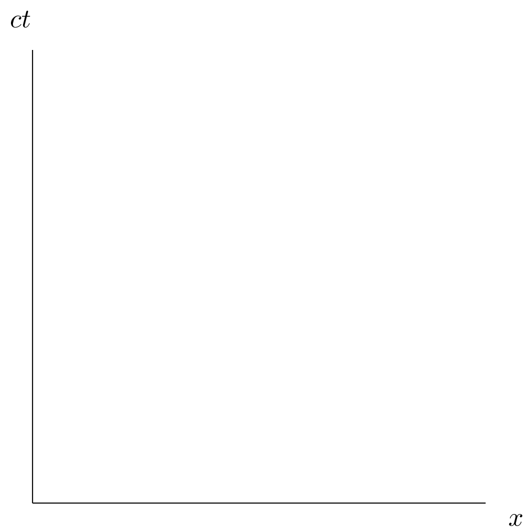
$$\begin{aligned}
\vec{u}' \cdot \vec{w}' &= \begin{pmatrix} u_0 & u_1 & u_2 & u_3 \end{pmatrix} \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \times \\
& \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \end{pmatrix} \\
&= \begin{pmatrix} u_0 & u_1 & u_2 & u_3 \end{pmatrix} \begin{pmatrix} -\gamma & -\frac{v}{c}\gamma & 0 & 0 \\ \frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \end{pmatrix} \\
&= \begin{pmatrix} u_0 & u_1 & u_2 & u_3 \end{pmatrix} \begin{pmatrix} -\gamma^2\left(1-\frac{v^2}{c^2}\right) & 0 & 0 & 0 \\ 0 & \gamma^2\left(1-\frac{v^2}{c^2}\right) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \end{pmatrix} \\
&= \begin{pmatrix} u_0 & u_1 & u_2 & u_3 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \end{pmatrix} \\
&= \begin{pmatrix} -u_0 & u_1 & u_2 & u_3 \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \end{pmatrix} \\
&= \vec{u} \cdot \vec{w}
\end{aligned}$$

Notice that, aside from the u and w terms, this looks very much like the proof that ${\Lambda^{\nu}}_{\beta'}$ is a rotation. This is a direct result of the invariance of a tensor under a Lorentz transformation.

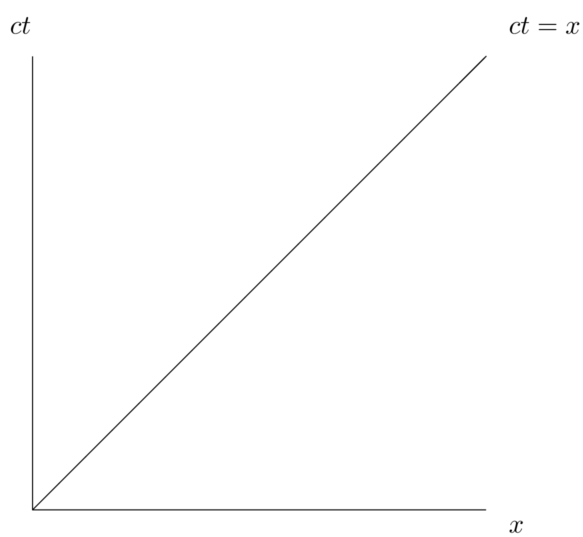
6.5 Minkowski

Minkowski was the first to discover an effective way to create coordinate axes and graph multiple reference frames on a single graph. This was such a useful discovery that they are now known as Minkowski diagrams.

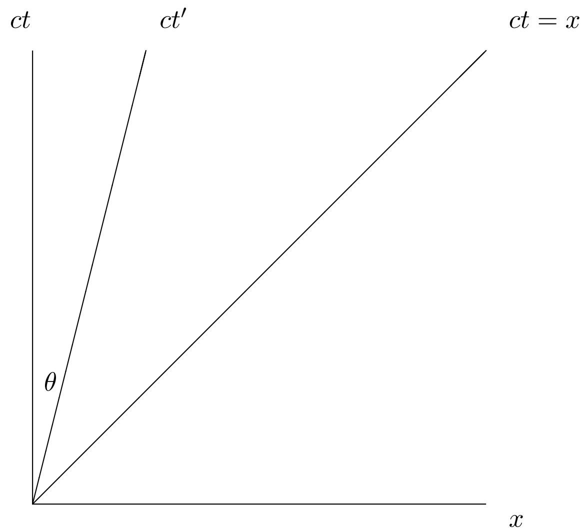
We start with the axes for reference frame S:



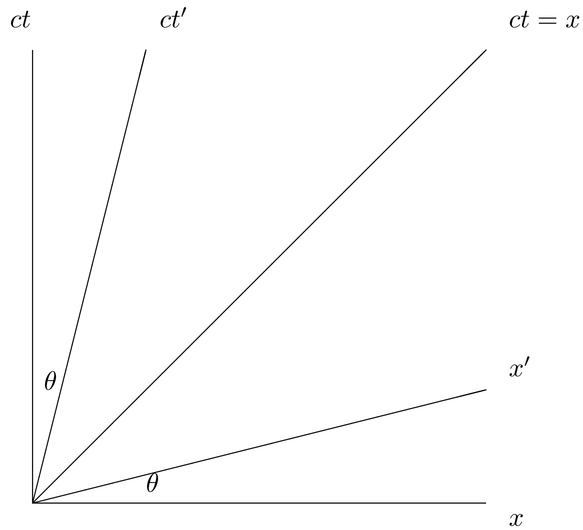
Note that time is on the vertical axis in a Minkowski diagram. Now, we can also insert a line that shows how a photon travels through the geometry of the universe:



The new question becomes: how do we plot axes for events that take place in the reference frame S'? One axis is quite straightforward: if we watch the spatial origin of this reference frame, we find that it will move in the x direction with speed $v = \frac{x}{t}$, so that it makes a line where $x' = y' = z' = 0$, which forms the ct' axis. In frame S, this line meets the ct axis at an angle θ. Using the definition of the tangent function from trigonometry, we see that $\tan\theta = \frac{x}{ct} = \frac{v}{c}$; thus, the relative speed v of the two reference frames provides all the information required to plot the ct' axis:



The question remains: where does the x' axis go? Well, one of the things every observers agrees on is orthogonality. If $\vec{u} \cdot \vec{w} = 0$, then $\vec{u}' \cdot \vec{w}' = 0$. So, which line is orthogonal with the axis we have already drawn? It is the line at which $ct' = 0$. In other words, it is the locus of points that those in frame S perceive as forming the x' axis at the time $t' = 0$. These points can be plotted using the same logic as above, knowing how points transform over time. Somewhat surprisingly, this orthogonal axis doesn't *look* orthogonal the first time it is viewed:



You don't need a protractor to see that the ct' and x' axes are not at a 90° angle to each other. Our exclusively Euclidean experience drives us to assume that right angles are the only form of orthogonality, but when time is measured on an imaginary axis, this is not the case. What matters is the angle made with respect to the $x = ct$ line, which also coincides with the $x' = ct'$ line. In fact, that link is the only line that remains invariant for every axis.

It should be noted that this diagram assumes $v > 0$m/s. If the value is negative, then the S' frame's axes are outside the S frame's axes, but the magnitude of the angle θ remains unchanged.