The Geometry and Trigonometry of Spacetime

John Denker

Abstract

It is well known in physics – and even in pop culture – that time is the fourth dimension.[2] Obviously the time dimension (t) is not exactly the same as the other three (x, y, and z), but it is more closely analogous than many people realize. The tx plane has a geometry and even a trigonometry, just as the xy plane does. For example, what would be called a *slope* in the xy plane is called a *velocity* when it occurs in the tx plane. The relativistic addition-of-velocities rule is nonlinear for the same reason that slope is a nonlinear function of angle: it's just trigonometry. Leveraging our understanding of the xy plane makes it easy to understand the essential features of special relativity, including the fact that no object can be accelerated to the speed of light, and the fact that the speed of light is the same for all observers.

The geometry of spacetime is remarkably similar to the geometry of ordinary space, with just one salient difference. This can be understood at many levels. In this paper, we focus on the simplest levels, using pictures, vectors, and a little bit of trigonometry. (In Ref. 1, we re-explain and validate the ideas using a more formal approach.)

Additional Keywords: rotation/boost; angle/rapidity; slope/velocity; correspondence principle; hyperbolic trigonometry; gorm.

I. REVIEW OF ROTATIONS

We assume the reader has a basic understanding of what it means to rotate something in the xy plane. It is, however, worth reviewing some of the basic ideas, to establish the notation and to lay the foundation for further developments. We begin by examining the effect of a small xy rotation, such as is shown in figure 1.
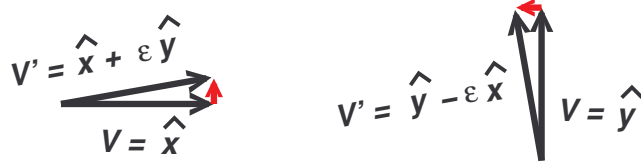

FIG. 1: Small Rotations

We start with a vector V pointing in the $\widehat{x}$ direction, as shown on the left side of figure 1. We rotate it by an angle ϵ to form another vector V'. When ϵ is small, the difference vector $(V' - V)$ is very nearly perpendicular to V. The difference vector is shown in red in the figure. We can express the effect of the rotation as:

$$\begin{aligned} \text{If} \quad & V = \widehat{x} \\ \text{then} \quad & V' = \widehat{x} + \epsilon\,\widehat{y} \end{aligned} \tag{1}$$

If instead we start with a vector pointing in the $\widehat{y}$ direction, as shown on the right side of figure 1, the effect is very similar: rotation causes the vector to pick up a component in a perpendicular direction. For our chosen direction of rotation, the new component is in the *negative x* direction:

$$\begin{aligned} \text{If} \quad & V = \widehat{y} \\ \text{then} \quad & V' = \widehat{y} - \epsilon\,\widehat{x} \end{aligned} \tag{2}$$

Rotating a sum of vectors is the same as rotating each of the summands separately. Since any vector in the xy plane can be written as a linear combination of $\widehat{x}$ and $\widehat{y}$, we can rotate any vector simply by combining equation (1) and equation (2). The overall effect is:

$$\begin{aligned} \text{If} \quad & V = \begin{pmatrix} q\,\widehat{x} \\ +\ r\,\widehat{y} \end{pmatrix} \\[6pt] \text{then} \quad & V' = \begin{pmatrix} q\,\widehat{x} - \epsilon\,r\,\widehat{x} \\ +\ \epsilon\,q\,\widehat{y} + \ r\,\widehat{y} \end{pmatrix} \end{aligned} \tag{3}$$
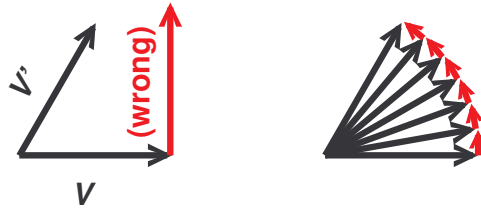
2

FIG. 2: Larger Rotations

When we consider larger rotations, it is not satisfactory to plug in a larger value of ϵ in these equations (perhaps as shown on the left side of figure 2). Instead, we break the large rotation into a sequence of small rotations – all in the xy plane – each of which is well described by equation (3), as shown on the right side of the figure.

If you do the math, combining N rotations each of size $\epsilon = \theta/N$, you get one overall rotation of size θ. In the limit of large N, the overall rotation can be expressed in the familiar form:

$$\text{If} \quad V = \begin{pmatrix} q\,\widehat{x} \\ + \ r\,\widehat{y} \end{pmatrix}$$

$$\text{then} \quad V' = \begin{pmatrix} \cos(\theta)\,q\,\widehat{x} \ - \ \sin(\theta)\,r\,\widehat{x} \\ + \ \sin(\theta)\,q\,\widehat{y} \ + \ \cos(\theta)\,r\,\widehat{y} \end{pmatrix} \tag{4}$$

If this expression for V' is not familiar, a helpful exercise is to set $q = 1$ and $r = 0$ and evaluate this expression for various values of θ from zero to $360°$ in steps of $30°$. Plot the resulting V' vectors.

It is easy to see that equation (3) is a consequence of equation (4), by considering what happens when θ becomes very small. (If that's not obvious, you can verify it by evaluating $\cos(\theta)$ and $\sin(\theta)$ for small values of θ using a calculator. Make sure the calculator accepts θ in radians.) It is also straightforward to prove the converse, namely that equation (4) is necessarily and uniquely the result of applying equation (3) N times ... but we defer the proof to Ref. 1.

II.   INTRODUCTION TO BOOSTS

A small boost is just a small change in velocity ... which is something we already understand in terms of basic physics. Let's see if we can parlay our understanding of small boosts

into an understanding of large boosts, by treating boosts the same way we treated rotations in the previous section. (A good discussion of boosts, with much additional detail, can be found in Ref. 3.)

We start by plotting the position of a particle as a function of time. Such a plot is familiar from high-school physics. We choose to call it by a slightly fancy name, namely the *world line* of the particle, plotted on a *spacetime diagram*. Note that we plot the t axis vertically, as is conventional in the field of relativity. This has several advantages, including the fact that the x axis is horizontal in both figure 3 and figure 1. (Beware that this conflicts with the convention commonly used in introductory physics, which plots motion as x versus t with the t axis horizontal.)
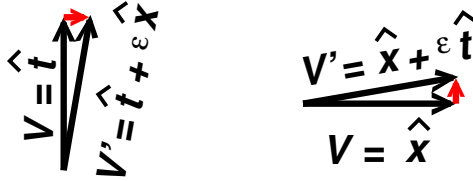

FIG. 3: Small Boosts

For starters, let's consider a particle at rest in our reference frame. If we take a step along the world line of the particle, the t coordinate changes, but the x coordinate does not. We can represent this act – stepping along the world line – by a vector that points in the $\widehat{t}$ direction, such as the vector V on the left side of figure 3.

Next consider a second particle that is moving with a small velocity relative to the first particle. (We say it has been boosted by an amount ϵ in the tx plane.) If we take a step along this particle's world line, the t coordinate changes in the usual way, but we also pick up a slight change in the x coordinate, as shown by the vector V' on the left side of figure 3. Such a step can be represented by a vector of the form $t + \epsilon\, x$. The boost operation itself can be represented as:

$$\begin{aligned} &\text{If} \quad V = \widehat{t} \\ &\text{then} \ \ V' = \widehat{t} + \epsilon\, \widehat{x} \end{aligned} \tag{5}$$

We emphasize that up to this point, we have done nothing subtle or tricky. We take a step in the V' direction along the particle's world line. It has a certain slope, a certain Δx over Δt. This is just velocity. This is just prosaic high-school physics.

4

The units in equation (5) bear some explaining. As a preliminary step, we choose to measure distance in meters and measure time in jiffies, where a jiffy is the time it takes for light to travel one meter. In these units, $c = 1$ meter per jiffy. As the next step, we take to heart the profound correspondence between space and time, and decide that a jiffy just another name for a meter; specifically, a meter of extent in the timelike direction. As the final step, we drop the distinction between ordinary meters (spacelike meters) and jiffies (timelike meters). In these units, $c = 1$. For the same reason, ϵ is dimensionless in equation (5), just as it was in equation (1).

Remark: This is the first subtle step in the derivation. It is not obvious from nonrelativistic mechanics what value to use for c. For now we can consider it an arbitrary conversion factor from spacelike units to timelike units. Later we can ascertain that c must be the speed of light. This can come from the Michelson-Morley experiment, or perhaps from an examination of the Maxwell equations; details are beyond the scope of this document.

Presumably you have noticed that equation (5) and equation (1) have essentially the same form and the same meaning. You may be wondering whether the discussion of boosts will simply retrace the discussion of rotations. It almost will, but not quite. The next equation we write will not quite be identical to equation (2).

It turns out that if you take a vector in the $\widehat{x}$ direction and boost it in the tx plane, the result is as shown on the right side of figure 3. It can be expressed as:

$$\text{If} \quad V = \widehat{x}$$
$$\text{then} \ V' = \widehat{x} + \epsilon \widehat{t} \tag{6}$$

Remark: The plus sign in equation (6) stands in contrast to the minus sign in equation (2). This is the second – and last – subtle step in the whole discussion. This seemingly-innocuous change of sign has the most profound consequences, as will be discussed in section V.

But first, we should finish working out the general formula for boosts. Boosting a sum of vectors is the same as boosting each of the summands separately, so we can combine equation (5) and equation (6) as follows:

$$\text{If} \quad V = \begin{pmatrix} p\,\widehat{t} \\ + \ q\,\widehat{x} \end{pmatrix}$$
$$\text{then} \ V' = \begin{pmatrix} p\,\widehat{t} \ + \ \epsilon\,q\,\widehat{t} \\ + \ \epsilon\,p\,\widehat{x} \ + \ q\,\widehat{x} \end{pmatrix} \tag{7}$$

To understand a large boost, we break it into a sequence of small boosts – all in the tx plane – each of which is well described by equation (7), as shown in figure 4.
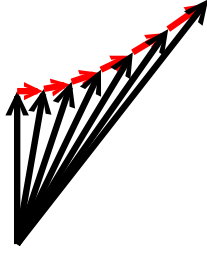

FIG. 4: Larger Boosts

If you do the math, combining N boosts each of size $\epsilon = \rho/N$, you get one overall boost of size ρ. In the limit of large N, the overall boost can be expressed in the form:

$$\text{If} \quad V = \begin{pmatrix} p\,\widehat{t} \\ +\ q\,\widehat{x} \end{pmatrix}$$

$$\text{then} \quad V' = \begin{pmatrix} \cosh(\rho)\,p\,\widehat{t} & + & \sinh(\rho)\,q\,\widehat{t} \\ +\ \sinh(\rho)\,p\,\widehat{x} & + & \cosh(\rho)\,q\,\widehat{x} \end{pmatrix} \tag{8}$$

It is easy to see that equation (7) is a consequence of equation (8), by considering what happens when ρ becomes very small. (If that's not obvious, you can verify it by evaluating $\cosh(\rho)$ and $\sinh(\rho)$ for small values of ρ using a calculator. Some students might benefit from making a graph of the cosh() and sinh() functions, perhaps by hand or perhaps by means of a spreadsheet.) It is also straightforward to prove the converse, namely that equation (8) is necessarily and uniquely the result of applying equation (7) N times ... but we defer the proof to Ref. 1.

Remark: In the relativity literature, the symbol β has long[4] been used to represent the quantity $\tanh(\rho)$. Similarly, γ is used to represent $\cosh(\rho)$.

III.   THE CORRESPONDENCE BETWEEN ROTATIONS AND BOOSTS

We now adopt the tactic of exhibiting the correspondence by putting rotations in the left column and boosts in the right column.

A rotation is a change of *angle* (θ).

The idea of angle is particularly useful, because a sequence of rotations of sizes θ_1, θ_2, θ_3 – all in the yx plane – has the same effect as one big rotation of size θ, where $\theta = \theta_1 + \theta_2 + \theta_3$.

Sometimes we are interested in the slope, which is equal to $\tan(\theta)$. Note that slope does not have the additivity property mentioned in the previous paragraph, except when the slope is very small. For instance, figure 2 can be thought of as six wedges piled on top of each other, tip atop tip. The top of the first wedge has a slope of 1 part in 6, but that does not mean that six wedges combine to make a 1:1 slope. In fact they make a slope that is more than 1.5:1, quite a bit more than a linear extrapolation would have predicted. As you add more wedges, the slope grows more and more quickly. Soon it passes through infinity and becomes negative.

The locus of the arrowheads in figure 2 is a circle. It is the result of rotating a given vector V by various amounts.

A boost is a change of *rapidity* (ρ).

The idea of rapidity is particularly useful, because a sequence of boosts of sizes ρ_1, ρ_2, ρ_3 – all in the tx plane – has the same effect as one big boost of size ρ, where $\rho = \rho_1 + \rho_2 + \rho_3$.

Sometimes we are interested in the 3-velocity, which is equal to $\tanh(\rho)$. Note that velocity does not have the additivity property mentioned in the previous paragraph, except when the velocity is very small. For instance, figure 4 can be thought of as a sequence of six boosts: Each particle in the sequence is moving relative to the previous one with a speed of 1/6th of the speed of light, but that does not mean that particle #6 is moving relative to particle #0 at 100% of the speed of light. In fact, the combined velocity is only 76% of the speed of light, quite a bit less than a linear extrapolation would have predicted. As you make more and more such boosts, the combined velocity keeps growing more and more slowly. It asymptotically approaches the speed of light.

The locus of the arrowheads in figure 4 is a hyperbola.[5] It is the result of boosting a given vector V by various amounts.

Not coincidentally, the coefficients appearing in equation (4) are circular trig functions:

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \qquad (9)$$

Not coincidentally, the coefficients appearing in equation (4) are hyperbolic trig functions:

$$\begin{bmatrix} \cosh(\rho) & \sinh(\rho) \\ \sinh(\rho) & \cosh(\rho) \end{bmatrix} \qquad (10)$$

IV.   ROTATED AND BOOSTED OBSERVERS

In previous sections, we have considered multiple vectors (notably V and V') using only one frame of reference. We now wish to change viewpoints, and henceforth consider only one vector at a time, using multiple frames of reference (notably Joe's frame and Moe's frame). This is sometimes called using passive transformations, in contrast to the previous active transformations. The observers will be rotated and boosted, but the vector V will not. The basis vectors in Joe's frame are $\widehat{t}_J$, $\widehat{x}_J$, and $\widehat{y}_J$. Similarly, the basis vectors in Moe's frame are $\widehat{t}_M$, $\widehat{x}_M$, and $\widehat{y}_M$.

Keep in mind that our spacetime diagrams have the t axis pointing upward, as is conventional in the field of relativity.

These ideas are depicted in figures 5, 6, and 7. Each figure has two diagrams, one for rotations (on the left) and one for boosts (on the right).

The lines in the diagram are drawn according to the following rules:

	constant x	constant y or t
Joe	solid blue	dashed blue
Moe	solid red	dashed red

Moe's frame is angled relative to Joe's frame. In the three figures, the angle is $\theta = 0.25$, 0.50, and 0.75 radians respectively.

Moe's frame is moving relative to Joe's frame. In the three figures, the rapidity is $\rho = 0.25$, 0.50, and 0.75 respectively.
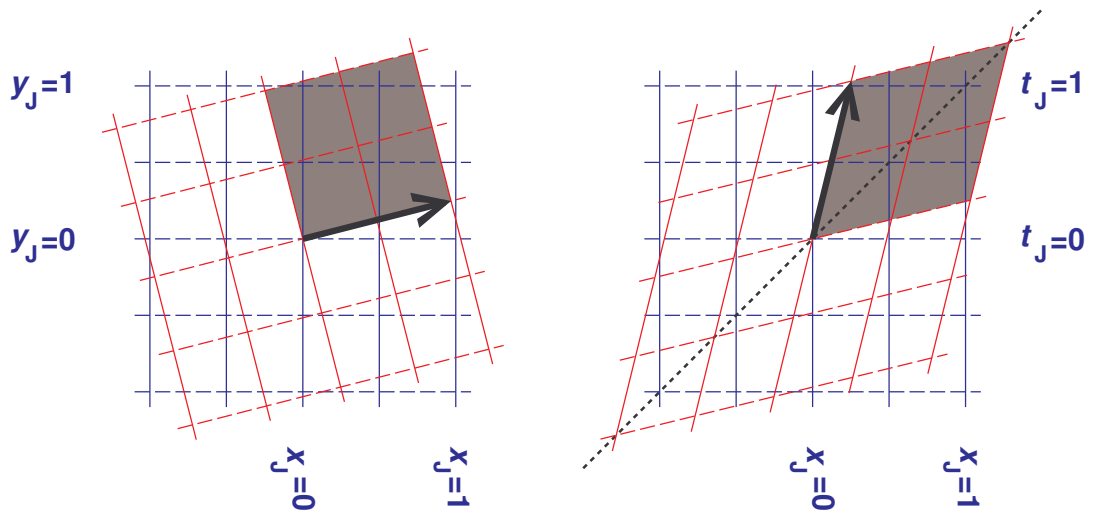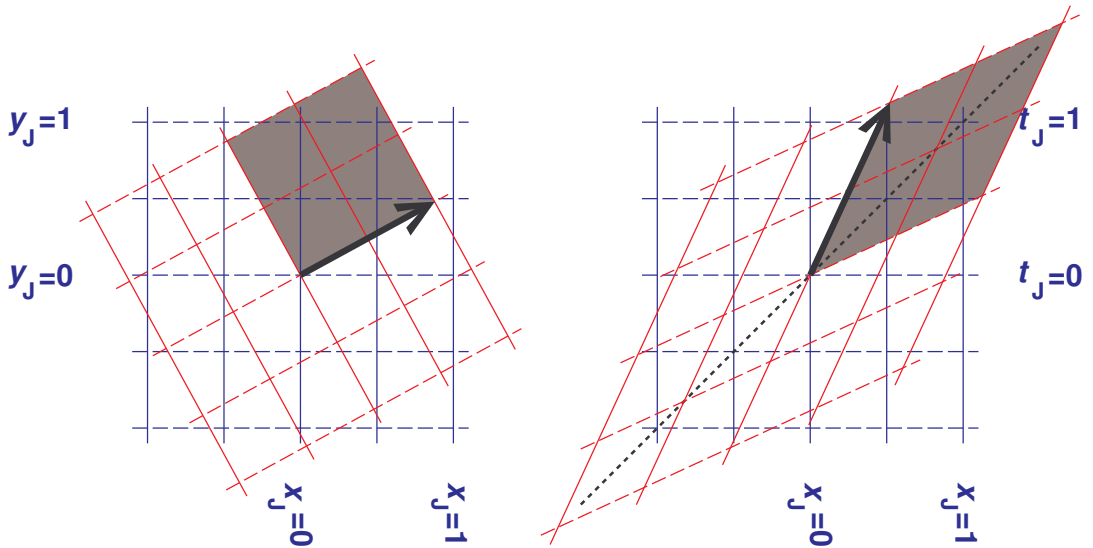
FIG. 5: Rotation and Boost (0.25)

FIG. 6: Rotation and Boost (0.50)

A rotation mixes x and y, in the following sense:

A vector that has only an x component in Moe's frame has both x and y components in Joe's frame. The y per unit x is what Joe calls the slope.

A boost mixes t and x, in the following sense:

A vector that has only a t component in Moe's frame has both t and x components in Joe's frame. The x per unit t is what Joe calls the 3-velocity.
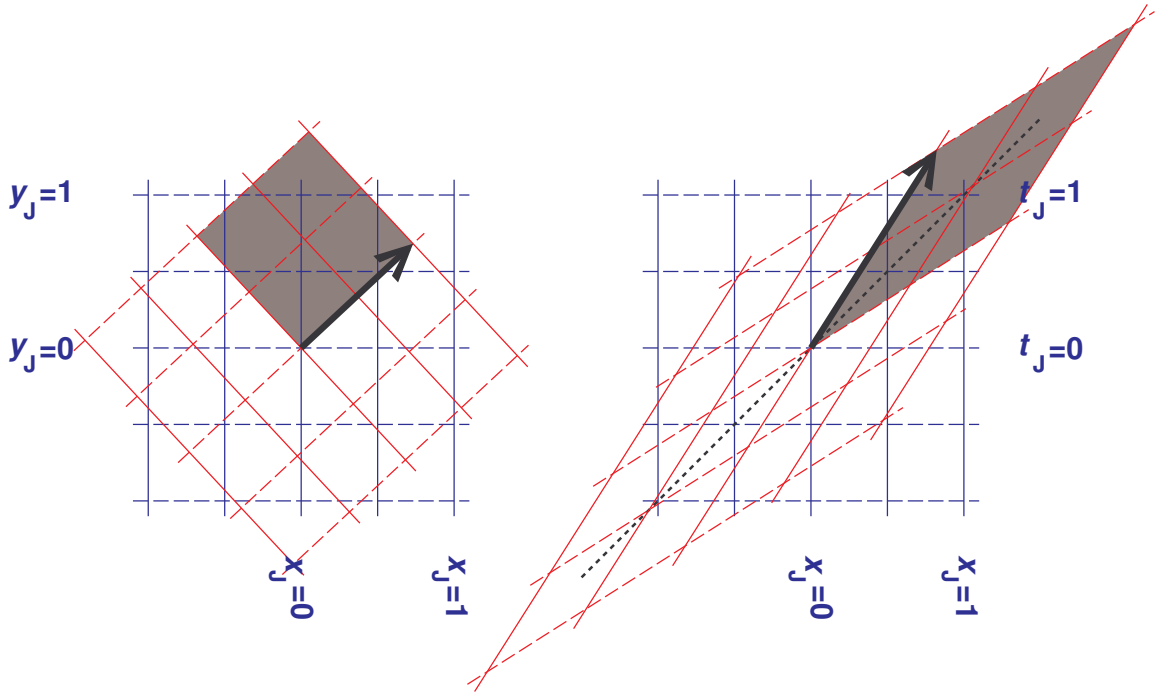
FIG. 7: Rotation and Boost (0.75)

We define the slope as $m := \tan(\theta)$. (Recall that we are measuring angles in radians.)

When θ is small, $\tan(\theta)$ is equal to θ to first order. That means slope is synonymous with angle, when the slope is small.

As θ increases from zero to 90 degrees, $\tan(\theta)$ starts out equal to θ but soon becomes larger – eventually very much larger than θ. This can be seen in the figures: Find the x and y components by projecting the black arrow onto the axes. As always, the slope is y per unit x.

We define the 3-velocity as $v = \tanh(\rho)$. (Recall that we are measuring velocities in units of meters per jiffy.)

When ρ is small, $\tanh(\rho)$ is equal to ρ to first order. That means velocity is synonymous with rapidity, when the velocity is small.

As ρ increases from zero to very large values, the 3-velocity starts out equal to ρ, but soon becomes smaller – eventually very much smaller than ρ. This can be seen in the figures: Find the x and t components by projecting the black arrow onto the axes. As always, the 3-velocity is x per unit t.

10

When Moe's frame is angled relative to Joe's frame by an amount θ, Moe will decompose the vector V into components one way, while Joe will decompose the same vector differently:

$$V = \begin{pmatrix} q\,\widehat{x}_M \\ +\ r\,\widehat{y}_M \end{pmatrix} \qquad (11)$$

$$= \begin{pmatrix} \cos(\theta)\,q\,\widehat{x}_J\ -\ \sin(\theta)\,r\,\widehat{x}_J \\ +\ \sin(\theta)\,q\,\widehat{y}_J\ +\ \cos(\theta)\,r\,\widehat{y}_J \end{pmatrix}$$

for arbitrary scalars q and r. This conveys the same physics as equation (4).

When Moe's frame is moving relative to Joe's frame by an amount ρ, Moe will decompose the vector V into components one way, while Joe will decompose the same vector differently:

$$V = \begin{pmatrix} p\,\widehat{t}_M \\ +\ q\,\widehat{x}_M \end{pmatrix} \qquad (12)$$

$$= \begin{pmatrix} \cosh(\rho)\,p\,\widehat{t}_J\ +\ \sinh(\rho)\,q\,\widehat{t}_J \\ +\ \sinh(\rho)\,p\,\widehat{x}_J\ +\ \cosh(\rho)\,q\,\widehat{x}_J \end{pmatrix}$$

for arbitrary scalars p and q. This conveys the same physics as equation (8).

We now wish to discuss lengths and angles. We have to be careful, because a spacetime diagram is not an entirely faithful representation of spacetime. That is, we wish to describe events that take place in the tx plane, which has one spacelike dimension and one timelike dimension. When we represent this on paper, though, the paper has two spacelike dimensions. The mapping from spacetime to paper distorts the lengths and angles.

In particular, in each of the six diagrams, the gray-shaded area is a unit square. That is, edges of each gray area all have unit length, and adjacent edges are perpendicular. This is obvious in the rotation diagrams, but somewhat less obvious in the boost diagrams. (Also we need to be careful what we mean by "unit" length, as discussed below.)

In all cases, it is obvious to Moe that the gray area is square, since it is aligned with his axes. Joe also recognizes the gray area as being square, but to confirm this he has to do some calculations.

As always, we use the *dot product* to formalize and quantify our notion of lengths and angles.

The familiar dot product can be defined by postulating

$$\widehat{x} \cdot \widehat{x} = \widehat{y} \cdot \widehat{y} = \widehat{z} \cdot \widehat{z} = 1 \qquad (13)$$
$$\widehat{x} \cdot \widehat{y} = \widehat{y} \cdot \widehat{z} = \widehat{z} \cdot \widehat{x} = 0$$

The four-dimensional dot product can be defined by postulating

$$\widehat{t} \cdot \widehat{t} = -1$$
$$\widehat{x} \cdot \widehat{x} = \widehat{y} \cdot \widehat{y} = \widehat{z} \cdot \widehat{z} = 1 \qquad (14)$$
$$\widehat{t} \cdot \widehat{x} = \widehat{x} \cdot \widehat{y} = \widehat{y} \cdot \widehat{z} = \widehat{z} \cdot \widehat{t} = 0$$

The dot product is commutative and distributes over addition in the usual way.

We hereby define the *gorm* of a vector to be the dot product of the vector with itself. If the gorm is positive, we say the vector is *spacelike*. If the gorm is negative, we say the vector is *timelike*. If the gorm is zero, we say the vector is *lightlike*, or equivalently *null*.

Let A and B be two points (i.e. events) in spacetime, and let V be the displacement vector, $V := B - A$. If V is spacelike, we define the *chord length* to be $\sqrt{V \cdot V}$, which is also called the *invariant spatial interval*. If on the other hand V is timelike, we define the *chord time* to be $\sqrt{-V \cdot V}$, which is also called the *invariant time interval*. For a null vector, the invariant interval is zero. Chord length and chord time are discussed in Ref. 8.

In equation (14), the minus sign in the definition of $\widehat{t} \cdot \widehat{t}$ tells us that our notion of perpendicularity in spacetime will be different from the corresponding notion in ordinary space. This peculiar minus sign is necessary to be consistent with the peculiar plus sign in equation (6), and vice versa.

We say two non-null vectors are *perpendicular* or equivalently *orthogonal* if and only if their dot product is zero.

In simple Euclidean space, only one notion of "norm" is needed, namely $|V| := \sqrt{V \cdot V}$. The quantity $V \cdot V$ is simply called the norm squared.

In Minkowski space, we need to think clearly about three different quantities: the gorm, the chord time, and the chord distance. It's not clear we should define a "norm" at all, but if we do, it must be defined differently for timelike and spacelike vectors. The gorm must not be thought of as the norm squared.

The vector $q\widehat{x} + r\widehat{y}$ is perpendicular to $r\widehat{x} - q\widehat{y}$, for any coefficients q and r, assuming both $\widehat{x}$ and $\widehat{y}$ are spacelike.

Among other things, this means that in figure 1, the difference vectors are perpendicular to the position vectors in accordance with equation (4).

It also allows Joe to verify that on the left side of figure 5, adjacent edges of gray squares are in fact perpendicular.

All observers agree that the edges of the gray squares have gorm equal to 1.

Beware that the literature is horribly inconsistent in its definition of "the" invariant interval, aka "the" spacetime interval, often blurring the distinction between timelike intervals and spacelike intervals.

The vector $p\widehat{t}+q\widehat{x}$ is perpendicular to $q\widehat{t}+ p\widehat{x}$, for any coefficients p and q, assuming $\widehat{t}$ is timelike and $\widehat{x}$ is spacelike.

Among other things, this means that in figure 3, the difference vectors are perpendicular to the position vectors in accordance with equation (8).

It also allows Joe to verify that on the right side of figure 5, adjacent edges of the gray square are in fact perpendicular, even though they may not look that way at first glance.

All observers agree that the spacelike edge of the gray square has gorm equal to 1. They also agree that the other edge – the timelike edge – has gorm equal to -1. We call this a unit square, since -1 fits the mathematical definition of *unit*.

V.   DISCUSSION

A.   Evidence Against Competing Hypotheses

The geometry of the xy plane is determined by the coefficients in equation (3).

$$\begin{bmatrix} 1 & -\epsilon \\ \epsilon & 1 \end{bmatrix} \quad (15)$$

The geometry of the tx plane is determined by the coefficients in equation (7).

$$\begin{bmatrix} 1 & \epsilon \\ \epsilon & 1 \end{bmatrix} \quad (16)$$

13

Our attention is captured by the element in the upper-right corner, since that is the only difference between equation (15) and equation (16).

This element tells us about the so-called *breakdown of simultaneity at a distance*. That is, it tells us that time as perceived by Moe is offset relative to the time perceived by Joe – offset by an amount depending on the distance (x), and also depending on Moe's velocity relative to Joe. This is illustrated by the right half of figure 3 and quantified by equation (6).

Throughout most of history, i.e. before relativity, this contribution was assumed to be zero. Everybody tacitly assumed that timekeeping was independent of location and independent of velocity. Indeed it is very nearly independent, unless you have a combination of large distances, multiple reference frames with high relative velocity, and highly accurate time-keeping – as you can see from the structure of equation (6). It was not until the end of the 19th century that Michelson and Morley did experiments that disproved the pre-relativity assumptions.

Michelson and Morley observed that the speed of light was the same in all reference frames. The geometry of the tx plane as defined by equation (8) has exactly this property, as you can see in e.g. figure 6. The loosely-dashed diagonal line represents the speed of light. Notice how Moe's t axis and x axis are both tilted toward the diagonal, so that anything that lies along the diagonal in Joe's frame also lies along the diagonal in Moe's frame.

The Michelson-Morley experiment, and innumerable others like it, are sufficient to tell us that the upper-right element in equation (16) must be $+\epsilon$, not zero or anything else.

Einstein said any theory should be as simple as possible, but not simpler. We have seen that special relativity is as simple as it possibly could be. The geometry of the tx plane is as closely analogous to the geometry of the xy plane as it possibly could be. The only difference is that the sign in equation (6) differs from the sign in equation (2). Virtually all of special relativity can be seen as arising from this one little change. Specifically:

  – This is what makes spacetime different from plain old space.

  – This is what makes the timelike dimension t different from the spacelike dimensions x, y, and z.

  – This is why you cannot accelerate an object to the speed of light or beyond.

  – This is why the speed of light is the same for all observers.

It is worth taking a moment to consider what would happen if the upper-right element in equation (16) were changed to a $-\epsilon$, thereby making boosts behave the same as ordinary

spacelike rotations. That would make the theory very simple ... but alas *too* simple, because it would make the time axis just another space axis. You would be able to interchange the t axis and the x axis by an ordinary rotation, and you could travel backward and forward in time just as easily as you move east and west.

B.   What is Invariant and What Isn't

In section IV, we used passive transformations. That is, in figure 6, there is one vector – only one vector – shown in black. If we switch from Joe's reference frame to Moe's reference frame, the vector is unaffected. The vector neither knows nor cares who, if anyone, is observing it. For background on this, see Ref. 7.

We did this to make the following point: According to the "geometric" approach, special relativity primarily describes the geometry of spacetime. It does not depend on anything funny happening to the vectors, clocks, rulers, odometers or other objects that inhabit spacetime. Objects are neither Lorentz contracted nor time dilated; they are completely unaffected by boosts. The *components* obtained by projecting a vector onto this-or-that reference frame are affected, but that is a property of the projective geometry of the situation, not a property of the vector itself. This idea and its consequences are discussed in Ref. 8.

VI.   EXERCISE

As suggested in Ref. 3, suppose we have a spaceship. Initially it is at rest with respect to the earth. It then accelerates steadily, such that the passengers aboard the spaceship feel a constant $1G$ acceleration, always in the same direction. Question: At the end of one year, how fast is the spaceship going, relative to the earth?

Solution: In the first second, the spaceship picks up a velocity of $v_1 = 9.8$ meters per second. We know this from basic physics, and the definition of $1G$. This corresponds to a rapidity of $\rho_1 = v_1/c = 3.27\text{e-}8$. The rapidity-change in each second thereafter is the same, as you can verify by considering how things look in a reference frame instantaneously comoving with the spaceship. After one year (3.16e7 seconds) this adds up to $\rho = 1.03$, which is an amusing result in itself: one G for one year gives you about one unit of rapidity.

Since $\tanh(1.03) = 0.77$, we conclude that the 3-velocity is 77% of the speed of light.

Our method of solution made heavy use of the additivity property discussed in section III. This is easy to understand by analogy to rotations.

There are other ways of solving this problem, but none easier or more elegant.

VII.   PEDAGOGICAL IMPLICATIONS

1. Spacetime diagrams are highly recommended as a way of introducing the main ideas of special relativity. Among other things, they can be used to introduce the idea of treating time and space on the same footing, and treating boosts and rotations on the same footing. Ref. 3 is the "gold standard" in this area.

   It is, alas, rather time-consuming to draw the spacetime diagrams with the required accuracy. Therefore it is worth finding a way to do some of the work in advance.

   A very useful first step is to prepare some "spacetime graph paper" containing two sets of axes: one set of ordinary unboosted (x, t) axes, plus another set of boosted axes, carefully drawn with the correct angles and the correct units.

   Often it is remarkably easy to *compute* the diagrams. Many drawing programs have a way of applying an arbitrary linear transformation to a sub-image. This can be used to boost vectors, or even more complex objects. All the diagrams in this paper were computed.

2. Writing the Lorentz transformation in terms of hyperbolic trig functions [equation (10)] makes it well-nigh unforgettable, in analogy to a rotation written in terms of the more-familiar circular trig functions [equation (9)].

3. A further advantage is that equation (10) gives a complete and correct description of special relativity. In contrast, saying "rulers contract" and "clocks slow down" has several problems. For starters, it gives an *incomplete* description, since it describes only two of the four terms in equation (10). This is a common source of confusion among non-experts. At the very least, you need to add a statement of *breakdown of simultaneity at a distance*.

4. The traditional approach to teaching relativity emphasizes that it is weird and rife with paradoxes, using rulers that can't be trusted and clocks that can't be trusted. In

contrast, we argue[8] that it is pedagogically better to teach students that relativity is not particularly weird or paradoxical; indeed it is in many ways closely analogous to the geometry of ordinary Euclidean space. Time and space are most easily understood in terms of proper distance and proper time, which do *not* contract or slow down.

5. For that matter, in an introductory discussion, you need not say anything whatsoever about clocks or rulers. All you really need is equation (5) which describes low-velocity behavior, as inherited from basic physics via the correspondence principle, plus equation (6) which describes the offset of simultaneity at a distance. Everything else follows from that.

6. In this paper we have considered only xy rotations and tx boosts. In general, one must consider all combinations of xy, yz, and zx rotations, plus tx, ty, and tz boosts. This is the topic of another paper[1].

   Note that when we consider combinations of boosts in more than a single direction, it is provably impossible to have a full discussion of boosts without also discussing rotations. That is because it is easy to construct a sequence of boosts whose net effect is a pure rotation (with no net boost). An example of this is the Thomas precession.

   To say the same thing another way, purely spacelike rotations form a subgroup of the Lorentz group, whereas boosts do not. (If you don't know what this means, don't worry about it.)

7. The sophisticated reader will have noticed that many of the equations in this paper could be expressed more compactly using matrices. The formatting of the equations is intentionally suggestive. However, we have avoided relying on matrices or linear-algebra ideas, so that the paper will be understandable to a wider audience. The geometry of spacetime can be understood using nothing beyond high-school-level vectors and trigonometry.

8. Similarly, the sophisticated reader will have noticed that what we have called an infinitesimal rotation is only a half-step removed from being a Lie derivative. As early as 1906, Poincaré[9] showed that the Lorentz transformations formed a group, and discussed the group's Lie algebra. However, if you don't know what that means, don't worry about it.

9. Our approach upholds the pedagogical principle that learning proceeds from the known to the unknown. We assume the students have some previous knowledge of rotations and vectors in ordinary 3-space. We build on this to explain the closely-analogous concepts of boosts and 4-vectors in spacetime. As usual, a side-benefit is that it broadens and deepens their understanding of the previously-known topics.

---

[1] John Denker, "Multi-Dimensional Rotations, Including Boosts"
   http://www.av8n.com/physics/rotations.htm

[2] H. Minkowski, "Raum und Zeit," 80. Versammlung Deutscher Naturforscher (Köln, 1908). Published in Physikalische Zeitschrift **10** 104-111 (1909) and Jahresbericht der Deutschen Mathematiker-Vereinigung **18** 75-88 (1909).
   http://de.wikisource.org/wiki/Raum_und_Zeit_(Minkowski)

[3] E.F. Taylor and J.A. Wheeler, **Spacetime Physics** (W.H. Freeman, San Francisco, CA, 1963).

[4] A. Einstein, "Zur Elektrodynamik bewegter Körper", Annalen der Physik **17**, 891 (1905).
   http://www.physik.uni-augsburg.de/annalen/history/papers/1905_17_891-921.pdf

[5] A.P. French **Special Relativity** (Norton, New York, 1968).

[6] C.W. Misner, K.S. Thorne, and J.A. Wheeler, **Gravitation** (W.H. Freeman, San Francisco, CA, 1973).

[7] John Denker, "Two Types of Vector : Physics and/or Components".
   http://www.av8n.com/physics/two-vector.pdf

[8] John Denker, "Odometers and Clocks in Introductory Relativity".
   http://www.av8n.com/physics/odometer.pdf

[9] H. Poincaré, "Sur la dynamique de l'electron," Rendiconti del Circolo matematico di Palermo, **21**, 129–176 (1906); see also Comptes Rendues **140**, 1504–8 (1905).
   http://www.soso.ch/wissen/hist/SRT/P-1905.pdf