

1 Introduction; Puzzle

Consider the following puzzle:

Suppose an interstellar rocket starts from rest and accelerates in a straight line such that the passengers feel one Gee for one year. How fast are they going at the end of the year?

There are at least two ways to solve the problem.

- You can grind it out using the Lorentz transformation equations – but I don’t recommend it.
- You can solve the puzzle in less time that it takes to state it, if you understand a few things about the structure of spacetime.

The steady acceleration of the rocket is profoundly analogous to a steady rotation. We can use what we know about ordinary rotations to learn some things that are useful for solving this puzzle and many others.

There are some really deep ideas involved, but once you learn to trust the formalism it is easy to remember and easy to use. I don’t even bother to remember the Lorentz transformation in terms of square roots or anything like that – equation 2 is incomparably easier to remember.

2 Rapidities and Boosts

We need to introduce two concepts:

- A *rapidity* will turn out to be analogous to an angle, and intimately related to velocity.¹
- A *boost* is a change in rapidity, just as a rotation is a change in angle.

Executive summary: This section will demonstrate that *rapidities are additive for compound boosts*. If you already know what that means, you can skip ahead; go directly to section 3. Otherwise, read on. Consider the following comparison; rotations are in the left column and boosts are in the right column:

Suppose Bob’s reference frame is rotated relative to Alice’s. In particular, let it differ by a rotation in the XY plane. This “mixes up” the X and Y coordinates of an object. That is, the X coordinate measured by Alice depends on the X coordinate *and the Y coordinate* measured by Bob. Similarly, the Y coordinate measured by Alice depends on the Y coordinate *and the X coordinate* measured by Bob.

The exact expression is:

$$\begin{bmatrix} P_{X(A)} \\ P_{Y(A)} \end{bmatrix} = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix} \begin{bmatrix} P_{X(B)} \\ P_{Y(B)} \end{bmatrix} \quad (1)$$

Suppose Bob’s reference frame is moving relative to Alice’s. In particular, let it differ by a boost in the X direction. This mixes up the X and T coordinates of an object. The X coordinate measured by Alice will depend on the X coordinate *and the T coordinate* measured by Bob. Similarly, the T coordinate measured by Alice depends on the T coordinate *and the X coordinate* measured by Bob – which is what we call the breakdown of simultaneity at a distance.

The exact expression is:

$$\begin{bmatrix} P_{T(A)} \\ P_{X(A)} \end{bmatrix} = \begin{bmatrix} \cosh(\rho) & \sinh(\rho) \\ \sinh(\rho) & \cosh(\rho) \end{bmatrix} \begin{bmatrix} P_{T(B)} \\ P_{X(B)} \end{bmatrix} \quad (2)$$

¹Throughout this document, the term “velocity” will refer to the ordinary velocity, also called the coordinate velocity, $v = dx/dt$, not to be confused with the proper velocity, $u = dx/d\tau$.

where ϕ is the angle of rotation. We recognize the rotation matrix for a rotation in the XY plane, namely

$$R(\phi) := \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix} \quad (3)$$

The slope of Bob's X-axis relative to Alice's is given by

$$\frac{dy}{dx} = m = \tan(\phi) \quad (5)$$

For small angles, the slope is equal to the angle (in radians). Equation 1 uses ordinary circular trig functions (sin and cos). The matrix element in the upper-right corner has a minus sign.

We see that a rotation in the XY plane changes the slope, whereas a rotation in the XT plane changes the velocity.

The slope is more directly analogous to the reduced velocity than to the 4-velocity.

If Bob is rotated relative to Alice, and Carol is rotated relative to Bob, the rotation matrix for the Alice \rightarrow Carol transformation is just the product of the Alice \rightarrow Bob and Bob \rightarrow Carol matrices.

If Bob is rotated relative to Alice by a tiny amount, and Carol is rotated relative to Bob 99 times as much (which we construct by compounding lots of tiny rotations), the rotation matrix for the Alice \rightarrow Carol transformation is just the 100th power of the Alice \rightarrow Bob matrix. That is,

$$R(N\epsilon) = [R(\epsilon)]^N \quad (8)$$

The rotation matrix for a tiny angle is:

$$R(\epsilon) = \begin{bmatrix} 1 & -\epsilon \\ \epsilon & 1 \end{bmatrix} \quad (10)$$

You can compute this as a direct consequence of equation 1, or you can take this as a starting point – the next few paragraphs will re-derive equation 1 starting from equation 10 and equation 8.

where ρ is the rapidity of the boost. We recognize the rotation matrix for a rotation in the XT plane (i.e. a boost), namely

$$B(\rho) := \begin{bmatrix} \cosh(\rho) & \sinh(\rho) \\ \sinh(\rho) & \cosh(\rho) \end{bmatrix} \quad (4)$$

The X-component of Bob's reduced velocity relative to Alice is given by

$$\frac{dx}{c dt} = \frac{v_x}{c} = \tanh(\rho) \quad (6)$$

Unlike equation 3, there are no minus signs in equation 4. Equation 2 uses hyperbolic trig functions (sinh and cosh).

The reduced velocity should not be confused with the 4-velocity. The X-component of the 4-velocity is given by:

$$\frac{dx}{c d\tau} = \frac{u_x}{c} = \sinh(\rho) \quad (7)$$

For small rapidities, both v_x and u_x are equal to the rapidity (in units where $c = 1$). It is not at all clear which should be considered “the” velocity.

If Bob is moving relative to Alice, and Carol is moving relative to Bob, the boost matrix for the Alice \rightarrow Carol transformation is just the product of the Alice \rightarrow Bob and Bob \rightarrow Carol matrices.

If Bob is moving relative to Alice by a tiny amount, and Carol is moving relative to Bob 99 times as much (which we construct by compounding lots of tiny boosts), the boost matrix for the Alice \rightarrow Carol transformation is just the 100th power of the Alice \rightarrow Bob matrix. That is,

$$B(N\epsilon) = [B(\epsilon)]^N \quad (9)$$

The boost matrix for a tiny rapidity is:

$$B(\epsilon) = \begin{bmatrix} 1 & \epsilon \\ \epsilon & 1 \end{bmatrix} \quad (11)$$

You can compute this as a direct consequence of equation 2, or you can take this as a starting point – the next few paragraphs will re-derive equation 2 starting from equation 11 and equation 9.

This can be written in the suggestive form

$$R(\epsilon) = \exp(\epsilon L_{xy}) = 1 + \epsilon L_{xy} \quad (12)$$

where

$$L_{xy} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad (13)$$

is called the *generator* of rotations in the XY plane.

This can be written in the suggestive form

$$B(\epsilon) = \exp(\epsilon L_{tx}) = 1 + \epsilon L_{tx} \quad (14)$$

where

$$L_{tx} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad (15)$$

is called the *generator* of rotations in the XT plane (i.e. boosts in the X direction).

Note: If you are not familiar with taking e to the power of a matrix, just write out e^x as a power series, and let x be a matrix. Then you're all set. It is straightforward to raise x to integer powers, just by repeated multiplication. Powers of L_{xy} and L_{tx} are particularly easy to compute.

By compounding N tiny rotations, we obtain:

$$R(N\epsilon) = \exp(N\epsilon L_{xy}) = (1 + \epsilon L_{xy})^N \quad (16)$$

where $N\epsilon$ is not necessarily small, even though ϵ is small.

Since any angles ϕ_1 and ϕ_2 can be written as multiples of epsilon, we can always write

$$R(\phi_1) = \exp(\phi_1 L_{xy}) \quad (18)$$

and

$$R(\phi_2) = \exp(\phi_2 L_{xy}) \quad (20)$$

By compounding N tiny boosts, we obtain:

$$B(N\epsilon) = \exp(N\epsilon L_{tx}) = (1 + \epsilon L_{tx})^N \quad (17)$$

where $N\epsilon$ is not necessarily small, even though ϵ is small.

Since any boosts ρ_1 and ρ_2 can be written as multiples of epsilon, we can always write

$$B(\rho_1) = \exp(\rho_1 L_{tx}) \quad (19)$$

and

$$B(\rho_2) = \exp(\rho_2 L_{tx}) \quad (21)$$

Exercise: calculate the N th power of L_{xy} for all N from 0 to 5. Then weight each one by $\phi^N/N!$ and add them up, matrix-element-by-matrix-element. It is interesting how the power series for sine and cosine emerge in the proper places, confirming that equation 18 is consistent with equation 3. We see that exponentials are intimately related to sines and cosines. This is reminiscent of Euler's formula, $\exp(i\theta) = \cos(\theta) + i \sin(\theta)$, but using real-valued matrices rather than complex numbers.

Since it is particularly easy to multiply exponentials (just add the exponents), we find that

$$\begin{aligned} R(\phi_1)R(\phi_2) &= \exp[(\phi_1 + \phi_2) L_{xy}] \\ &= R(\phi_1 + \phi_2) \end{aligned} \quad (22)$$

This proves something that everybody takes for granted, namely that angles are additive for compound rotations (in the same plane). This is exceedingly useful. This is why angles were invented.

Angles are much better behaved than, say, slopes. If A is sloping relative to B, and B is sloping relative to C, the compound slope is *not* the sum of the contributions (except when all slopes are small).

and since it is particularly easy to multiply exponentials (just add the exponents), we find that

$$\begin{aligned} B(\rho_1)B(\rho_2) &= \exp[(\rho_1 + \rho_2) L_{tx}] \\ &= B(\rho_1 + \rho_2) \end{aligned} \quad (23)$$

This proves something that might not have been 100% obvious a few minutes ago, namely that the rapidities are additive for compound boosts (in the same direction). This is exceedingly useful. This is why rapidities were invented.

Rapidities are much better behaved than, say, velocities. If A is moving relative to B, and B is moving relative to C, the compound velocity is *not* the sum of the contributions (except when all velocities are small).

3 Solution of the Puzzle

Returning to the puzzle posed in section 1: After one second, the passengers have picked up a velocity of 9.80 meters per second. That means they have a rapidity of $9.80\text{m s}^{-1}/c$, since for small boosts the rapidity is equal to the velocity (in the appropriate units). There are just over 3×10^7 seconds in a year, so at the end of the year the rapidity is about $3 \times 10^8\text{m s}^{-1}/c$. But since c is about $3 \times 10^8\text{m s}^{-1}$, the rapidity is $\rho = 1$. That's a remarkable coincidence: The earth's surface gravity times the earth's year is about 1, in the appropriate units.

To answer the puzzle: After one year, the ordinary speed is $|v| = \tanh(\rho) = \tanh(1) = 0.8c$.

[After two years, it would be $\tanh(2) = 0.96c$.]

4 Accelerated Reference Frames

One thing about this puzzle that sometimes throws people for a loop is that involves an accelerated reference frame. A key part of the statement is that the passengers feel an acceleration of one Gee. That is an observation made in an accelerated reference frame. Most people have relatively little experience working in accelerated reference frames, and they are justifiably uncertain as to which laws of physics can be relied on in such a frame. Some can, and some can't

Let me coin a scale of complexity or sophistication:

Level 1 := Newtonian mechanics

Level 2 := Special relativity, as it is usually presented.

Level 3 := General relativity

Then I would say that accelerated reference frames are at level 2.1 or some such – more sophisticated than the usual introductory SR discussion, but certainly not requiring the heavy-duty machinery of GR (curved spacetime and all that).

I remember being bothered by accelerated frames for a day or so, back when I was a student. At first, nobody offered an explanation of why SR should work for accelerated reference frames. The discussion consisted of asking “why the heck shouldn't it work” and asserting that there's no reason why it shouldn't. That was completely unconvincing, because there are some laws of physics that do not work in accelerated frames. In particular, the three laws of motion (in their usual basic form) don't work in rotating reference frames; you need to more complex laws including centrifugal effects and Coriolis effects.

Eventually, the real explanation emerged: The trick is to arrange for a succession of *instantaneously comoving* unaccelerated observers. We know SR works for each observer separately. We then arrange to have enough observers at the right places at the right velocities, so that they can observe the action. Afterwards, we collect all their observations and integrate them.

By the correspondence principle, everything (including a modest acceleration) will look Newtonian to the instantaneously comoving observers, so life is easy for them.