Special Relativity

Nic Ford

Mathcamp 2022

Introduction

These are notes from a class called "Special Relativity" that I taught at Canada/USA Mathcamp 2022. The target audience is someone who has seen some high school physics concepts like momentum and kinetic energy but who has had no exposure to special relativity. Mostly there's no math more complicated than high school algebra, although a bit of calculus shows up in the last section. These notes were designed to be somewhat readable on their own, but I make no promises that they'll make nearly as much sense outside the context of that class — some parts stand a good chance of seeming quite unmotivated!

1 Why We Need Special Relativity

1.1 Two Principles in Conflict

I'd like to start by giving a brief history of the ideas leading up to the development of special relativity. This history is going to be something of a cartoon version of the story, since I care more about the logical development of the ideas than about historical accuracy. But my hope is that it will still make what we're about to do feel more motivated than it might otherwise.

The 19th century was an incredibly productive time for physics. So much got done that we don't have time to summarize here, but the part of the story that's going to be the most relevant for us involves the theory of electromagnetism. After many decades of experiments established lots of quantitative facts about the behavior of various electrical and magnetic phenomena, James Clerk Maxwell summarized them all in a unified description of the electric and magnetic fields, what we now call *Maxwell's equations*. One consequence of these equations — and the only one that will concern us for this class — was a beautiful description of light as a *wave* in the electric and magnetic fields. The speed at which such a wave travels can even be extracted directly from the equations and shown to be the quantity we now recognize as the speed of light, $c = 299,792,458 m/s$.

It is possible, therefore, to extract a *fundamental speed* directly from the laws of physics. This is a big problem for another fundamental physical principle: if two observers are moving at a constant velocity relative to each other, they are supposed to see the same laws of physics. Think about what happens if you are sitting in a car and toss a small object into the air — even though the car is moving relative to the Earth, the object looks to you to be travelling straight up and straight down, just as it would if you were outside and standing still.

We can describe this change of coordinates quantitatively. Suppose (x, y, z) are the coordinates that an observer outside the car would use, and (x', y', z') are the coordinates that an observer inside the car would use, and say we decide to make the two coordinate systems agree at time $t = 0$. Then, if the car is travelling in the negative x direction at speed v, we can say

$$x' = x + vt \qquad y' = y \qquad z' = z \qquad t' = t.$$

This coordinate change is sometimes called a **Galileian transformation**, and the principle we're talking about here can be summarized by saying that Galileian transformations are supposed to be symmetries of the laws of physics.

Suppose there were a bird flying outside the car at speed w in the positive x direction. The x coordinate of the bird as a function of time is then given by $x = x_0 + wt$, where x_0 is its position at time $t = 0$, which means that $x' = (x_0 + wt) + vt = x_0 + (v + w)t$. In other words, the Galileian transformation leads us to the unsurprising conclusion that the observer in the car would see the bird travelling at a speed of $v + w$. In particular, the two observers will *disagree* about the speed that ought to be assigned to the bird.

The upshot is that these two physical observations — that the speed of light follows directly of the laws of physics, and that Galileian transformations preserve the laws of physics — are in direct conflict with each other! If instead of a bird we considered a light wave travelling at speed c, then the same logic would lead us to conclude that the observer in the car would measure its speed as $v + c$, which conflicts with the value they would get from Maxwell's equations. There are a couple of ways to resolve a conflict like this:

- **Maxwell's equations are just false.** As far as I know, this option never got that much purchase historically. The problem is that Maxwell's equations don't just capture the same general kind of phenomena as earlier electromagnetic experiments; they make specific, verifiable, quantitative predictions that hold up very well.

- **Maxwell's equations are true, but only in one coordinate system.** This was in fact the dominant hypothesis for quite a while. It's not that hard to see why — every other wave anyone had encountered was a wave in some sort of material substace, so it was easy to imagine that light waves were as well. This substance eventually got the same "luminiferous aether," and it was assumed for quite a while that the theory of electromagnetism should be interpreted as a theory of how the aether behaves.

  The problem with this theory wasn't so much that there was anything wrong with the basic idea, but that more detailed experiments kept requiring more and more complicated descriptions of the aether to account for them. Experiments involving the speed of light through moving water didn't quite agree with the simplest version of the aether theory, which led some theorists to propose "aether drag" theories, in which the water pulls the aether along with it. These had to be refined even further to account for the fact that the apparent amount of aether drag seemed to depend on the frequency of the light, which meant adding more bells and whistles to the theory.

  Finally, the famous *Michelson–Morley experiment* was conducted in 1887 in an attempt to experimentally confirm the existence of the aether. To simplify slightly, the idea was to measure the speed of light at two different times of year, when the Earth was presumably moving at different speeds relative to the aether. After the experiment found no evidence of any such difference, aether theories began to look much less attractive.

  The other possibility is that...

- **The speed of light really *is* the same in every coordinate system.** In particular, this means that the Galileian transformation is *not* an exact symmetry of the laws of physics! This is where we're going to start.

1.2 Our Fundamental Assumptions

So where does this leave us? If we're throwing out the Galileian transformation, we're going to need a replacement for it — in other words, a way to determine how to change coordinates from one observer's coordinate system to another when they are moving at a constant velocity relative to each other. In order to make this more precise, we'll need to talk just a bit more about which coordinate systems we're interested in.

There are some pairs of observers who should *not* be expected to measure the same laws of physics as each other. If, for example, you are sitting on a stool that's spinning around very fast and try to hold up a small weight suspended by a string, you'll find that the weight seems to be pulled away from you. In order to exclude the coordinate systems of observers like this one, we'll restrict our attention to what we call **inertial** coordinate systems. Inertial coordinate systems are the ones in which objects will keep travelling at a constant velocity if they're not being acted on by any outside forces. If two observers are travelling at a constant velocity relative to each other and one of their coordinate systems is inertial, then the other one is too, and every pair of inertial coordinate systems is related in this way.

With this language in hand, we can summarize the fundamental assumptions we're working from by stating that the laws of physics are the same in every inertial coordinate system. In particular, this means that, in every inertial coordinate system, the speed of light will always be measured as $c = 299,792,458 m/s$. Our job now is to use this to derive a formula for moving from one inertial coordinate system to another. That's what we'll start on in the next section.

2 Time Dilation and Length Contraction

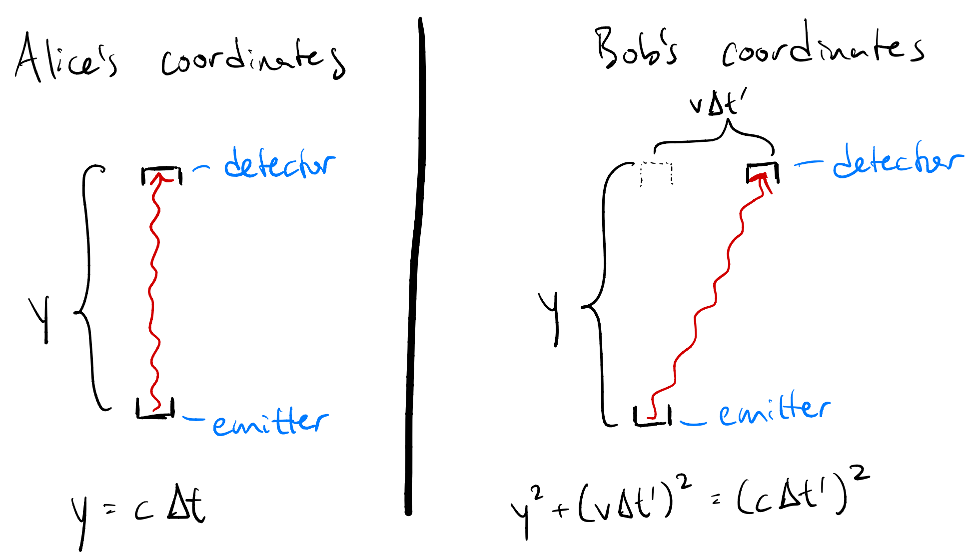
We're ready now to explore the consequences of the fact that the speed of light is the same in every inertial coordinate system. We'll do this by following a classic series of thought experiments which demonstrate all of the phenomena we're going to be interested in.

Imagine a train moving past a platform at speed v in the positive x direction. Alice is sitting on the train performing some physics experiments, and Bob is watching her from the platform outside. We'll use (t, x, y, z) for Alice's coordinate system and (t', x', y', z') for Bob's. (You may be wondering why we're including t alongside x, y, and z in these lists. Remember, though, that we're looking to describe the right formula for changing coordinates between these two coordinate systems, and this coordinate change will depend on time, because the point Alice calls $(0,0,0)$ appears to Bob to be moving to the right.)

2.1 Time Dilation

Alice's first experiment will be to find out how long it takes a flash of light to travel a fixed distance. To do this, she holds up a small device consisting of a light emitter and a light detector. She holds this device vertically, so that the emitter sits at the bottom and fires a flash of light directly upward to the detector. If the distance from the emitter to the detector is y and the light spends time Δt in transit, she will conclude that $y = c\Delta t$.

Now let's consider what Bob sees when he watches this experiment from the platform. Because both coordinate systems are inertial, he will also see the light travel in a straight line. But, because the detector at the top has moved to the right in the time it takes for the light to travel, this line will be slightly diagonal:



Therefore, the distance the light travels in Bob's coordinate system is longer than the distance in Alice's. But, because both of them need to agree about the speed of light, this means they must disagree about how *long* it takes the light to get there! If $\Delta t'$ is the amount of time Bob's coordinate system assigns to this process, then the detector has moved a distace of $v\Delta t'$ while the light is in transit, so we conclude that $c\Delta t' = \sqrt{y^2 + v^2(\Delta t')^2}$.[1]

We can solve for $\Delta t'$ relatively easily:

$$\begin{aligned} c^2(\Delta t')^2 &= y^2 + v^2(\Delta t')^2 \\ &= c^2(\Delta t)^2 + v^2(\Delta t')^2 \\ (c^2 - v^2)(\Delta t')^2 &= c^2(\Delta t)^2 \\ \Delta t' &= \frac{c}{\sqrt{c^2 - v^2}}\Delta t. \end{aligned}$$

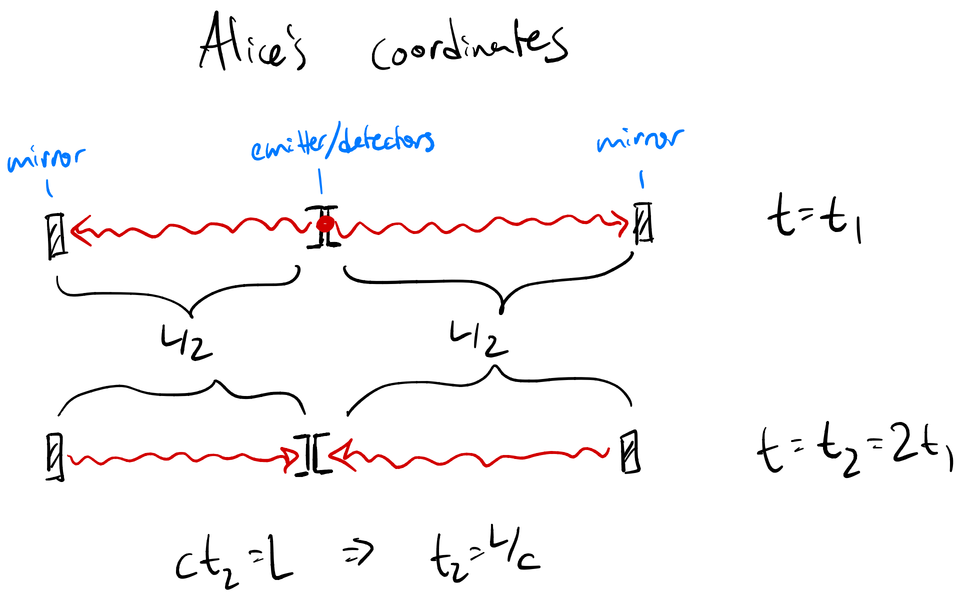
We'll write $\gamma = c/\sqrt{c^2 - v^2}$, so that this formula just becomes $\Delta t' = \gamma \Delta t$. (This quantity will come up a couple more times, so it's nice to give it a name!) This phenomenon, where two inertial coordinate systems disagree about the amount of time passes between two events, is called **time dilation**, and it's the first of the surprising consequences of the constancy of the speed of light we'll see in this section. Notice that if v is very small relative to c, then $c^2 - v^2$ is very close to c^2, which means that γ is very close to 1. So for small speeds — in particular, for the speeds we are likely to encounter in everyday life — the effect of time dilation is very small, so small that it's basically not noticeable at all. As v increases, though, γ does as well, approaching ∞ as $v \to c$. We'll talk more about the physical significance of this in the next section.

2.2 Length Contraction and Relativity of Simultaneity

If our two coordinate systems disagree about the amount of time passes between two events, then a bit of thought should lead you to conclude that they probably ought to disagree about lengths as well — after all, you could measure a length by measuring how long it takes to travel from one point to another at a known speed.

Alice and Bob decide to set up another experiment to measure this. Alice now has two emitters, one pointing to the left and one pointing to the right. She also sets up two mirrors, one to the left of the emitters and one to the right, both at distance $L/2$. At time $t = 0$, both emitters fire out a flash of light. Later, at time $t = t_1$, the light hits both mirrors simultaneously and is reflected back toward the center. Finally, at time $t = t_2$, they both arrive back at the center. Let's say that these emitters also serve as detectors, so that Alice is able to precisely record how long this whole process takes. Each flash of light travels a total distance of L over the course of this experiment, and so we conclude that $t_2 = L/c$.

[1] There is one question you might have about this derivation: why is y the same between these two coordinate systems? There are a couple ways to see this; here's one. If switching from Alice's coordinate system to Bob's changed y, it would either make it larger or smaller. Let's suppose for simplicity that it made it larger. The laws of physics are rotationally symmetric, so if I take this whole system and rotate it 180 degrees around the y axis, nothing ought to change. But if we do this, then it is *Bob* travelling at speed v in the positive x direction relative to *Alice*, and so if we switched back to Alice's coordinate system, symmetry would demand that y has to increase yet again. But of course now we are back where we started, so this is impossible.



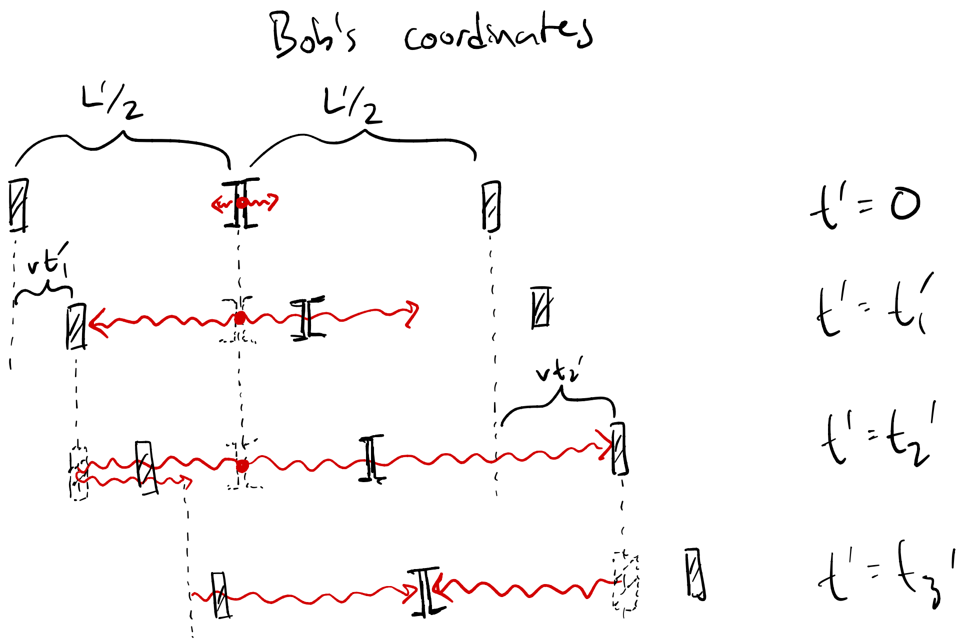
Bob, on the other hand, sees a fairly different sequence of events. Let's suppose that Alice and Bob have synchronized their clocks, so that Bob refers to the moment when the light is emitted as time $t'=0$. We'll write L' for the length that Bob will assign to Alice's machine, and our goal will be to write L' as a function of L and v. Our fundamental assumption means that Bob needs to see both flashes of light travelling at speed c. But, during the time the light is in transit, Bob sees that both mirrors have moved some distance to the right. In particular, then, the light moving to the left will hit its mirror *before* the light moving to the right, *even though Alice's coordinate system had these events happening at the same time*!

This phenomenon is called the **relativity of simultaneity** — two different inertial coordinate systems can disagree about whether spatially separated events happen at the same time. Our experiment to determine how our coordinate change affects lengths appears to have bumped into another phenomenon that needs to be quantified! We can in fact handle both at the same time if we are sufficiently careful about labelling all the events from Bob's perspective.

So let's say that the moment the light hits the left mirror is time $t'=t'_1$. Because the left mirror has travelled a distance of vt'_1 to the right during this process, we conclude that $ct'_1=L'/2-vt'_1$, so that $t'_1=L'/2(c+v)$. A similar computation tells us that, if t'_2 is the moment that the light hits the right mirror, then $t'_2=L/2(c-v)$. Finally, let's write t'_3 for the time both light flashes arrive back at the middle.[2] By adding together the transit times for the left flash to and from its mirror, we can see that

$$ct'_3=\left(\frac{L'}{2}-vt'_1\right)+\left(\frac{L'}{2}+v(t'_3-t'_1)\right).$$

[2] Why must this be the same for both of them? You could deduce this from a computation, but there's a more conceptual way to arrive at the same conclusion. As we've just seen, it's not actually possible to determine whether two spatially separated events are simultaneous by running an experiment that both Alice and Bob's coordinate systems will describe the same way, but it *is* possible to experimentally determine whether two events happen at the same time *and* the same place! Imagine, for example, a light on the detector that turns a different color based on which of the two flashes of light arrives first.



By our earlier discussion about time dilation, though, we have another expression for this amount of time: since Alice's coordinate system has this experiment taking time $t_2 = L/c$, we must also have $t_3' = \gamma t_2 = L/\sqrt{c^2 - v^2}$.

It's now just a matter of some obnoxious algebra to solve for L' in terms of L. The above expression for $c\,t_3'$ becomes (plugging in our expression for t_1' in the middle):

$$\begin{aligned}
c\,t_3' &= L' - 2v\,t_1' + v\,t_3' \\
(c-v)t_3' &= L'\left(1 - \frac{v}{c+v}\right) = L' \cdot \frac{c}{c+v} \\
t_3' &= L' \cdot \frac{c}{c^2 - v^2} \\
\frac{L}{\sqrt{c^2 - v^2}} &= L' \cdot \frac{c}{c^2 - v^2} \\
L' &= \frac{\sqrt{c^2 - v^2}}{c} \cdot L = \frac{L}{\gamma}.
\end{aligned}$$

This final expression, $L' = L/\gamma$, summarizes the phenomenon of **length contraction**: if Alice and Bob measure the length of the same object along the direction in which the train is moving, then Bob's measurement will be shorter than Alice's by a factor of γ.

We can also extract a quantitative version of the relativity of simultaneity from this same experiment. Alice's coordinate system has the two flashes of light hitting the mirrors at the same time, whereas Bob's has them happening at different times $t_2' - t_1'$ apart. Using the expressions for t_1' and t_2' derived above and our new expression for L', we see that

$$\begin{aligned} t_2' - t_1' &= \frac{L'}{2}\left(\frac{1}{c-v} - \frac{1}{c+v}\right) \\ &= \frac{vL'}{c^2 - v^2} \\ &= \frac{vL}{c} \cdot \frac{\sqrt{c^2-v^2}}{c^2-v^2} \\ &= \frac{vL}{c^2} \cdot \frac{c}{\sqrt{c^2-v^2}} \\ &= \frac{\gamma v L}{c^2}. \end{aligned}$$

In other words, if Alice's coordinate system has two events happening simultaneously and separated by a distance L along the x axis, then Bob's coordinate system will describe them as being separated in time by $\gamma vL/c^2$.

Recall that our goal was to come up with the coordinate-change formula that turns Alice's coordinates into Bob's. The three phenomena we've just discussed — time dilation, length contraction, and the relativity of simultaneity — will turn out to be exactly what we need to accomplish this. We'll take up this task in the next section.

Exercises

2.1. I'd like to travel to the star Antares, which is 600 light-years away. How fast do I need to go to get there in three years?

2.2. Very thin cookie dough is moving at $0.5c$. A cookie cutter in the shape of a circular cylinder drops down periodically, cutting cookies out of the dough. Later on, the dough is brought to rest and you inspect the cookies that were cut out during this process. What shape are the cookies?

2.3. Susan is running at relativistic speeds carrying a long ladder pointed in the direction of her motion. Steve, stationary with respect to the Earth, is standing near a barn with open doors on opposite sides. If the doors remain open, Susan (and the ladder) will pass through the barn. If Susan were stationary, the ladder would be too long to fit inside the barn. But due to length contraction, Steve sees Susan and the ladder contracted enough that they can fit inside the barn. When Steve sees that the ladder is entirely inside the barn, he closes the doors and then immediately opens them, trapping Susan and the ladder inside for just an instant.

   (a) According to Susan's coordinate system, the ladder is too long to fit inside the barn. (In fact, this is even more true than it would be if she weren't running — she sees the barn contracted too!) So she should think that what Steve is trying to do is impossible. Who is right? What happens?

   (b) Now suppose that Steve wants to close the doors but not open them afterward. Susan, who is standing near the front of the ladder, stops immediately before the ladder would hit the front door (that is, the last doorway she encounters). What does Susan see? What does Steve see? What happens?

2.4. A stick of length l is sitting above the x axis at an angle θ to the horizontal. If it then starts moving horizontally relative to you at a speed of v, how long does the stick appear in your coordinate system? What angle do you see it forming?

3 Minkowski Space

3.1 The Lorentz Transformation

In the last section, we considered two coordinate systems. The first, corresponding to Alice on the train, was labelled (t, x, y, z); and the second, corresponding to Bob on the platform, was labelled (t', x', y', z'). The train is moving relative to the platform at speed v in the positive x direction, and we wrote $\gamma = c/\sqrt{c^2 - v^2}$. We derived the following three facts:

- If two events happen at the same place in Alice's coordinate system but separated in time by Δt, then Bob's coordinate system has them happening $\gamma\Delta t$ apart. This was called *time dilation*.
- If two events happen at the same time in Alice's coordinate system but separated in space by Δx (and the displacement is along the x axis), then Bob's coordinate system has them separated in time by $\gamma v \Delta x / c^2$. This was called the *relativity of simultaneity*.
- If some object has length L along the x axis in Alice's coordinate system, it has length L/γ in Bob's. (Lengths along the y and z axes are the same.) This was called *length contraction*.

We would like to compile all of this into a coordinate change formula to replace the Galileian transformation we have concluded can't work. It will be easier to write down the formula if we turn the length contraction result into the answer to a slightly different question. If two events happen at the same time in Alice's coordinate system and are separated in space by Δx, how far apart in space does Bob's coordinate system have them happening?

Notice that this is *not* the same question as length contraction. When Bob measures the length of an object, he is looking for two events — say the left side of the object arriving at a certain point and the right side arriving at another point — that are happening at the same time in *his* coordinates. In this question, we are instead considering two events that happen at the same time in *Alice's* coordinates. As the relativity of simultaneity tells us, these are not the same!

One way to do this is to imagine Alice standing on the train holding an object of length Δx horizontally so that its left end arrives at $x = 0$ at time $t = 0$. What we're after is the x' coordinate corresponding to the position of the right end at (according to Alice's coordinates) this same moment. Bob's coordinate system assigns this object the length $\Delta x/\gamma$, but by the relativity of simultaneity, he places the event of the right end's arrival at time $\gamma v \Delta x / c^2$. Its x coordinate is then

$$\Delta x' = \frac{\Delta x}{\gamma} + \frac{\gamma v \Delta x}{c^2} \cdot v = \Delta x \left(\frac{1}{\gamma} + \gamma \cdot \frac{v^2}{c^2} \right),$$

and after some tedious algebra of the sort we did in the last section, we can see that answer to our question is that, in Bob's coordinates, the two events are $\gamma\Delta x$ apart.

To compile this all into a coordinate change formula, let's imagine that Alice and Bob synchronize their coordinates to agree on which point to call $t = x = y = z = 0$. Our formula then looks like this:

$$t' = \gamma t + \frac{\gamma v}{c^2} x \qquad x' = \gamma v t + \gamma x \qquad y' = y \qquad z' = z.$$

The first time of t' comes from considering time dilation of events which all have $x = 0$. The second term comes from the relativity of simultaneity for events which all have $t = 0$. The first term of x' is essentially the definition of what it means for Alice's coordinates to be moving at speed v relative to Bob's: the origin of Alice's coordinates must, after time t' has passed, lie at a point where $x' = vt' = \gamma vt$. Finally, the second term of x' comes from the logic we just went through in the previous paragraph.

This coordinate change is called a **Lorentz transformation**, and it's the fundamental symmetry of special relativity — special relativity can in fact be summarized by the statement that Lorentz transformations preserve the laws of physics.[3] This will be our main object of study for the rest of these notes.

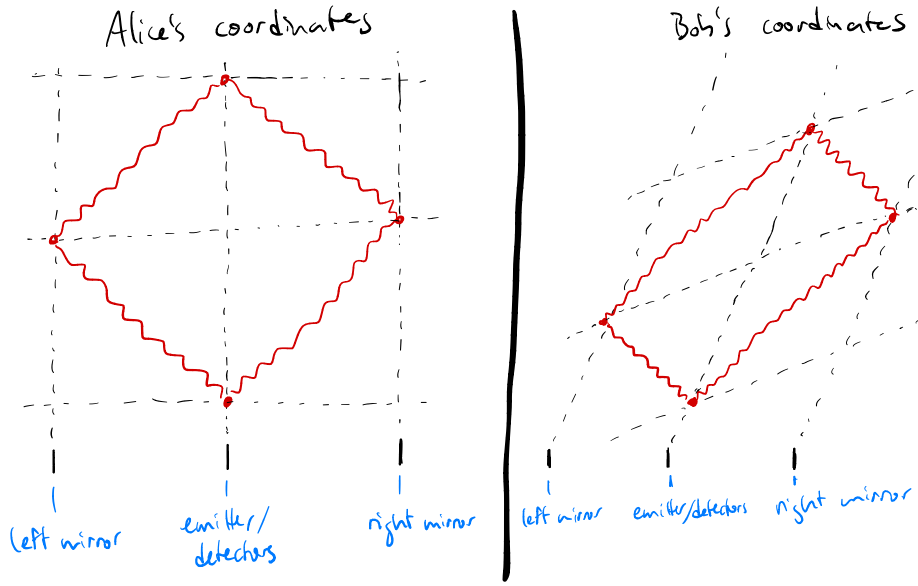
One thing to note about this formula is the relationship between v, c, and γ. As we discussed briefly in the last section, γ takes its minimum value of 1 at $v = 0$, and goes to ∞ as $v \to c$. This means that the Lorentz transformation is only defined when $|v| < c$ — if $|v| = c$ we would be dividing by zero and if $|v| > c$ then the formula for γ involves the square root of a negative number. The physical interpretation of this is quite striking: since we are interpreting the Lorentz transformation as the coordinate change formula between two coordinate systems moving at speed v relative to each other, we conclude that there is *no such thing* as a pair of objects whose relative speed is faster than the speed of light! We'll talk more about what this means in the next section.

3.2 Spacetime

It can be very helpful to imagine drawing pictures of scenarios like the ones we talked about yesterday in $\mathbb{R}^4$, where we think of the four coordinates as t, x, y, and z. This copy of $\mathbb{R}^4$ is called **Minkowiski space** or **spacetime**, and points in Minkowski space are often called **four-vectors** to distinguish them from the purely spatial **three-vectors** in $\mathbb{R}^3$. We often write the coordinates of a four-vector as (ct, x, y, z) so that they all have the same units.

Here, for example, is a depiction of the second experiment from the last section — the one with the horizontal light apparatus — in both Alice and Bob's coordinate systems. We show only the x and ct axes, with x running horizontally and ct running vertically. (Note that, because we are using ct to label the time axis, objects travelling at the speed of light correspond to lines of slope 1.)

[3] What about coordinate systems that are moving relative to each other but not along the x axis? You can extend the formula for the Lorentz transformation to handle this case with a bit of extra work, or you can simply rotate until the direction points along the x axis, perform this version of the Lorentz transformation, and then rotate back. The term "Lorentz transformation" generally refers to this broader class of coordinate changes, not just the ones along the x axis.



The path that a single object follows in a diagram like this is called the object's *world-line*. Just as rotations in ordinary Euclidean space preserve lengths and angles, there is a quantity that is similarly preserved by Lorentz transformations: you will check in Exercise 3.4 that, if two points in Minkowski space are separated from each other by $\Delta w = (c\Delta t, \Delta x, \Delta y, \Delta z)$, then the quantity

$$(c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2$$

is preserved by Lorentz transformations. I encourage you to convince yourself that this quantity is also preserved by space translations, time translations, and rotations, which are in fact all of the symmetries of spacetime.

We'll call this quantity the **spacetime norm** of the four-vector Δw, and we'll denote it by $N(\Delta w)$.[4] Unlike distances in ordinary Euclidean space, the spacetime norm can be negative or zero. The best way to interpret it depends on this sign:

- First, suppose $N(\Delta w)$ is positive. By performing an appropriate rotation (which doesn't affect the spacetime norm) we can assume that Δy and Δz are both zero. Now, the fact that the spacetime norm is positive implies that $c\Delta t > \Delta x$, that is, $|\Delta x/\Delta t| < c$. So, if we then perform a Lorentz transformation with $v = -\Delta x/\Delta t$, we'll end up with $\Delta x = 0$ as well.

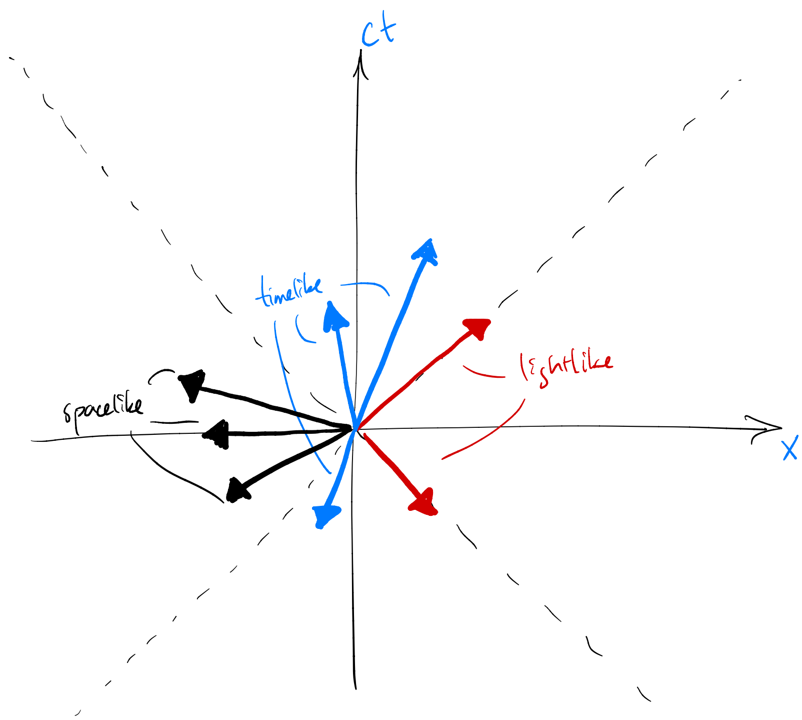
  The spacetime norm — which, again, is the same as the spacetime norm of our original four-vector — is now just $c^2(\Delta t)^2$, from which we can extract our interpretation: if $N(\Delta w)$ is positive, then you should interpret it as (c^2 times the square of) the amount of time that passes in the coordinate system of a particle that moves in a straight line from the first point to the second. In this scenario, we call Δw **timelike**.

- If $N(\Delta w)$ is negative, on the other hand, you will check in Exercise 3.2 that this means we can make $\Delta t = 0$ with an appropriate Lorentz transformation. The spacetime norm is then just $-((\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2)$, and you should therefore interpret it as (the square

[4] This notation is nonstandard. The more common name for this quantity is *spacetime interval*, and it's common to write it as $(\Delta w)^2$. I find both of these conventions confusing, so I made up a new one for these notes.

of) the distance between the two points according to a coordinate system that has them happening at the same time. In this scenario, we call Δw **spacelike**.

- Finally, if $N(\Delta w)$ is zero (and the original two points aren't equal) then it is not possible to perform a Lorentz transformation to make the space coordinates or the time coordinate vanish. In this scenario, we call Δw **lightlike**, because it represents a path that light might travel. In particular, this means that there is no coordinate system in which light is at rest, a natural consequence of the fact that the speed of light is the same in every coordinate system! We'll talk about some of the consequences of this in the next section.



Exercises

3.1. Prove the *velocity addition formulas*:

(a) Suppose, in some coordinate system, that some object is moving at speed u. If I'm moving at speed v in the opposite direction, prove that the speed I see the object moving at is

$$\frac{u+v}{1+uv/c^2}.$$

(b) Now suppose that I'm instead moving at speed v in a direction perpendicular to the object. How fast do I see it moving?

(c) Now suppose I'm moving at speed v in the x direction, and the x and y components of the object's velocity in the first coordinate system are u_x and u_y. Find a formula for what I think the x and y components of the object's velocity are.

3.2. Prove that if the spacetime norm of a four-vector (ct, x, y, z) is negative, then there is a Lorentz transformation we can apply to make $t = 0$.

3.3. It's tempting to say that length contraction causes an object moving relative to you to look shorter than it would at rest, but this ignores an important fact: the way you see things is by detecting the light that comes from the object, and the image you see is made up of light rays coming from different parts of it arriving at your eye at the same time.

With this in mind, imagine a thin rod moving directly toward you, pointing along the direction of its motion. How does the length you see the rod having compare to its length measured at rest? What if the rod is moving away from you instead? (If it helps to visualize the situation, imagine you're standing just slightly off to the side, so you can see the length of it and not just the tip.)

3.4. Minkowski space comes with an **inner product**, very similar to the dot product of vectors you might be used to. The inner product of two four-vectors is:

$$(ct, x, y, z) \cdot (ct', x', y', z') = c^2 t t' - x x' - y y' - z z'.$$

(a) Prove that this inner product is **Lorentz-invariant**, that is, if I apply the same Lorentz transformation to two vectors, it doesn't change their inner product.

(b) Conclude from this that the spacetime norm is preserved by Lorentz transformations.

3.5. Imagine you have access to a device that is capable of reliably sending a signal faster than light, say at $1.2c$. This device can be replicated as many times as necessary, and it is portable. If you also have access to spaceships that can travel at any speed *slower* than light, devise a way to send a signal back in time. What are the limitations of your method?

4 Mass, Energy, and Momentum

We're going to close these notes by saying just a bit about how to work with a few basic physics concepts in special relativity. We'll be guided by one unifying principle: because the Lorentz transformation mixes space and time coordinates and it's supposed to be a symmetry of the laws of physics, four-vectors (rather than three-vectors) are the appropriate mathematical object to use to represent physically meaningful quantities.

4.1 Four-velocity

Let's start by applying this to the concept of velocity. Let's suppose that, according to some inertial coordinate system, some particle is moving with constant three-velocity (v_x, v_y, v_z). If t is the time coordinate of our coordinate system, then this means that (for example) the x coordinate of our particle will change by $v_x \Delta t$ whenever t increases by Δt, that is, $v_x = \Delta x / \Delta t$. This is just not going to be a workable definition of velocity for other coordinate systems, because other coordinate systems will disagree about how much time has passed! We're going to need something better.

One way to solve this problem is to use a time coordinate that every coordinate system will agree on, and the best candidate in sight for this is the time coordinate of the coordinate system of the particle itself. We will call this the particle's **proper time**, and denote it by τ. By the interpretation we gave to timelike spacetime norms in the last section, we can see that, if Δw is some four-vector, then $\Delta\tau = \sqrt{N(\Delta w)}/c$. The particle's **four-velocity** will then be defined as

$$u := \left(c\frac{\Delta t}{\Delta \tau}, \frac{\Delta x}{\Delta \tau}, \frac{\Delta y}{\Delta \tau}, \frac{\Delta z}{\Delta \tau} \right).$$

(Notice that, now that we're using τ as our "clock" rather than t, the time coordinate is also nontrivial!)

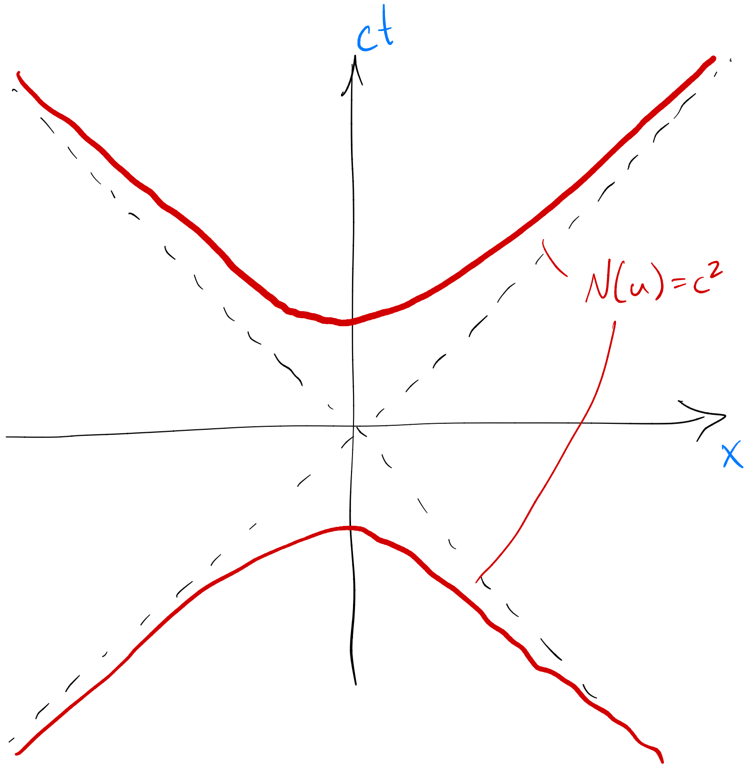
What does this look like in terms of our original three-velocity v? One way to answer this is to start from the particle's coordinate system and perform the right series of transformations until we get to the coordinate system in which the particle is moving with three-velocity v. In the particle's coordinate system, x, y, and z never change, and t is the same as proper time, so the four-velocity is just $(c, 0, 0, 0)$. Now, let's perform a Lorentz transformation in the x direction to a coordinate system in which the particle is moving at speed $|v|$ in the x direction. With the c's inserted on the t coordinates, the formula for the Lorentz transformation looks like

$$(ct)' = \gamma(ct) + \frac{\gamma v}{c^2} x \qquad x' = \frac{\gamma v}{c}(ct) + \gamma x \qquad y' = y \qquad z' = z,$$

so our four-velocity becomes $(\gamma c, \gamma|v|, 0, 0)$. Finally, we can rotate this four-vector in the spatial coordinates, which gives us

$$u = (\gamma c, \gamma v_x, \gamma v_y, \gamma y_z).$$

Because this whole procedure preserves the spacetime norm, we see that the four-velocity of a particle always has spacetime norm c^2.



It's pretty straightforward to generalize this to particles that aren't moving at constant velocity.[5] Imagine that our particle is travelling along some path $w(s)=(ct(s),x(s),y(s),z(s))$ in spacetime. (We will, for the moment, not assume anything about the parameter s that we're using, although some restrictions on the path as a whole will appear momentarily.) We'd like to say that the four-velocity is $dw/d\tau$. This is not necessarily the same as dw/ds — the parameter s doesn't have to have anything to do with proper time.

We can determine the relationship between s and τ by considering a small segment of the path, say from s to $s+\Delta s$. The time that will elapse over the course of this segment according to the particle's coordinate system is $\Delta\tau = \frac{1}{c}\sqrt{N(w(s+\Delta s)-w(s))}$, and so, dividing through by Δs, we can conclude that $d\tau/ds = \frac{1}{c}\sqrt{N(dw/ds)}$. We can therefore say that the four-velocity is

$$\frac{dw}{d\tau} = \frac{dw}{ds}\cdot\frac{ds}{d\tau} = \frac{c\frac{dw}{ds}}{\sqrt{N\left(\frac{dw}{ds}\right)}}.$$

This expression works perfectly fine so long as $N(dw/ds)$ is positive — that is, so long as dw/ds is timelike. Paths w for which this is not true are therefore not paths that it is possible for a particle to take. (Or at least dw/ds can't be spacelike; we'll discuss the lightlike case momentarily.) This corresponds to the prohibition we have already seen against particles travelling faster than light.[6]

[5] There is a common misconception that special relativity can only handle particles moving at constant velocity, and that to describe accelerating particles you need general relativity. This is, in fact, completely false — general relativity is a theory of gravity, but there is nothing about special relativity that precludes talking about particles accelerating due to non-gravitational forces.

[6] There is one more restriction on the possible paths w that make sense for a particle to travel along: we would like it

We can also use our expression for $d\tau/ds$ to determine how much time elapses in the particle's coordinate system over any segment of the path. If we integrate our expression for $d\tau/ds$ over the segment of the path from $s = s_0$ to $s = s_1$, we get

$$\tau_1 - \tau_0 = \frac{1}{c}\int_{s_0}^{s_1} \sqrt{N\left(\frac{dw}{ds}\right)}\,ds.$$

4.2 Energy-momentum

Consider a particle moving in some coordinate system with four-velocity $u = (\gamma c, \gamma v_x, \gamma v_y, \gamma v_z)$. What should be the relativistic analogue of its momentum? The naive answer would be to simply multiply the four-velocity by the mass m of the particle, giving

$$p = (\gamma mc, \gamma m v_x, \gamma m v_y, \gamma m v_c).$$

Let's see what happens when we try this.

If v is much smaller than c, then, as we've said before, γ is very close to 1, and so the spatial components of p will be close to (mv_x, mv_y, mv_z), the coordinates of nonrelativistic momentum. So far, so good! What about the time component? For this, it will be helpful to use the power series for $1/\sqrt{1-x^2}$, which begins

$$\frac{1}{\sqrt{1-x^2}} = 1 + \frac{1}{2}x^2 + \frac{3}{8}x^4 + \cdots.$$

Using the fact that $\gamma = c/\sqrt{c^2 - v^2} = 1/\sqrt{1-(v/c)^2}$, we get that

$$p_t = mc \cdot \frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}} = mc\left(1 + \frac{1}{2}\frac{v^2}{c^2} + \cdots\right) = \frac{1}{c}\left(mc^2 + \frac{1}{2}mv^2 + \cdots\right).$$

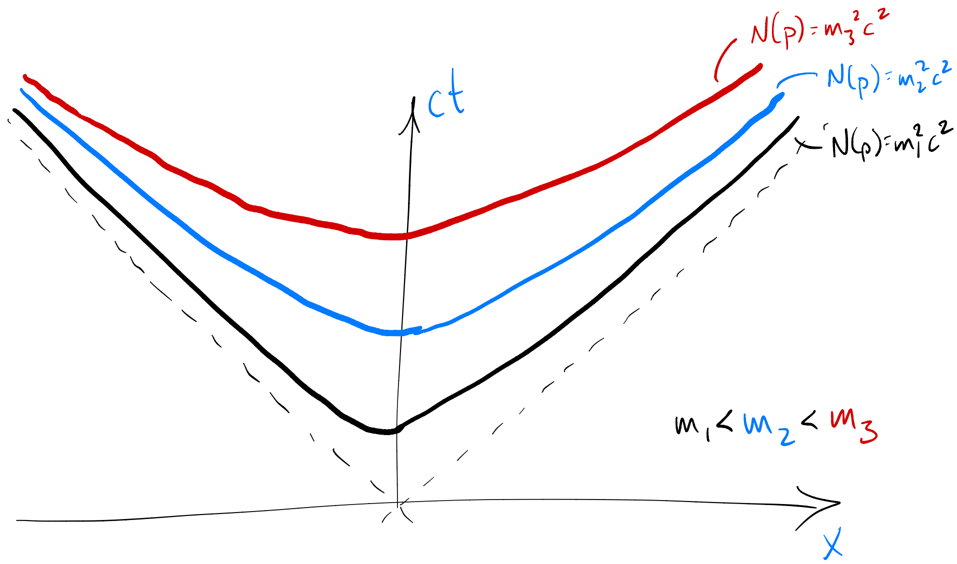
There are a couple of interesting things we can see from this expression. First, the second term is exactly the nonrelativistic expression for the kinetic energy of a particle. This strongly suggests that cp_t ought to be interpreted as an energy, and (for this reason and many other reasons) this is exactly what we are going to do. The first term, mc^2, doesn't depend on the velocity of the particle, but if we are going to interpret cp_t as the particle's energy, then this first term seems to give a contribution to the energy purely from the particle's mass. This quantity is referred to as the particle's **rest energy**. This is the origin of the famous equation $E = mc^2$, although, since that equation actually only holds in a coordinate system in which the particle is at rest, it might be more correct to say $E = \gamma mc^2$ or $\sqrt{E^2 - c^2(p_x^2 + p_y^2 + p_z^2)} = mc^2$.[7]

The four-vector p is referred to as **energy-momentum** or **four-momentum** and, just as the first name suggests, it is the relativistic analogue of *both* energy and momentum. Just like energy and momentum in nonrelativistic physics, energy-momentum is a conserved quantity in special

to point forward rather than backward in time! Taking the negative of a vector preserves its spacetime norm, so just requiring the four-velocity to be timelike isn't quite enough to make this happen. I encourage you to check that Lorentz transformations take forward-pointing timelike vectors to forward-pointing timelike vectors, so this restriction is the same in every inertial coordinate system.

[7] There is an older convention, which I don't especially like, of using the word "mass" to refer to p_t/c, even in coordinate systems where the particle isn't at rest. This convention seems to have mostly fallen out of use, and I personally dislike it — it means that the mass of a particle changes depending on its speed. We are going to use "mass" to refer to the constant quantity I've called m here, and use "energy" to refer to the quantity that changes.

relativity — that is, the total energy-momentum of a system can't change except under the influence of an external force. Perhaps surprisingly, though, mass by itself is *not* conserved. Both mass and kinetic energy contribute to the total energy of a particle, but there definitely are interactions in nature that exchange one of these forms of energy for the other.



Just as four-velocities were required to satisfy the equation $N(u) = c^2$, the four-momentum of a particle of mass m has to satisfy the equation $N(p) = m^2c^2$, or

$$E^2 - c^2(p_x^2 + p_y^2 + p_z^2) = m^2c^4.$$

I encourage you to also verify that, if $p_x = \gamma m v_x$ is the x component of energy-momentum, then $p_x c = E v_x / c$.

What does this mean if $m = 0$? The definition $p = (\gamma mc, \gamma m v_x, \gamma m v_y, \gamma m v_z)$ is now obviously broken.[8] But we can still look at the two relations we just proved in the $m > 0$ case and see what they would imply about the energy-momentum of a massless particle. Applying the two equations in order, we see that

$$E^2 = c^2(p_x^2 + p_y^2 + p_z^2) = E^2 \frac{v_x^2 + v_y^2 + v_z^2}{c^2}.$$

At least if you accept the assumptions underlying this whole exercise, this implies that $|v|^2 = c^2$, that is, every massless particle must always be travelling at the speed of light! (I encourage you to check that this works in the other direction too.)

This is in fact the way massless particles are modelled in special relativity. In particular, even when a particle doesn't have mass, it can have energy and momentum. (Indeed, we know light itself has to have energy — exposing things to light can make them hotter, and I get a bill every month in part for the energy it takes to light my apartment.) Unlike for massive particles, though, just knowing the mass and velocity of a massless particle doesn't tell you its energy-momentum — indeed, if p is the energy-momentum of a massless particle travelling in some fixed direction, then so is any positive multiple of p.

[8] Indeed, our derivation of the expression for four-velocity, which was used to write this expression for p, relied on starting in a coordinate system in which the particle is at rest. But, as we will see momentarily, if $m = 0$ then there *is* no coordinate system where the particle is at rest!

Exercises

4.1. Recall that, in Exercise 3.4, we defined an inner product on four-vectors and showed in was Lorentz-invariant. Consider two objects, a travelling with the four-velocity v, and b travelling with the four-velocity w. Prove that $v \cdot w = \gamma c^2$, where γ is the Lorentz factor I would use to write a's velocity in b's coordinate system, or to write b's velocity in a's coordinate system. *[Hint: Imagine one of them points along the world-line of a particle and imagine performing a Lorentz transformation to a frame in which this particle is at rest.]*

4.2. Suppose two objects with masses m_1 and m_2, travelling directly toward each other with relative velocity v, collide and stick together. What is the mass of the resulting object?

4.3. A rocket ship operates by converting some of its mass into photons and propelling them out of the back of the ship. How much of the mass needs to be converted to take the ship from rest to $0.5c$?

4.4. An electron at rest relative to you is hit by a moving electron, and after the collision there are three electons and one positron moving at some undetermined velocities. What is the minimum amount of kinetic energy, in your coordinate system, that the moving electron must have in order for this process to be able to occur? (The mass of a positron is the same as the mass of an electron.)