APPENDIX A: SPECIAL RELATIVITY

MICHEL JANSSEN

1. The relativity postulate, the light postulate, and their strange consequences

1.1. **The two postulates of special relativity and the tension between them.** When Einstein first presented what came to be known as special relativity, he based the theory on two postulates or principles, called the "relativity postulate" or "relativity principle" and the "light postulate." Both postulates are supported by a wealth of experimental evidence. The combination of the two, however, appears to lead to contradictions. To avoid such contradictions, Einstein argued, we need to change some of our fundamental ideas about space and time.

Einstein formulated the relativity postulate as follows: "The same laws of electrodynamics and optics will be valid for all frames of reference for which the equations of mechanics hold good" (Einstein 1905r, 891). Such frames of reference are called inertial frames and an observer at rest in one of them is called an *inertial observer*. A few examples will suffice to understand both the concept of an inertial frame and the meaning of the relativity postulate. First consider a plane which starts out sitting on the tarmac, proceeds to fly through clear skies, and eventually hits turbulence. All the while a passenger is nursing a cup of coffee. Sipping coffee without spilling is easy during the smooth portion of the flight. This is because the laws governing the behavior of the coffee in the frame of reference of the plane flying at constant velocity are the same as in the frame of reference of the airport.[1] In fact, these same laws hold in any frame moving uniformly (i.e., with constant velocity) with respect to the frame of the airport. Drinking coffee without spilling when the plane ride gets bumpy is much harder. The laws for the coffee in non-inertial frames, such as the frame of a plane encountering turbulence, are more complicated than in inertial frames. As a second example consider a cruise ship that sets out from its port of origin, sails smoothly on a calm sea, and eventually is caught in a storm. All the while two passengers engage in a drawn-out tennis match on the ship's upper deck. No matter whether the ship lies anchored in the harbor or is sailing on calm seas, the balls will bounce and spin according to the same physical laws. Playing tennis on the deck of a cruise ship is no different from playing tennis on any other court. Once the ship hits bad weather, however, the laws governing the motion of the balls will get more complicated and the players will have to start adjusting their strokes accordingly.[2]

For purely mechanical phenomena such as the bounce of balls or the flow of fluids, the relativity principle had been known since the days of Galilei. For reasons that will become clear below, 19th-century physicists expected the principle to break down for electromagnetic and optical phenomena. But no violations were ever found. Einstein could thus extend the principle to all of physics, in particular to electromagnetism and optics.

To the relativity postulate Einstein added the light postulate: “light propagates through empty space with a definite velocity [c, about 186,000 miles per second or 669,600,000 miles per hour] which is independent of the state of motion of the emitting body” (Einstein 1905r, 891). That the velocity of light is independent of the velocity of its source is one of the key features of the electromagnetic theory developed in the second half of the 19th century by Maxwell, Lorentz, and others. The light postulate is thus indirectly supported by all the evidence amassed during the 19th century for this powerful theory.

When Einstein introduced his second postulate, he immediately warned his readers that it is “apparently irreconcilable” (ibid.) with the first. That the two postulates would seem to be incompatible with one another is not difficult to see; the hard part is to see that this incompatibility is only apparent.

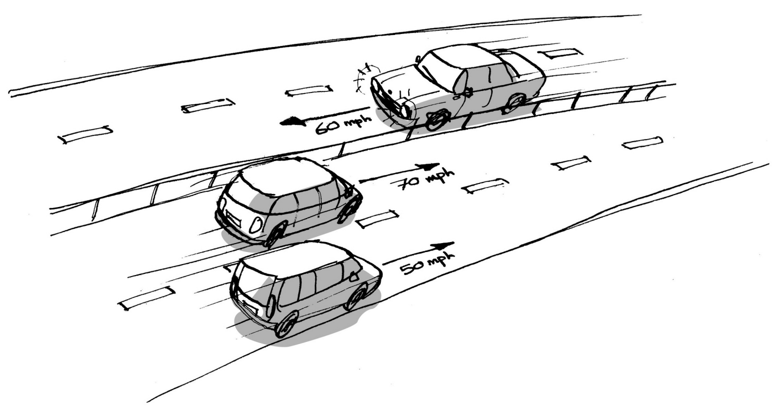

FIGURE 1. The postulates and their strange consequences for the behavior of light. *This and all other diagrams in the appendix were drawn by Laurent Taudin.*

Consider the situation in Fig. 1, showing two SUVs and a sedan driving down the highway. The sedan is moving at 60 m.p.h. in one direction, the SUVs are moving in the other direction, one at 50 m.p.h., the other at 70 m.p.h. Common sense tells us that the drivers of the SUVs will give different answers when asked how fast the sedan is approaching them. The first driver, going 50 m.p.h., will say 110 m.p.h.; the second one, going 70 m.p.h., 130 m.p.h. Now ask both drivers how fast they think the light from the headlights of the sedan is coming toward them? The answers will depend on whether the drivers think of light as waves in some medium, like sound in air, or as particles, emitted like bullets from a gun. If light consists of waves in a medium, the velocity of light is the velocity with which these waves propagate through the medium. If light consists of particles, the velocity of light is the velocity with which these particles are emitted from their source.

If the drivers of the SUVs think of light as particles, they have to take into account both the velocity of their SUVs and the velocity of the sedan. The velocity of the sedan, 60 m.p.h., will be a component of the velocity of the light particles emitted by its headlights. For the drivers of the SUVs approaching the light their own velocity will be another component of the light's velocity. Hence, the driver going 50 will say that the light is approaching him with $c + 110$ m.p.h., while the one going 70 will say that the light is approaching him with $c + 130$ m.p.h.

If the drivers of the SUVs think of light as a wave in medium, they have to take into account the velocity of the medium, the velocity of the SUVs, but not the velocity of the sedan. It is a general property of waves in a medium that their velocity is independent of the velocity of their source. The ripples spreading across a pond from the point where a rock hit the water travel at the same speed regardless of whether you skipped the rock on the water or lobbed it into the pond at that point. Likewise, the velocity of the sedan will not be a component of the velocity of the light waves leaving its headlights. The velocity of the medium, however, will be a component of the velocity of the waves. A wave is a disturbance of the medium and is carried along with it. Assuming the medium for the light waves to be at rest with respect to the highway, the driver going 50 m.p.h. will say that the medium is approaching him with 50 m.p.h. and the light with $c + 50$ m.p.h., while the other one will say that the medium is approaching him with 70 m.p.h. and the light with $c + 70$ m.p.h.

If Einstein's postulates are correct, both sets of answers are wrong. The relativity postulate says that the laws of physics in the frame of reference of any one of the cars shown in Fig. 1 are the same as in the frame of reference of the highway. The light postulate is an example of such a law. The light postulate thus holds in the frame of reference of both SUVs. The velocity of the sedan in the frame of the SUV going 50 m.p.h. is different from its velocity in the frame of the SUV going 70 m.p.h., but according to the light postulate the velocity of the sedan does not affect the velocity of the light coming from its headlights. According to the postulates of special relativity, both SUV drivers will find that the light from the sedan's headlights, or any other light for that matter, approaches them at the same speed c!

This example offers some insight into why the combination of the two postulates seems to lead to contradictions. The light postulate suggests that light is a wave in some medium. This is how light is pictured in the electrodynamics of Maxwell and Lorentz. They saw light as an electromagnetic wave in the ether, a substance thought to fill all of space with no internal motion like a perfectly calm sea filling every nook and cranny of the universe. This is why adherents of this theory expected the relativity principle to break down. An observer in uniform motion through the ether is not equivalent to an observer at rest in the ether. A wave in the ether only moves with the same constant velocity in all directions with respect to the latter. Yet experiment suggested that the relativity principle holds for light as well. The relativity principle is compatible with light consisting of particles moving through empty space. But then the velocity of light depends on the velocity of the source from which the light particles are emitted, which contradicts the light postulate. If we insist on having both the relativity and the light postulate, light, it seems, can neither be a particle nor a wave.

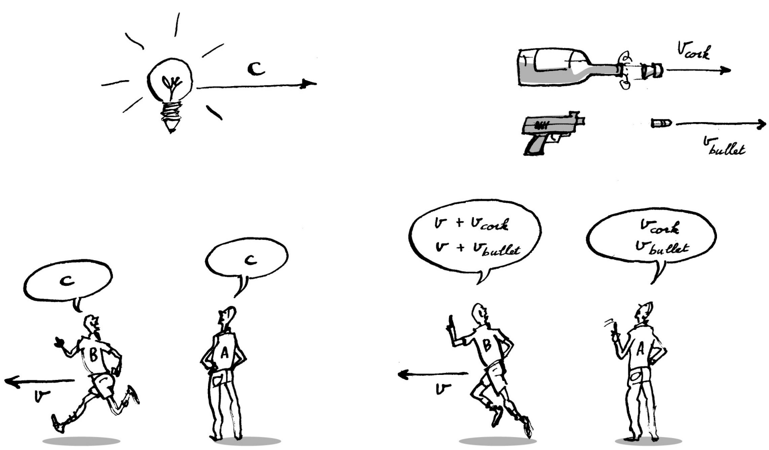

FIGURE 2. The postulates and adding velocities.

Fig. 2 illustrates the problem of combining Einstein's two postulates in a slightly different way. We have two observers, Al and Bob, moving with respect to one another (like the drivers of the SUVs in Fig. 1). They examine a flash of light emitted by a light bulb, a cork shooting out of a champagne bottle, and a bullet fired from a gun. Common sense tells us that the light, the cork, and the bullet will have different velocities for Al and Bob. Special relativity confirms this in the cases of the cork and the bullet, even though, as we shall see, it calls for corrections of the common-sense values $v + v_{cork}$ and $v + v_{bullet}$ shown in Fig. 2. In the case of light, however, it is a direct consequence of Einstein's two postulates that Al and Bob, despite being in motion with respect to one another, register the exact same velocity! How can this possibly be? It is clear that something will have to give.

What Einstein showed in his 1905 paper is that once we accept his two postulates—as we should given all the empirical evidence backing them up—we have to give up some of our common-sense ideas about space and time. Now it is one thing to concede when confronted with the relentless logic of Einstein's 1905 paper that our old ideas were indeed nothing but prejudices; getting comfortable with the new ideas that Einstein put in their place is a different matter. Here the 1905 paper is of little help. It does not tell us how to visualize the new relativistic ideas about space and time. Such visualization was provided a few years later by Minkowski (1909). Minkowski's geometrical formulation of special relativity is still standard today. It took Einstein several years to appreciate this contribution (Pais 1982, 152). In Section 2 we shall turn to Minkowski's geometry of relativistic space-time. In the remainder of this section we stay true to the spirit of Einstein's more abstract approach.

In a nutshell, the argument in Sections 1.2–1.5 is as follows. A direct consequence of the two postulates[3] is that two inertial observers moving with respect to one another will disagree on whether events at different locations happen at the same time or not. This phenomenon is called the *relativity of simultaneity* (Section 1.2). Judgments about the simultaneity of events at different locations are involved in measuring the rate of moving clocks and the length of moving rods. As a consequence, we find that moving clocks must tick at a lower rate than those same clocks at rest (Section 1.3) and that the length of moving rods must be less than the length of those same rods at rest (Section 1.4). Otherwise we end up with violations of the relativity principle. What makes these phenomena—called *time dilation* and *length contraction*, respectively—especially baffling is that which clocks and which rods are moving and which ones are at rest depends on whose point of view we adopt. Two observers in relative motion will both claim that the other observer's rods are contracted and that the other observer's clocks run slow. As we shall see in Section 2, Minkowski's geometrical interpretation of length contraction and time dilation makes this much easier to understand. We conclude Section 1 by examining the consequences of these new and unexpected phenomena for the addition of velocities (Section 1.5). This will resolve the apparent contradictions found above.

Qualitatively, this resolution goes as follows. Once again consider Fig. 2. According to Al, the cork, the bullet, and the light flash have velocities v_{cork}, v_{bullet}, and c, respectively. According to Bob, however, Al has determined these velocities using rods that are contracted and clocks that are not properly synchronized and are running slow to boot. If Bob wants to know how fast the objects are moving with respect to him given how fast they are moving with respect to Al, he cannot simply add the velocities of the objects reported by Al to Al's own velocity v. He first needs to correct Al's results. The general rule for adding velocities measured by two observers in relative motion that takes into account such corrections is called the *relativistic addition theorem for velocities*. In the case of light, it turns out, the corrected value for the velocity reported by Al is $c - v$ (see note 15). Adding Al's own velocity to this corrected value, Bob finds that the light is moving with velocity with respect to him as well. In other words, even though they are in motion with respect to one another, both observers find that one and the same light flash is moving with velocity with respect to them. This takes care of the problem brought out with the help of Fig. 2.

The analysis of Fig. 1 suggested that light could neither be a wave nor a particle if we accept Einstein's two postulates. It turns out, however, that special relativity is compatible with both views. Picture light as a wave in the ether of 19th-century physics. For an observer at rest in the ether, the light will have velocity c. Because of the way velocities are added in special relativity, however, it will have the exact same velocity for any observer in uniform motion through the ether. This means that it is impossible to tell with respect to which observer the ether is truly at rest. For this and other reasons, Einstein preferred to do away with the ether altogether, calling it "superfluous" in the introduction of his 1905 paper. Like Einstein, we now think of electromagnetic waves as propagating through empty space rather than through an ether. If we picture light as a wave, with or without the ether, the light coming from the headlights of the sedan in Fig. 1 has velocity c with respect to the drivers of both SUVs, independently of the velocity of the sedan. If we picture light as a particle emitted with velocity c by the headlights of the sedan, we need to add the velocity of the sedan to the velocity of the light. But because of the way velocities are added in special relativity, the result will once again be that these particles move with velocity c with respect to the two SUVs as well. Before we factored in the relativity of simultaneity, length contraction, and time dilation, the wave picture of light and the particle picture of light gave different answers for the velocities with which light from the sedan's headlights is approaching the two SUVs. Once these three phenomena are taken into the account both pictures give the same result, which, moreover, is precisely the result we found on the basis of a direct application of the postulates. Both the wave and the particle picture are thus compatible with the postulates. Special relativity is agnostic about the nature of light.

1.2. **The relativity of simultaneity.** Whether two events occurring at different locations happen simultaneously or one after the other depends on the state of motion of the person making the call. This is the key insight of special relativity. It dawned on Einstein about six weeks before he published his theory and made everything fall into place.[4]

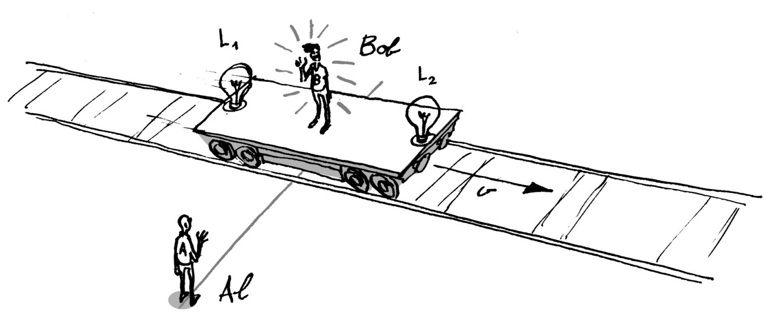

FIGURE 3. Two light flashes reach Bob at the same time.

Consider Fig. 3. Bob is standing exactly in the middle of a railroad car moving to the right at constant velocity v. Al is standing by the tracks. To both ends of the railroad car a light bulb is attached, labeled L_1 and L_2, respectively. Both light bulbs flash once. At exactly the moment that Al and Bob come face to face with one another (see the solid line connecting them), the two flashes reach Bob. Al and Bob agree that these two flashes hit Bob at the same time. They will always agree on what happens at one and the same location at one and the same instant. They will not agree, however, on whether the light bulbs flashed at the same time. The flashing of L_1 and the flashing of L_2 are events happening at different locations. Appealing to Einstein's two postulates, Bob will say that these two events happened simultaneously. Appealing to the same postulates, Al will say that L_1 flashed before L_2.

Bob offers the following impeccable argument: "I am an inertial observer, so the laws of nature are the same for me as they are for any other inertial observer. The light postulate is one such law. Hence, the light flashes from L_1 and L_2 have velocity c with respect to me. I am standing halfway between L_1 and L_2. Hence, the light flashes from L_1 and L_2 had the same distance to cover to get to me. They hit me at the same time. *Ergo*, they must have left L_1 and L_2 a little earlier at the same time."

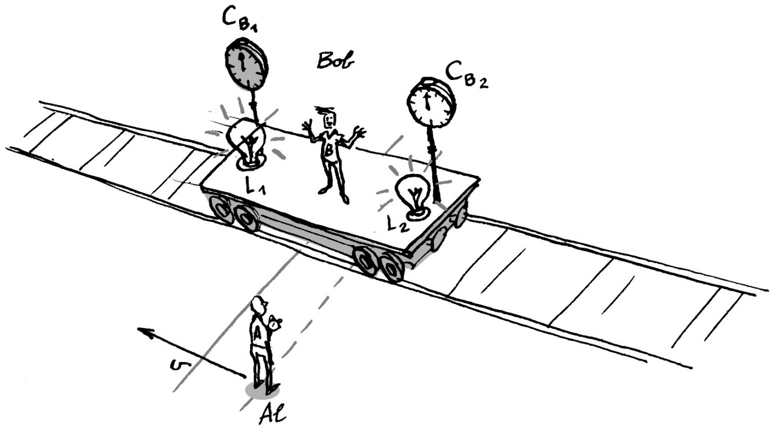

Figure 4. According to Bob, light bulbs L_1 and L_2 flashed at the same time.

Fig. 4 illustrates, from Bob's perspective, the moment that L_1 and L_2 flashed. From this perspective, Bob and the railroad car are at rest while Al is moving to the left with velocity v. At this point Al is slightly to the right of the solid line marking the point where he will come face to face with Bob a split-second later, when the two flashes hit Bob (see Fig. 3).[5] The figure also shows two clocks, labeled C_{B_1} and C_{B_2}, fastened to the railroad car at the positions of the light bulbs. If these clocks are properly synchronized according to Bob (hence the subscript

'B'), the time on C_{B_1} when L_1 flashes is the same as the time on C_{B_2} when L_2 flashes. In the figure it is 12:00 on both clocks when the light bulbs flash.

Starting from the same information (the two light flashes hit Bob simultaneously), appealing to the same postulates, and with a logic as impeccable as Bob's, Al reaches a very different conclusion. Al argues as follows: "According to the light postulate, the velocity of light is not affected by the velocity of its source. Hence, the light flashes of the moving light bulbs L_1 and L_2 both have velocity c with respect to me. Bob, standing in the middle of the railroad car, was rushing away from the flash coming from L_1 and rushing toward the flash coming from L_2. The flash from L_1 thus had a larger distance to cover to get to Bob than the flash from L_2. Yet the two flashes hit Bob at the same time. That means that the flash from L_1 must have started to make its way over to Bob before the flash from L_2 did. *Ergo*, L_1 flashed before L_2."

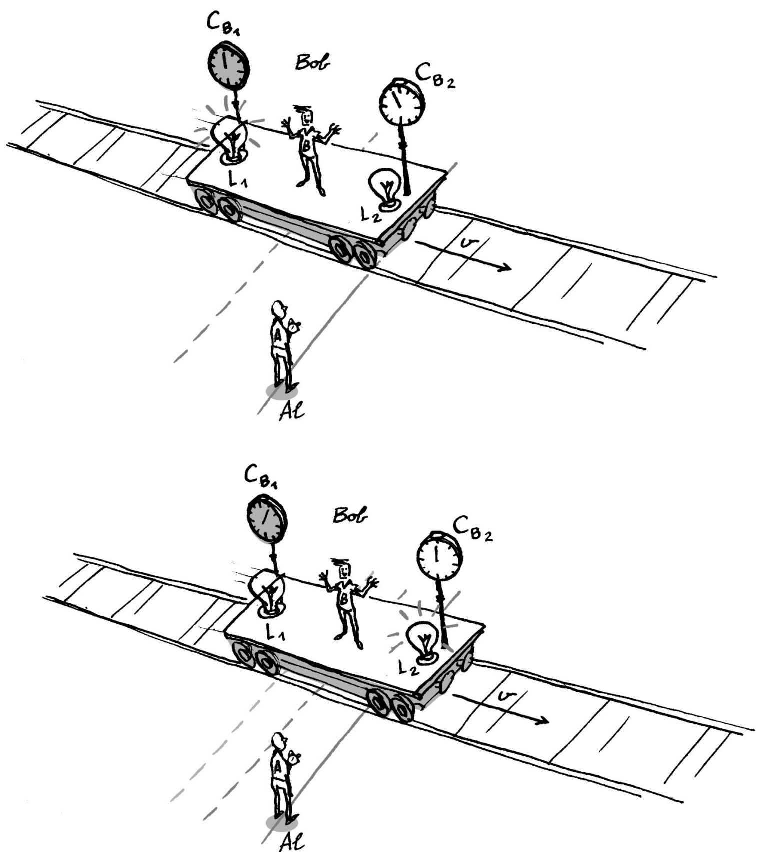

FIGURE 5. According to Al, light bulb L_1 flashed before light bulb L_2.

Fig. 5 illustrates the flashing of the two light bulbs from Al's perspective. The top half shows the flashing of L_1 at the rear of the railroad car; the bottom half the flashing a split-second later of L_2 at the front. It also shows the readings on Bob's clocks C_{B_1} and C_{B_2}. Note that Al agrees with Bob that C_{B_1} read 12:00 when L_1 flashed and, likewise, that C_{B_2} read 12:00 when L_2 flashed. The flashing of a light bulb and the reading on a clock right where that light bulb is are part of one event happening at an instant at a particular location. The disagreement between Al and Bob is never about such individual events, but always about what other events are happening elsewhere at the same time. According to Al, the event 'L_1 flashes and C_{B_1} reads 12:00' at the rear of the railroad car happened before the event 'L_2 flashes and C_{B_2} reads 12:00' at the front. When L_1 flashed, L_2 had not flashed yet and C_{B_2} did not read 12:00 yet but, say, 11:55.[6] Likewise, when L_2 flashed, L_1 had already flashed and C_{B_1} no longer read 12:00 but 12:05. In other words, when Bob has synchronized C_{B_1} and C_{B_2} properly, Al will say that C_{B_1} is five minutes fast compared to C_{B_2}.[7] This is true in general. If one observer synchronizes two clocks at different locations, both at rest with respect to her, another observer with respect to whom she is moving at some constant velocity will find that her rear clock is fast compared to her front clock: the greater the velocity and the farther the two clocks are apart, the greater the discrepancy between the two clocks (the exact formula will be derived in Section 1.5; see Fig. 13).

1.3. **Time dilation: the rate of moving clocks and other processes in systems in motion.** A direct consequence of the relativity of simultaneity is that we have to give up the common-sense notion that clocks in uniform motion run at the same rate as clocks at rest, or, more generally, that processes in systems in uniform motion happen at the same rate as those same processes in systems at rest. Special relativity tells us that processes in moving systems take longer than in systems at rest. This phenomenon is called time dilation. Although it is completely negligible at everyday velocities, time dilation affects all physical processes, mechanical, electrodynamical as well as biological ones such as the metabolism of organisms or the aging of human beings (see Section 2.7).

Fig. 6 brings out the problem with the common-sense assumption about the rate of moving clocks. Suppose Bob wants to measure the rate of one of Al's clocks moving to the left with velocity v (see the first column of the table in Fig. 6). For this measurement he needs at least two stationary clocks at different locations. These clocks are shown in the figure, one with a white, one with a shaded front. Bob makes sure that they are properly synchronized. When Al's clock passes Bob's white clock at 12:00, Al's clock also happens to read 12:00. Half an hour later, at 12:30 on Bob's clocks, Al's clock passes Bob's shaded clock. Our common-sense assumption tells us that at that point Al's clock will read 12:30 as well.

Now look at the situation from Al's perspective (see the second column of the table in Fig. 6). Al and Bob agree that Al's clock and Bob's white clock both read 12:00 when these two clocks meet. They also agree that Al's clock and Bob's shaded clock both read 12:30 when the two clocks meet. These are events taking place at one instant at one place. They are circled with dashed lines in Fig. 6. What Al and Bob disagree about is what Bob's shaded clock reads when Al's clock

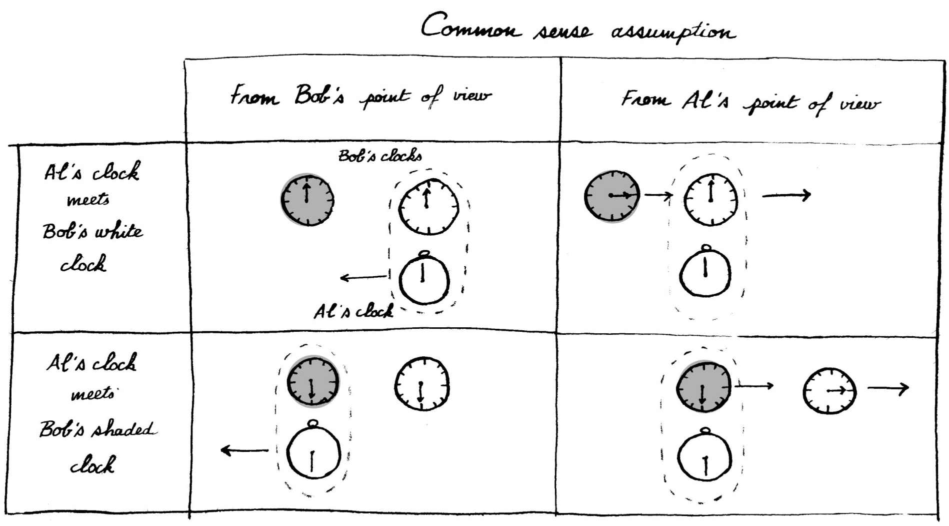

FIGURE 6. Common sense: a moving clock runs at the same rate as the same clock at rest.

meets Bob's white clock and what Bob's white clock reads when Al's clock meets Bob's shaded clock. This, after all, involves judgements about the simultaneity of events at different locations. As we saw in Section 1.2, such judgements depend on one's state of motion. More specifically, Al finds that Bob's clocks, moving to the right with velocity v, are not properly synchronized: the shaded clock at the rear is fast compared to the white clock at the front (see Fig. 5), say, given their velocity and the distance between them, by 15 minutes. According to Al, Bob's shaded clock therefore reads 12:15 (and not 12:00 as Bob claims) when his own clock meets Bob's white clock and Bob's white clock reads 12:15 (and not 12:30 as Bob claims) when his own clock meets Bob's shaded clock.

As a result of this, Al and Bob will disagree about how much time elapses between the event 'Al's clock meets Bob's white clock' and the event 'Al's clock meets Bob's shaded clock'. According to Bob, 30 minutes pass between these two events on all three clocks, in accordance with the common-sense assumption that moving clocks tick at the same rate as clocks at rest. According to Al, however, 30 minutes pass on his own clock, but only 15 on those of Bob: the shaded one goes from 12:15 to 12:30 between the two events, the white one from 12:00 to 12:15. Hence, Al concludes, moving clocks tick at a lower rate than clocks at rest.

This flatly contradicts the relativity postulate. Bob and Al are fully equivalent inertial observers and should judge the rate of each other's clocks in exactly the same way. What is responsible for this contradiction between our common-sense assumption about the rate of moving clocks and the relativity postulate is the relativity of simultaneity. Time dilation provides the escape from this contradiction. Fig. 7 shows how this phenomenon restores the symmetry between Al and Bob.

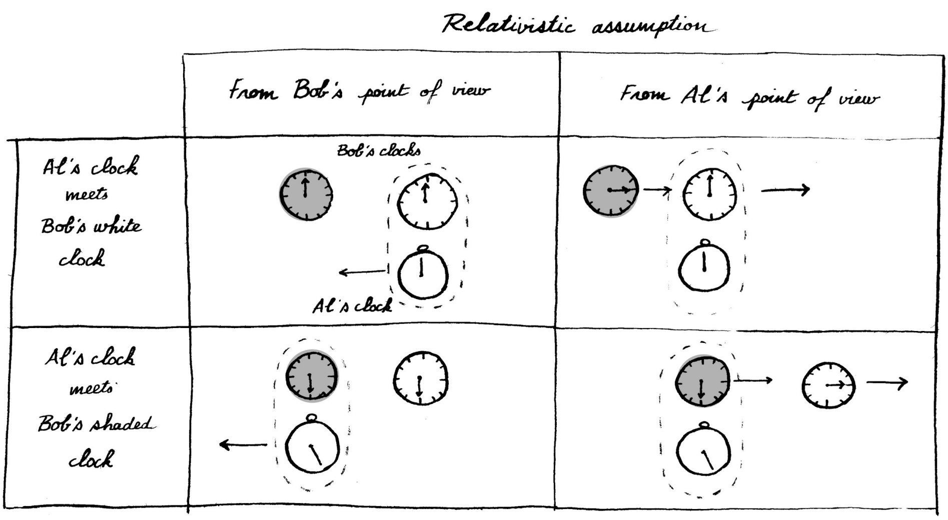

FIGURE 7. Special relativity: moving clocks run slow.

Fig 7 retains as much as possible from Fig. 6. We assume (a) that, for Al, Bob's shaded clock is 15 minutes fast compared to Bob's white clock; (b) that Al's clock and Bob's white clock both read 12:00 as they pass each other; and (c) that it takes 30 minutes on Bob's clocks for Al's to get from one to the other. As we just saw, Al's clock cannot read 12:30 when it passes Bob's shaded clock. That is incompatible with the relativity postulate. So what should Al's clock read at that point? Let x be the as yet unknown amount of time that elapses on Al's clock between the events 'Al's clock meets Bob's white clock' and 'Al's clock meets Bob's shaded clock'. For Bob, the rate of Al's clock is to the rate of his own as $x : 30$. For Al, the rate of Bob's clocks is to the rate of his own as $15 : x$. Because of the relativity postulate, these two ratios must be the same: $x : 30 = 15 : x$. It follows that $x^2 = 30 \cdot 15 = 450$, so $x = \sqrt{450} \approx 21$. This is the number used in the construction of Fig. 7.

First consider the situation from Bob's point of view (see the first column of the table in Fig. 7). For Bob, 30 minutes pass on his own clocks between the two circled events, whereas only 21 minutes pass on Al's. Bob concludes that the rate of Al's clock is 21/30 or about 70% the rate of his own.

Now look at the situation from Al's point of view (see the second column of the table in Fig. 7). For Al, 21 minutes pass on his own clock between the two circled events, whereas only 15 minutes pass on Bob's. Al concludes that the rate of Bob's clocks is 15/21 or about 70% the rate of his own.

Time dilation thus restores the symmetry between Al and Bob. The combination of relativity of simultaneity and time dilation ensures that both observers will claim that the other person's clocks tick at a lower rate than their own.

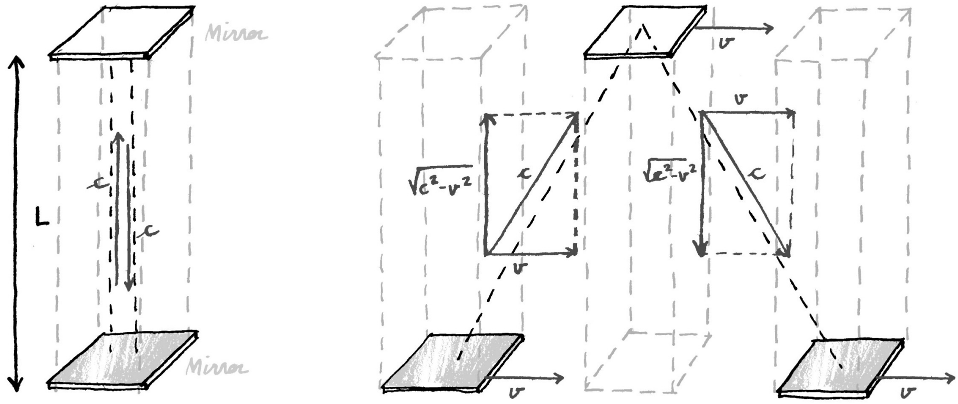

FIGURE 8. Moving clocks run slow by a factor $\sqrt{1-v^2/c^2}$.

The rate of any process in a system moving with velocity v will be $\sqrt{1-v^2/c^2}$ times the rate of that same process in the system at rest. This factor equals 1 for $v = 0$, steadily decreases as v increases, and goes to 0 as v approaches c, the speed of light. A simple way to derive this factor is by examining a so-called "light clock" (Fig. 8). This clock works by having a light signal go back and forth between two mirrors a distance L apart. One tells time by counting how many round-trips of the signal fit into the time interval of interest. Consider this clock from the point of view of two observers. For the first observer, the clock is at rest (see the drawing on the left in Fig. 8); for the second, it is moving with a velocity v perpendicular to L (see the drawing on the right in Fig. 8).[8] For the first observer, the signal simply goes up and down and one round-trip takes $2L/c$. For the second observer, the motion of the signal has a horizontal as well as a vertical component. According to the light postulate the net velocity of the signal is c for both observers. This means that, for the second observer, the signal's velocity in the vertical direction is only $c\sqrt{1-v^2/c^2}$.[9] For this observer, one round-trip of the signal thus takes longer than $2L/c$, namely $(2L/c)/\sqrt{1-v^2/c^2}$. This means that the rate of the light clock in motion is a fraction $\sqrt{1-v^2/c^2}$ of the rate of the light clock at rest. What is true for light clocks will be true for any clock or any process subject to laws compatible with the postulates of special relativity.

1.4. **Length contraction: the length of moving rods and other objects in motion.** Another direct consequence of the relativity of simultaneity is that we have to give up the common-sense notion that uniform motion does not affect the length of measuring rods and other objects. According to special relativity, moving objects must be shorter in the direction of motion than those same objects at rest. This effect is called length contraction. Like time dilation, it is negligible at everyday velocities, but gets large at velocities approaching the speed of light.

The argument for length contraction in this section will have the same structure as the argument for time dilation in Section 1.3. Because of the relativity of simultaneity, the common-sense assumption about the length of moving objects leads to a contradiction with the relativity postulate and length contraction provides the escape from this contradiction.

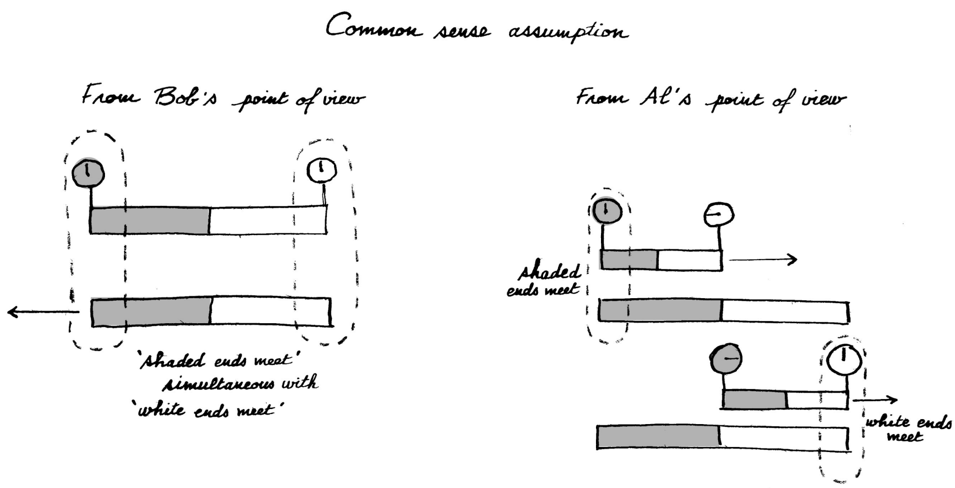

FIGURE 9. Common sense: a moving rod has the same length as the same rod at rest.

Fig. 9 brings out the problem with the common-sense assumption. Suppose Al and Bob are given identical measuring rods. One half of these rods is white, the other half is shaded. Bob attaches clocks to both ends of his rod and carefully synchronizes them. For Bob, Al's rod is moving with velocity v to the left. If the length of rods does not depend on their velocity, there will be one moment in time such that the two rods line up perfectly. This is illustrated in the picture on the left in Fig. 9. At 12:00 on Bob's clocks both the white ends and the shaded ends of the rods meet.

Now look at this situation from Al's perspective (see the drawings on the right in Fig. 9). Bob's rod and the two clocks are moving with velocity v to the right. Al and Bob agree that the shaded ends meet when Bob's shaded clock reads 12:00 and that the white ends meet when Bob's white clock reads 12:00. But Al does not agree with Bob that the events 'shaded ends meet' and 'white ends meet' happen simultaneously. According to Al, Bob's shaded clock is fast compared to his white clock by, say, 15 minutes. When the shaded ends meet at 12:00 on the shaded clock, it is only 11:45 on the white clock. The white ends only meet when the white clock reads 12:00, at which point the shaded one reads 12:15. The way Al sees it, the white ends have yet to meet when the shaded ends meet and the shaded ends have already met when the white ends meet.

Al and Bob will therefore draw very different conclusions about the length of moving rods. According to Bob, moving rods are as long as rods at rest. According to Al, however, moving rods are shorter than rods at rest. This flatly contradicts the relativity postulate. Bob and Al are fully equivalent inertial observers and they should judge the length of moving rods in exactly the same way. Length contraction restores the symmetry between Al and Bob.

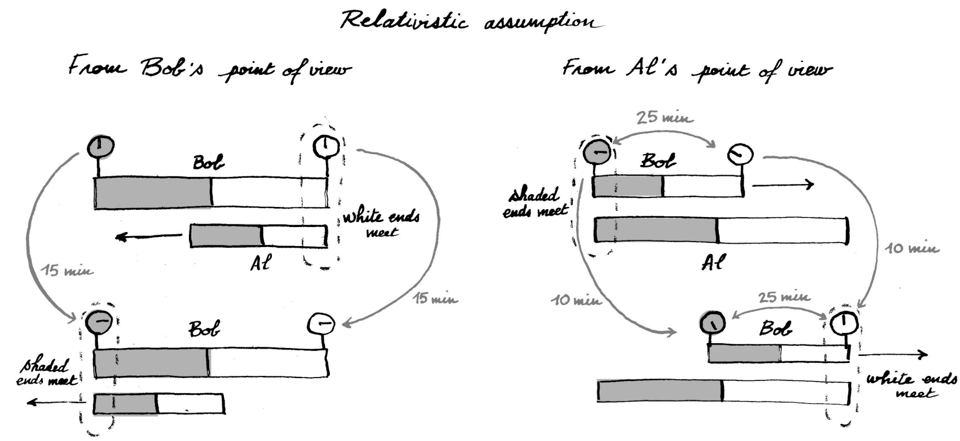

FIGURE 10. Special relativity: moving rods contract.

This is shown in Fig. 10, which not only illustrates the effect of length contraction in this situation but also the effects of relativity of simultaneity and time dilation. Assume that the relative velocity v of Al and Bob and the length of their identical rods are such that both of them will say that (a) the length of the other person's rods is 2/3 the length of their own, (b) the rate of the other person's clocks is 2/3 the rate of their own (as we shall see shortly, these ratios must be the same), and (c) a clock attached to the rear of the other person's rod is 25 minutes fast compared to the clock attached to the front. These are the numbers used in in the construction of Fig. 10.

First consider the situation from Bob's point of view (see the drawings on the left in Fig. 10). For Bob, Al's rod is moving to the left at velocity v and its length is only 2/3 the length of his own. The white ends meet at 12:00; 15 minutes later, at 12:15, the shaded ends meet. Both events are circled with dashed lines.

Now consider the situation from Al's point of view (see the drawings on the right in Fig. 10). Al agrees with Bob that the white ends meet when Bob's white clock reads 12:00, and that the shaded ends meet when Bob's shaded clock reads 12:15. According to Al, however, the shaded clock is 25 minutes fast compared to the white clock. So, according to Al, when the white clock reads 12:00, it is already 12:25 on the shaded clock, and when the shaded clock reads 12:15, it is only 11:50 on the white clock. That means that for Al the order of the events 'white ends meet' and 'shaded ends meet' is just the reverse of their order for Bob.

For Al, the shaded ends meet first (at 12:15 on the shaded and 11:50 on the white clock) and then the white ends meet (at 12:25 on the shaded and 12:00 on the white clock). The 10 minutes that pass on Bob's clocks between these two events (from 12:15 to 12:25 on the shaded clock and from 11:50 to 12:00 on the white one) correspond to 15 minutes on Al's own clocks (not shown in the figure[10]). For Al, after all, Bob's clocks run slow by a factor of 2/3. The situation of Al and Bob is thus fully symmetric. For both of them the front end of the other person's rod meets the corresponding end of their own rod 15 minutes after the rear end meets the corresponding end of their own rod. If Bob found that the length of Al's rod is 2/3 the length of his own, Al will likewise find that the length of Bob's is 2/3 the length of his own.

Length contraction thus restores the symmetry between Al and Bob. The combination of relativity of simultaneity, time dilation, and length contraction ensures that both observers will claim that the other person's rods are shorter than their own.

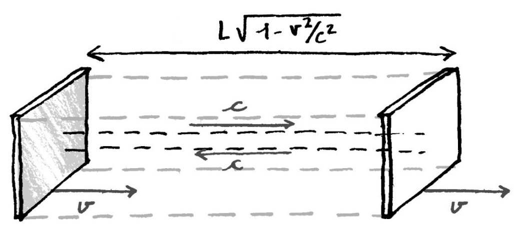

FIGURE 11. Moving rods contract by a factor $\sqrt{1-v^2/c^2}$.

The length of any object moving with velocity v will be shortened by a factor $\sqrt{1-v^2/c^2}$ in the direction of motion compared to that same object at rest. This factor can be derived by considering the light clock introduced in Fig. 8 again. Fig. 11 shows this clock from the point of view of an inertial observer with respect to whom it moves at velocity v in the direction parallel to the line connecting the two mirrors. The light clock, like all other objects, will be shortened in the direction of motion and the distance between the two mirrors will only be xL, where x is some yet to be determined factor between 0 and 1. When the velocity v is parallel to L, the outbound leg of the light signal's round-trip between the two mirrors will take longer than the inbound leg. During the outbound leg, the light signal is heading for a mirror moving away from it; during the inbound leg, it is heading for a mirror moving toward it.

The round-trip will thus take a total time of

$$\frac{xL}{c-v}+\frac{xL}{c+v}=\frac{xL(c+v)+xL(c-v)}{c^2-v^2}=\frac{2xLc}{c^2\left(1-v^2/c^2\right)}=\left(\frac{2L}{c}\right)\frac{x}{1-v^2/c^2}.$$

This result allows us to determine the factor by which moving objects must contract. The rate of a clock in motion should only depend on the magnitude of its velocity, not on its direction. In the case of the light clock, it should therefore not make any difference whether is perpendicular or parallel to L. As we saw in Section 1.3, the rate of a light clock moving perpendicular to L is a fraction $\sqrt{1-v^2/c^2}$ of its rate at rest. The equation above shows that the rate of a light clock moving parallel to L is a fraction $\left(1-v^2/c^2\right)/x$ of its rate at rest. Since these two fractions must be the same, the factor x by which moving objects contract in their direction of motion must be equal to $\sqrt{1-v^2/c^2}$.[11]

1.5. **The addition of velocities.** With the three consequences of the postulates of relativity derived in Sections 1.2–1.4 (relativity of simultaneity, time dilation, and length contraction), we have all the necessary ingredients to explain how two inertial observers moving with respect to one another can both find that light has velocity c with respect to them.

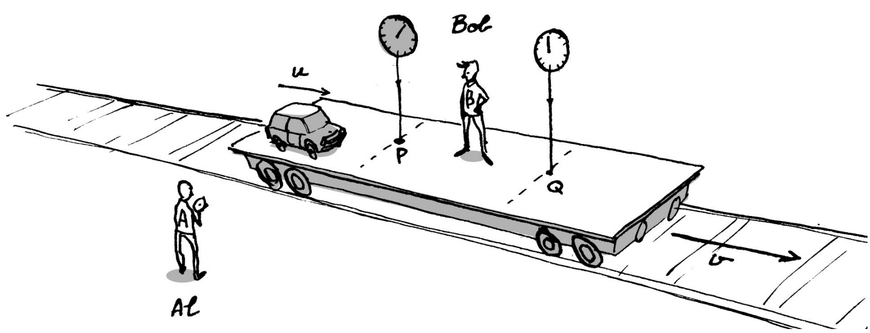

FIGURE 12. The velocity of the little car according to Al and Bob.

Consider the situation in Fig. 12. As in Figs. 3–5, Bob is standing on a railroad car moving to the right with respect to Al at a constant velocity v. On the railroad car is a remote control car also moving to the right with some constant velocity. According to Bob, the car is moving with velocity u with respect to the railroad car. What is the velocity w of the car with respect to Al? The common-sense answer would be $u + v$. This answer is wrong. Al cannot accept the velocity u reported by Bob. As far as Al is concerned, Bob has determined this velocity with rods that are contracted and clocks that run slow and are out of sync. Al needs to correct u for all three of these effects before he can add it to v.

In Sections 1.3 and 1.4 we saw that moving clocks run slow by a factor $\sqrt{1-v^2/c^2}$ and that moving rods contract by that same factor. To calculate Al's corrections to velocities reported by Bob we also need to know by how much two clocks are out of sync when they are both moving at some velocity v a certain distance apart.

Fig. 13 shows how Bob checks whether his clocks at P and Q are properly synchronized. At $t = 0$ on the clock at Q he sends out a light signal from Q to P. If the distance between P and Q is D (according to Bob), the travel time of the signal is D/c (according to Bob). Hence, Bob concludes that P and Q are properly synchronized if the clock at P reads $t = D/c$ the moment the signal arrives at P.

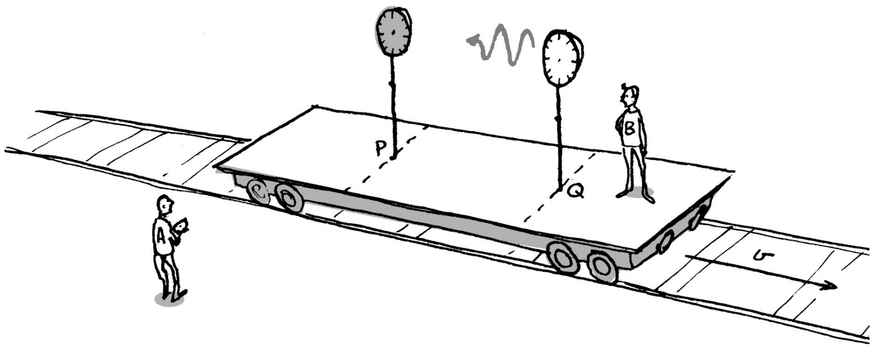

FIGURE 13. Bob checks whether his clocks are properly synchronized.

Consider Bob's synchronization check from Al's perspective. According to Al, the travel time of the signal is less than D/c. Bob measures the distance between P and Q with contracted rods, so the real distance is only $D\sqrt{1 - v^2/c^2}$. Moreover, the clock at P is rushing toward the light signal with velocity v. According to Al, the travel time is therefore only:

$$\frac{D\sqrt{1 - v^2/c^2}}{c + v}.$$

Since Bob's clocks, according to Al, run slow, they only register a fraction $\sqrt{1 - v^2/c^2}$ of this travel time. Al concludes that the clocks at P and Q are properly synchronized if the one at P reads

$$t = \frac{D\,(1 - v^2/c^2)}{c + v}$$

when the light signal arrives at P. Using that

$$1 - \frac{v^2}{c^2} = \frac{1}{c^2}\left(c^2 - v^2\right) = \frac{1}{c^2}(c + v)(c - v),$$

we can rewrite this equation as

$$t = \frac{D}{c^2}\frac{(c + v)(c - v)}{c + v} = \frac{D}{c} - \left(\frac{v}{c^2}\right) D.$$

If Bob has properly synchronized his clocks, the one at P will, in fact, read $t = D/c$ when the light signal arrives. Al concludes that, since it *should* read $(v/c^2)D$ *before* D/c when the light signal arrives, the clock at P is *fast* compared to the one at Q by $(v/c^2)D$.

Now return to Fig. 12. To determine the velocity of the remote control car with respect to the railroad car, Bob measures the time T it takes the car to get from P to Q and divides the result into the distance D between these two points. To find T, Bob subtracts the reading on his clock at P when the car passes P from the reading on his clock at Q when the car passes Q. According to Al, the clock at P is $(v/c^2)D$ fast compared to one at Q. Al therefore has to *add* the difference to the time reported by Bob.[12] He then needs to divide the result by $\sqrt{1-v^2/c^2}$ since Bob is measuring times with clocks that run slow. According to Al, it thus takes the car longer than T to get from to P to Q, namely

$$\frac{T+(v/c^2)D}{\sqrt{1-v^2/c^2}}.$$

Finally, Al has to multiply the distance D reported by Bob by $\sqrt{1-v^2/c^2}$ since Bob is measuring distances with rods that are contracted. Dividing this shorter distance by the longer time above, Al finds that the velocity of the remote control car with respect to the railroad car is less than the velocity $u = D/T$ reported by Bob, namely

$$u_{\text{corrected}} = \frac{D\,(1-v^2/c^2)}{T+(v/c^2)D}.$$

Dividing numerator and denominator by T and using that $D/T = u$, we find[13]

$$u_{\text{corrected}} = \frac{u\,(1-v^2/c^2)}{1+uv/c^2}.$$

Adding the velocity v of the railroad car, Al finds that the velocity w of the remote control car with respect to him is not $u+v$, as suggested by common sense, but:

$$w = u_{\text{corrected}} + v = \frac{u\,(1-v^2/c^2)+v\,(1+uv/c^2)}{1+uv/c^2} = \frac{u+v}{1+uv/c^2}.$$

This result is known as the relativistic addition theorem of velocities. It gives the rule for the composition of velocities such as u and v measured by different inertial observers. Adding such velocities directly would be like adding apples and oranges. The addition theorem is such that as long as the velocities u and v are both subluminal (i.e., less than the velocity of light) their composite $u_{\text{corr}}+v$ is also subluminal. This means that an object can never be accelerated from subluminal to superluminal velocities and that its inertial mass, a measure of its resistance to acceleration, must increase without limit as its velocity approaches the speed of light.[14]

Suppose that Bob determines the velocity of the light from the headlights of the remote control car. In that case, $u = c$. Inserting this into the formula above, we find that this light also has velocity c with respect to Al:[15]

$$u_{\text{corrected}} + v = \frac{c+v}{1+v/c} = \frac{c(1+v/c)}{1+v/c} = c.$$

This is just as it should be according to the postulates of special relativity.

2. Special relativity and Minkowski space-time

2.1. Minkowski or space-time diagrams. Fig. 14 shows a series of snapshots illustrating, from Al's point of view, an experiment like the one analyzed in Section 1.2 with Bob standing on a railroad car moving to the right at velocity v. This sequence of snapshots should be read from bottom to top.

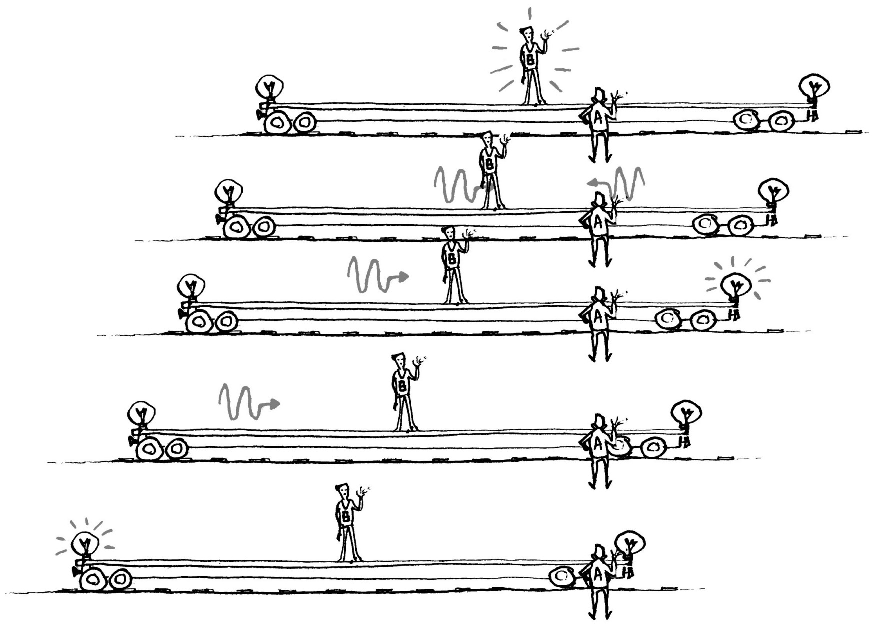

Figure 14. An experiment like the one in Figs. 3–5 from Al's perspective.

The first snapshot shows the flashing of the light bulb at the rear of the railroad car. The next four show the light flash from this light bulb catching up with Bob, who is rushing away from it. The third shows the flashing of the light bulb at the front. The next two show the light flash from this light bulb making its way over to Bob, who is rushing toward it. The two light flashes both hit Bob in the final snapshot. Bob is still somewhat to the left of Al when this happens.

In Figs. 15 and 16, the snapshots of Fig. 14 are used to construct a *Minkowski diagram* or *space-time diagram* for this situation. These diagrams picture the way space and time must be in a world in accordance with the postulates of special relativity. This relativistic space-time is called *Minkowski space-time* or, to distinguish it from the curved space-times of general relativity, *flat space-time*. It is standard practice to suppress two spatial dimensions in space-time diagrams. That leaves only one spatial dimension in addition to the time dimension, so motion can only be to the left or to the right. This suffices to illustrate all salient features of Minkowski space-time.

We begin our construction of a space-time diagram for the situation shown in Fig. 14 by picking space and time axes for a frame of reference in which Al is at

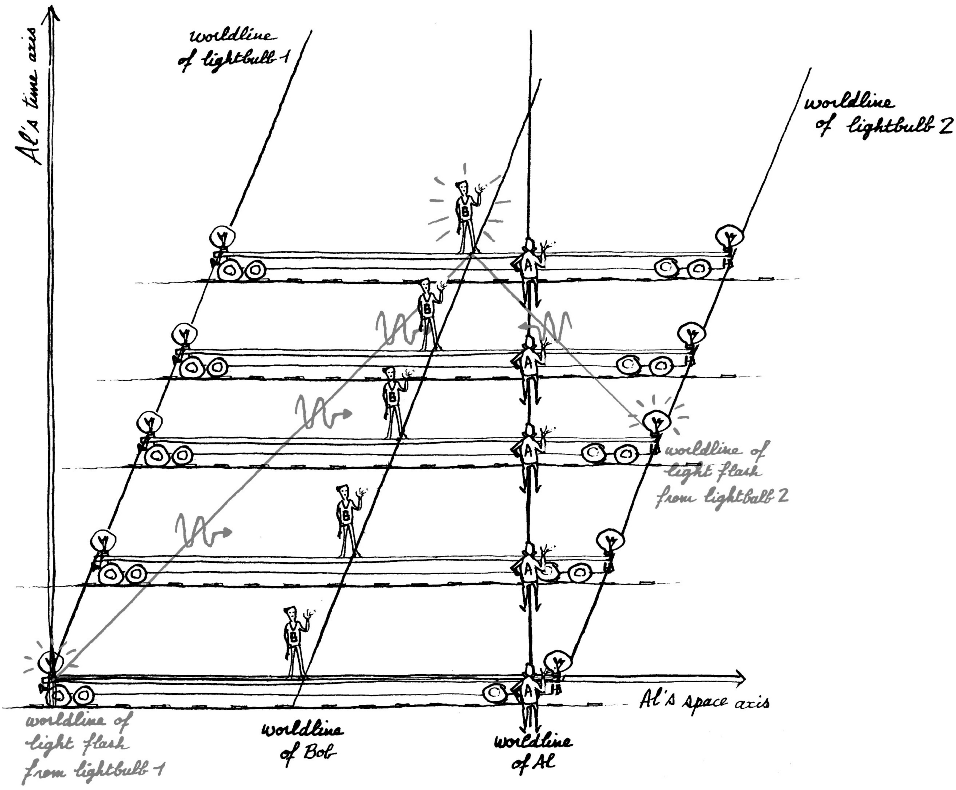

FIGURE 15. Constructing a Minkowski or space-time diagram.

rest. Let the event 'light bulb L_1 flashes' be the origin of this space-time coordinate system. Al's space and time axes are the horizontal and the vertical lines through the origin. We then trace the trajectories of the various elements pictured in Fig. 14 through the space-time spanned by these two axes: the observers Al and Bob, the light bulbs L_1 and L_2, and the light flashes emitted by them. Such trajectories are called *world lines.* In general, world lines will not be straight. Only those of objects at rest or in uniform motion will be. Vertical lines are the world lines of objects at rest in Al's frame. The world line of Al himself provides an example of this. The other world lines in Fig. 15 are tilted, i.e., they are at an angle to the time axis of Al's frame. These are the world lines of objects moving at constant velocity in Al's frame, to the right if the world line is tilted to the right, to the left if it is tilted to the left. The greater the tilt, i.e., the greater the angle between the world line and the time axis, the greater the velocity. The world lines of Bob and the light bulbs L_1 and L_2 are tilted to the right at the same angle. They are parallel to one another. This reflects that these three objects move at the same velocity. The world lines of the light flashes from L_1 and L_2 are tilted at a larger angle, to the left and the right, respectively. We shall use units for measuring times and distances such that the world lines of light will always be tilted at 45°.

Where two world lines intersect, the corresponding objects meet. For instance, the world lines of the light flashes from L_1 and L_2 intersect Bob's world line at the event 'the two light flashes hit Bob'.

The icons depicting Al, Bob, the light bulbs, and the light signals in Fig. 15 are nothing but window dressing in the end. Stripping Fig. 15 of such unnecessary detail, we are left with the space-time diagram shown in Fig. 16. Every horizontal slice of Fig. 16 corresponds to a snapshot like the ones shown in Fig. 14.

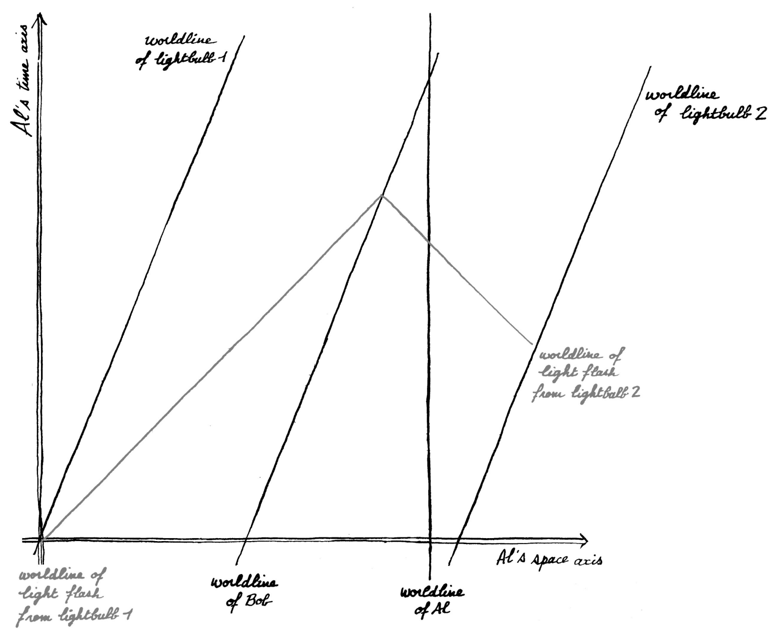

FIGURE 16. Minkowski or space-time diagram.

In the internationally accepted system of units, distance is measured in meters and time in seconds. In these units c, the velocity of light, is about $3 \cdot 10^8$ m/s. Following standard practice in special relativity, we use a different time unit, which can be called—although the name is not commonly used—the 'light meter' (cf., e.g., Mermin 1968, 180). One light meter is defined as the time it takes light to travel one meter. It follows that one second is about $3 \cdot 10^8$ light meters. The light meter is thus a much smaller unit of time than the second. To convert seconds to light meters we need to multiply by $c \approx 3 \cdot 10^8$ m/s: t seconds correspond to ct light meters. The result of multiplying seconds and meters per second is meters. If distances are measured in meters and times in light meters, then both quantities have the same dimension and can meaningfully be added to one another. As we shall see in Sections 2.5–2.6, this is an important advantage of using light meters. Another advantage is that the velocity of light in these units is equal to 1. From the definition of a light meter it follows that light travels one meter per light meter. This is why the world lines of light are always tilted at 45° in our space-time diagrams.[16]

In the remainder of this section we develop the geometry of Minkowski spacetime with the help of space-time diagrams. In Section 2.2, the key insight of special relativity, the relativity of simultaneity, is given a geometrical representation: different inertial observers have different ways of carving up space-time into slices of simultaneous events. In a space-time diagram one can see at a glance how it can be that light has the same velocity for all inertial observers. They also make it easy to see that objects moving faster than light in one frame of reference go backward in time in another (Section 2.3). This amounts to a strong argument against the existence of so-called tachyons, particles moving faster than light, for they would open up a Pandora's box of causal paradoxes. In Section 2.4, the arguments for time dilation and length contraction of Sections 1.3–1.4 are rephrased in geometrical language. In Sections 2.5–2.6, the spatio-temporal distance between events in Minkowski space-time is introduced in analogy with the distance between points in Euclidean space. Section 2.7 is devoted to the famous 'twin paradox', which shows vividly that the time it takes to go from one event to another in Minkowski space-time depends on the space-time trajectory between those two events, just as the distance covered in going from one point to another in ordinary space depends on the trajectory between those two points.

2.2. **Relativity of simultaneity: different observers carving Minkowski space-time into simultaneity slices.** Fig. 17 once again shows the space-time diagram corresponding to the series of snapshots in Fig. 14. The events 'light bulb L_1 flashes', 'light bulb L_2 flashes' and 'the light flashes from L_1 and L_2 hit Bob', are labeled O, P, and Q, respectively. The event O is the origin of Al's space-time coordinate system. The point R on Al's space axis and the point S on his time axis mark the space and time coordinates that Al assigns to the event Q. According to Al, Q happens a distance OR to the right of O and a time OS after O. According to Al, the light signal from O to Q covers the distance OR in a time OS. In the figure OR (representing meters) is equal to OS (representing light meters). The light thus has the velocity $OR/OS = 1$ meter/light meter, the velocity of light, with respect to Al.

Fig. 17 also shows the space and time axes for Bob and the points U and V marking the space and time coordinates of the event for Bob. The origin of Bob's space-time coordinate system coincides with the origin of Al's. How do we find the space and time axes for Bob?

The time axis is a line through O that, according to Bob, is purely in the time direction. In other words, it is a line connecting events that, according to Bob, all happen at the same location. This means that Bob's time axis is a line through O parallel to Bob's world line. Bob's time axis is tilted to the right of Al's time axis at an angle $\angle SOV$.

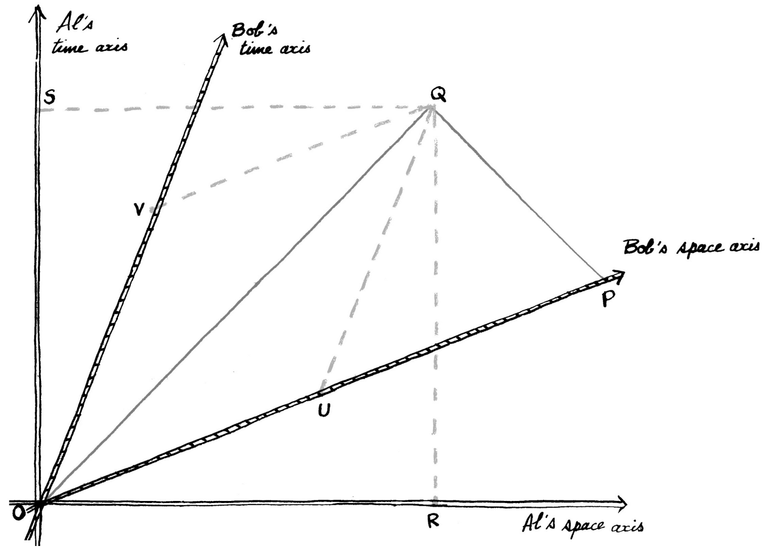

Figure 17. Space and time axes for Al and Bob.

Bob's space axis is a line through O that, according to Bob, is purely in the spatial direction. In other words, it is a line connecting events that, according to Bob, all happen at the same time. According to Bob, the events O ('L_1 flashes') and P ('L_2 flashes') happen at the same time (see Figs. 3–4). Bob's space axis is therefore the line through O and P. This line is tilted with respect to Al's space axis at an angle $\angle ROP$. As we shall see shortly, Bob's space axis must be tilted with respect to Al's space axis at the same angle as Bob's time axis with respect to Al's time axis ($\angle ROP = \angle SOV$).

For Bob, as for Al, the time direction is orthogonal (i.e., perpendicular) to the space direction: $\angle VOU = \angle SOR = 90°$. In Fig. 17, however, the angle between Bob's space and time axes is acute. This is because the geometry of Minkowski space-time is not the same as ordinary Euclidean geometry. Using a Euclidean sheet of paper to represent a two-dimensional Minkowski space-time, we cannot capture all features of the latter. An example of this complication is that the angle $\angle VOU$ between Bob's space and time axes, which is a right angle in space-time, is represented by an acute angle in Fig. 17.

The different directions of the space axes of Al and Bob reflect their disagreements about simultaneity. Al and Bob each carve up Minkowski space-time into slices of simultaneous events by drawing lines parallel to their own space axis. If two events at different locations lie on the same simultaneity slice for one of them, they cannot lie on the same simultaneity slice for the other.

The points U and V mark the space and time coordinates of the event Q in Bob's coordinate system. The point U is found by drawing a line through Q parallel to Bob's time axis and determining where this line intersects Bob's space axis. (Note that the events U and Q happen at the same location according to Bob.) The point V is found by drawing a line through Q parallel to Bob's space axis and determining where this line intersects Bob's time axis. (Note that the events V and Q happen at the same time according to Bob.) From Bob's point of view, the light signal from O to Q covers the distance OU in a time OV and its velocity is thus OU/OV. According to the postulates of relativity, this velocity must be 1 meter/light meter. Hence, OU must be equal to OV. Since $OU = OV$, UQ is parallel to OV, and VQ is parallel to OU, the quadrangle $OUQV$ is a rhombus and $\angle SOV = \angle ROU$. We can now understand why the angle between the space axes of Al and Bob must be the same as the angle between their time axes: this is what guarantees that Al and Bob both find that the velocity of light with respect to them is 1 meter/light meter.[17]

The conversion of the description of a physical situation in Al's space-time coordinate system to a description of that same situation in Bob's space-time coordinate system is an example of a (passive) *Lorentz transformation*. The same transformation equations can be used to turn the description of a system at rest into the description of a system in motion in the same space-time coordinate system (*active Lorentz transformation* or *Lorentz boost*). Lorentz had introduced the transformation equations in this latter sense before 1905 but the former sense had eluded him.[18] The basic transformation equations for the space-time coordinates incorporate the relativity of simultaneity, time dilation, and length contraction. The relativity principle requires that all physical laws can be expressed in the same way in any of the space-time coordinate systems related by Lorentz transformations. Special relativity, in other words, requires all physical laws to be *Lorentz invariant*. This, in turn, guarantees that all physical systems obeying these laws will indeed exhibit the effects deduced from the postulates in Section 1.

There is nothing sacred about representing Al's space and time axes by horizontal and vertical lines. We might just as well represent Bob's space and time axes by horizontal and vertical lines, in which case we arrive at the space-time diagram in Fig. 18. As in Fig. 17, the events O ('L_1 flashes'), P ('L_2 flashes') and Q ('the light flashes hit Bob') are labeled, and so are the points R, S, U, and V marking the coordinates of the event Q on the space and time axes of Al and Bob.

Al is moving to the left with respect to Bob at velocity v, so Al's time axis is tilted to the left at the same angle that Bob's time axis is tilted to the right in Fig. 17. Al's space axis must be tilted at that same angle with respect to Bob's space axis ($\angle SOV = \angle ROU$) so that the quadrangle $ORQS$ is a rhombus and OR/OS, the velocity of the light signal from to according to Al, equals 1 meter/light meter. Note that O and P lie on the same horizontal line, which in this space-diagram is what tells us that these events ('L_1 flashes' and 'L_2 flashes') are simultaneous for Bob. For Al, P comes after O: the simultaneity slice containing P (a line through P parallel to Al's space axis) lies above the simultaneity slice containing O (Al's space axis).

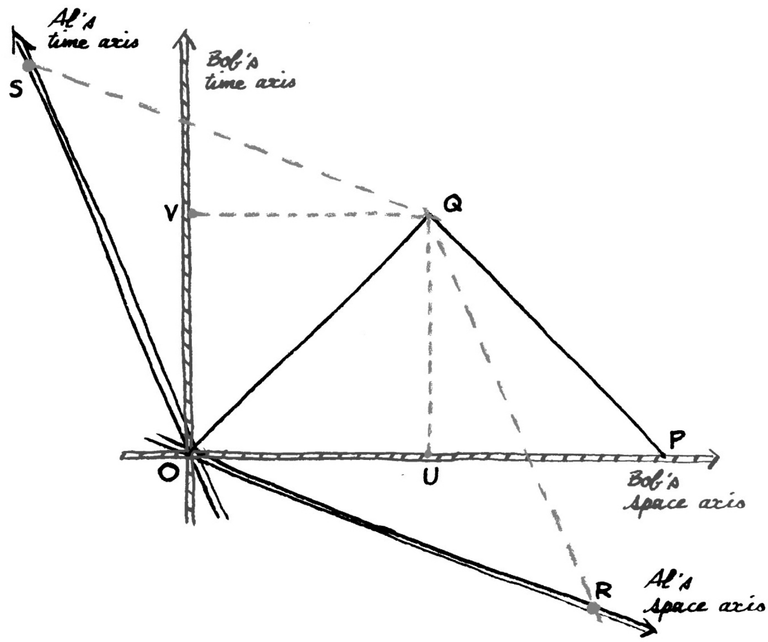

FIGURE 18. Space-time diagram with the space and time axes of Bob represented by horizontal and vertical lines.

With the help of Figs. 17 and 18 a complete inventory can be made of two-dimensional space-time coordinate systems with origin O and orthogonal axes moving with respect to one another at arbitrary but constant subluminal velocities. Choose one such coordinate system and represent its space and time axes by horizontal and vertical lines. The space and time axes of equivalent coordinate systems moving to the right with respect to the chosen one will look like Bob's space and time axes in Fig. 17. The right angle between the space and time axes of such coordinate systems will be represented by an acute angle. The time axis is tilted to the right from the vertical by some angle between 0° and 45° and the space axis is tilted up from the horizontal by that same angle. The space and time axes of equivalent coordinate systems moving to the left with respect to the chosen one will look like Al's space and time axes in Fig. 18. The right angle between the space and time axes of such coordinate systems will be represented by an obtuse angle. The time axis is tilted to the left from the vertical by some angle between 0° and 45° and the space axis is tilted down from the horizontal by that same angle.

2.3. **Tachyons, causal loops, and causal paradoxes.** A tachyon is a hypothetical particle moving faster than light. Suppose Al creates a tachyon at space-time point P. The tachyon moves to the right and is annihilated at space-time point Q. Since its velocity with respect to Al is superluminal (i.e., greater than that of light), its world line is tilted at an angle somewhere between 45° and 90° with respect to Al's time axis. This world line is shown in the first space-time diagram in Fig. 19. The space and time axes of Al are represented by horizontal and vertical lines, those of Bob by tilted lines. The event P ('tachyon is created') is chosen as the origin of both coordinate systems. Bob is moving to the right with respect to Al. The velocities of Bob and the tachyon with respect to Al are such that the world line of the tachyon lies between the space axes of Al and Bob. For Bob, the tachyon is therefore annihilated (at Q) before it is created (at P)! For Bob, in other words, the tachyon is not only moving faster than light, it is also moving backward in time!

The second space-time diagram in Fig. 19 shows the same situation but now Bob's space and time axes are represented by horizontal and vertical lines. The tachyon is created at P, travels backward in time (for Bob), and is annihilated at Q. If it is possible to create a tachyon at P, it should also be possible to create one at Q. There are no preferred points in space (i.e., space is homogeneous). Likewise, if it is possible to create a tachyon moving to the right, it should also be possible to create one moving to the left. There are no preferred directions in space (i.e., space is isotropic). Hence, if it is possible to have a tachyon going from P to Q, it should also be possible to have one just like it going from Q to R, where R is an event that, for Bob, happened at the same location as P two lifetimes of these tachyons before P. This is shown in the third space-time diagram in Fig. 19. If Bob's world line goes through R and P, he can use these tachyons to send information acquired at P to his former self at R. The figure $PQRP$ forms what is called a *causal loop*. The presence of causal loops puts tight constraints on what can happen in the space-time points that are part of them. Given a causal loop it is easy to construct scenarios that lead to *causal paradoxes*, scenarios in which we are driven to the logically impossible conclusion that one and the same state of affairs both does and does not obtain.

An example may help to appreciate what it would take to rule out causal paradoxes in the presence of causal loops. With the help of a tachyon phone, a gadget for sending and receiving messages carried by tachyons, Bob is planning a museum heist in broad daylight. He parks his car outside the museum and checks his tachyon phone for messages from his future self. If there are none, he goes inside, takes his favorite painting, and walks out again with the painting under his arm. Should the alarm go off or should he be stopped by security, he uses his tachyon phone to tell his former self to call off the heist for now and try again later. One day, Bob reasons, security at the museum will fail and the heist will succeed. Clearly, Bob's reasoning is fallacious. Suppose the alarm does go off, as it probably will during most attempts. This would prompt Bob to send his message back in time. But then he would have received it before he entered the museum and, heeding the warning, would have driven off; in which case he would not have sent his message to begin with. This scenario is not just physically but logically

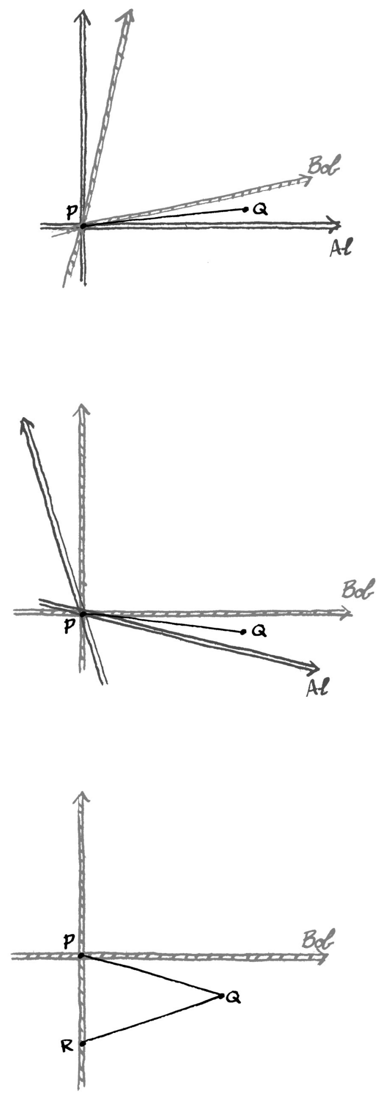

FIGURE 19. Tachyons and causal paradoxes.

impossible. Bob cannot both send and not send his message. This particular chain of events on the causal loop must be broken somewhere. For some reason, for instance, his tachyon phone malfunctioned. Or it did work and he did get the message but Bob foolhardily went ahead and tried to steal the painting anyway. If the tachyon phone works properly and Bob sticks to his plan, something else must go wrong. Bob cannot possibly have left the scene tipped off by a message he is never in a position to send.

If there were tachyons we could produce many causal paradoxes such as Bob's bungled heist. The existence of tachyons thus requires elaborate conspiracies in nature that prevent the chain of events in such loops to turn into logically contradictory scenarios. Believing in such conspiracies is the price we have to pay for believing in tachyons. For most physicists this price is too high. There is broad consensus that tachyons do not exist and that no signal capable of transmitting information from one place to another can travel faster than the speed of light.

2.4. **Time dilation and length contraction: different observers using different line segments to determine the rate of clocks and the length of rods.** The space-time diagrams in Figs. 20–23 in this section are the counterparts of Figs. 6–7, 9–10 that formed the basis for the arguments for time dilation and length contraction in Sections 1.3–1.4. This section offers geometrical versions of those arguments.

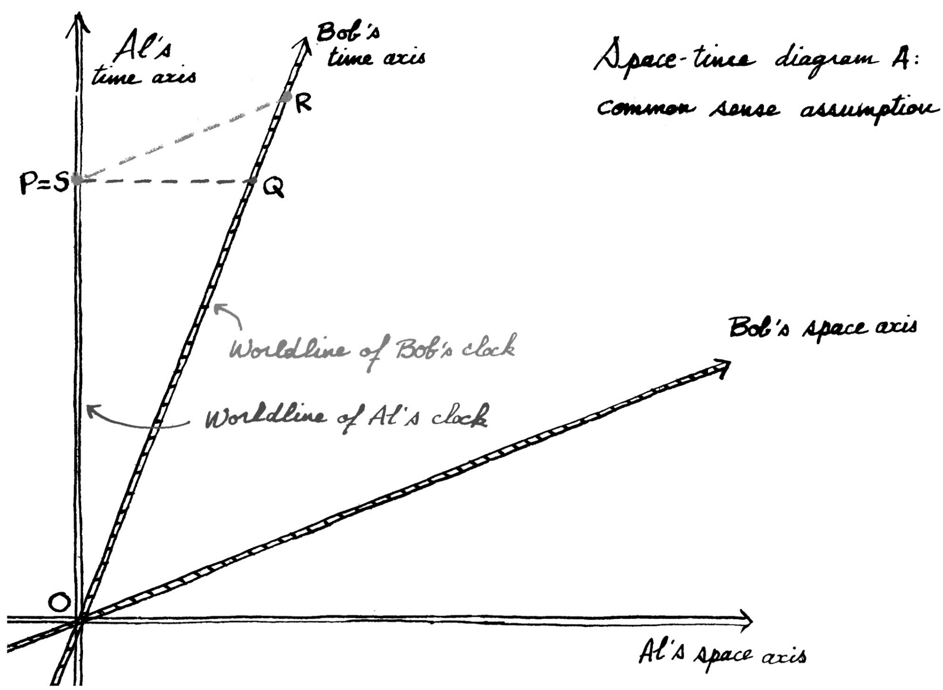

FIGURE 20. A moving clock ticks at the same rate as that same clock at rest.

Fig. 20 brings out the problem with the common-sense assumption that moving clocks tick at the same rate as clocks at rest. Al and Bob have identical clocks with world lines that coincide with their respective time axes. Al and Bob both pick the event O as their zero point of time. Let P be the point on Al's time axis where Al's clock reads 1 light meter and let R be the point on Bob's time axis that, according to Bob, is simultaneous with P. Similarly, let Q be the point on Bob's time axis where Bob's clock reads 1 light meter and let S be the point on Al's time axis that, according to Al, is simultaneous with Q.

On the basis of the common-sense assumption about moving clocks, Al would insist that S (simultaneous to Q according to Al) coincides with P, as shown in Fig. 20. In that case, after all, the two clocks would take the same time to go from reading 0 to reading 1 for Al: $OS = OP$. For Bob, however, Al's clock would then take longer to go from 0 to 1 than his own clock: $OR > OQ$. In other words, according to Bob, Al's clock would run slow.

On the basis of the same common-sense assumption about moving clocks, Bob would insist that R (simultaneous to P according to Bob) coincides with Q. In that case, the two clocks would take the same time to go from 0 to 1 for Bob: $OR = OQ$. For Al, however, Bob's clock would then take longer to go from 0 to 1 than his own clock: S would be above P which means that $OS > OP$. Hence, according to Al, Bob's clock would run slow.

Either way we run into a contradiction with the relativity principle. Al and Bob are equivalent inertial observers. They should judge the rate of moving clocks the same way.

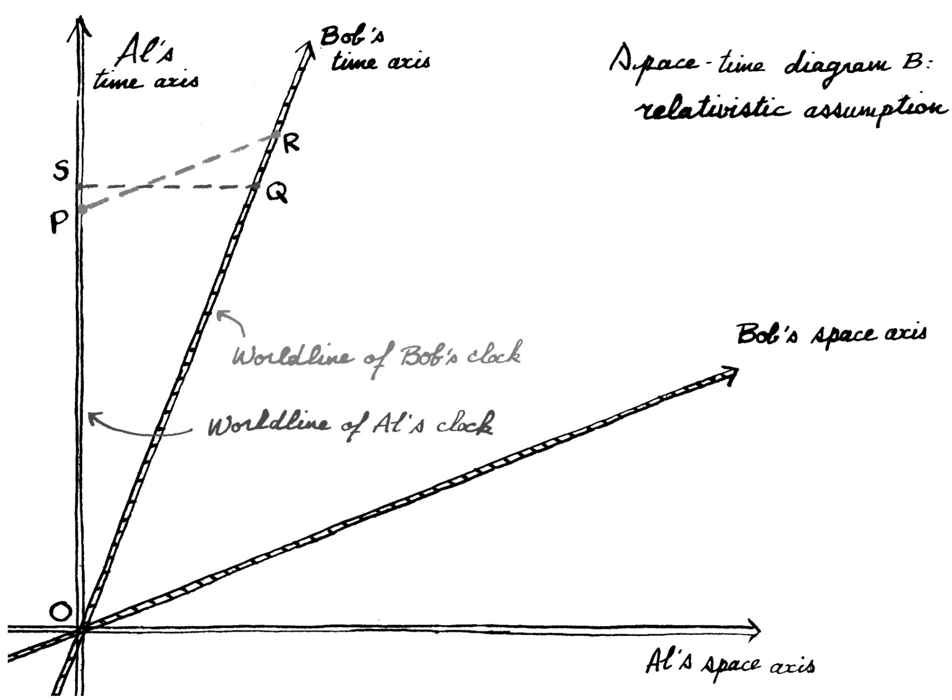

Figure 21. Moving clocks run slow.

The solution to the problem is to choose P and Q in such a way that the line segments PR and QS intersect somewhere in between the time axes of Al and Bob, as shown in Fig. 21. In that case Al and Bob both conclude that the other person's clock runs slow, Al because $OS > OP$, Bob because $OR > OQ$. Time dilation thus restores the symmetry between Al and Bob.

Fig. 21 also provides the solution to the puzzle of how Al and Bob can both claim that the other person's clock runs slow. They agree that $OS > OP$ and that $OR > OQ$, but they disagree about which line segments represent the time elapsed between the readings 0 and 1 on their two clocks, OP and OS on Al's time axis or OQ and OR on Bob's.

Although the space-time diagram in Fig. 21 solves this puzzle, it seems to leave us with another. Since the clocks of Al and Bob are identical, the line segments on their own time axes representing the time elapsed between the readings 0 and 1 on their own clocks must be equally long, i.e., $OP = OQ$. In Fig. 21, however, OQ is longer than OP. This is another feature of Minkowski space-time that space-time diagrams on Euclidean sheets of paper fail to capture. One light meter on a tilted time axis is represented by a larger line segment than one light meter on a vertical axis. The more tilted the axis, the larger the segment corresponding to one unit of time.

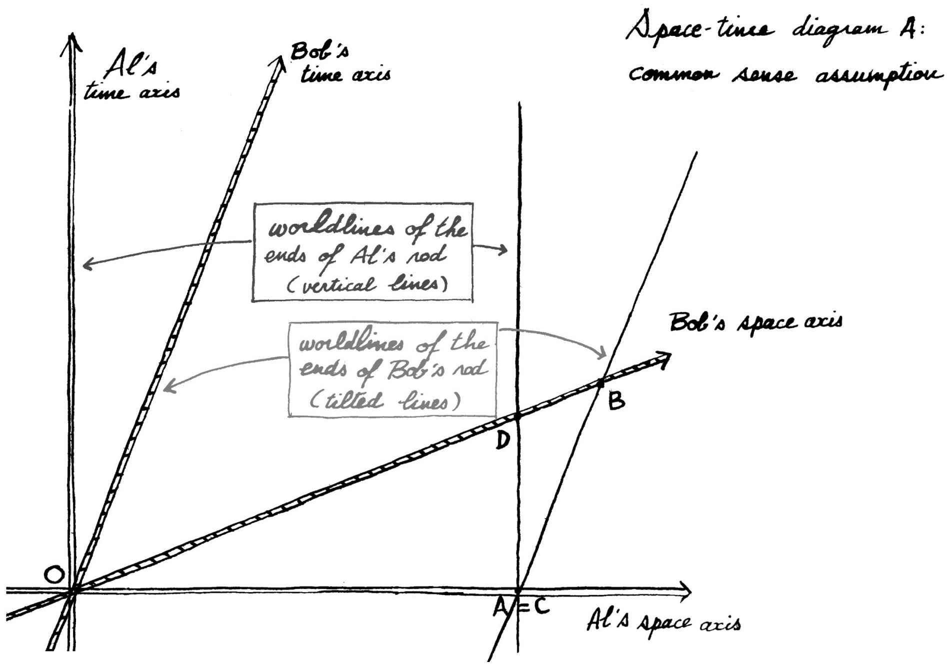

FIGURE 22. A moving rod has the same length as that same rod at rest.

A geometrical argument fully analogous to the one for time dilation can be given for length contraction. Consider Fig. 22. Al and Bob have identical rods—half white, half shaded (cf. Figs. 9–10). The shaded ends meet at the common origin O of their space-time coordinate systems. The world lines of the shaded ends of the rods thus coincide with their times axes. The world lines of the white ends are parallel to these time axes. Call the points where the world line of the white end of Al's rod intersects the space axes of Al and Bob A and D, respectively, and the points where the world line of the white end of Bob's rod intersects the space axes of Al and Bob C and B, respectively. Consider the segments AD and BC of the world lines of the white ends of the two rods.

On the basis of the common-sense assumption that moving rods have the same length as rods in motion, Al would insist that AD and BC intersect at $A = C$, as shown in Fig. 22. In that case, after all, the two rods would be equally long for Al: $OA = OC$. But then Bob would say that his rod is longer than Al's: $OB > OD$.

On the basis of the same common-sense assumption about moving rods, Bob would insist that AD and BC intersect at $B = D$. In that case, the two rods would be equally long for Bob: $OB = OD$. But then Al would say that his rod is longer than Bob's: C would be to the left of A which means that $OA > OC$.

Either way we run into a contradiction with the relativity principle. Al and Bob are equivalent inertial observers. They should judge the length of moving rods the same way.

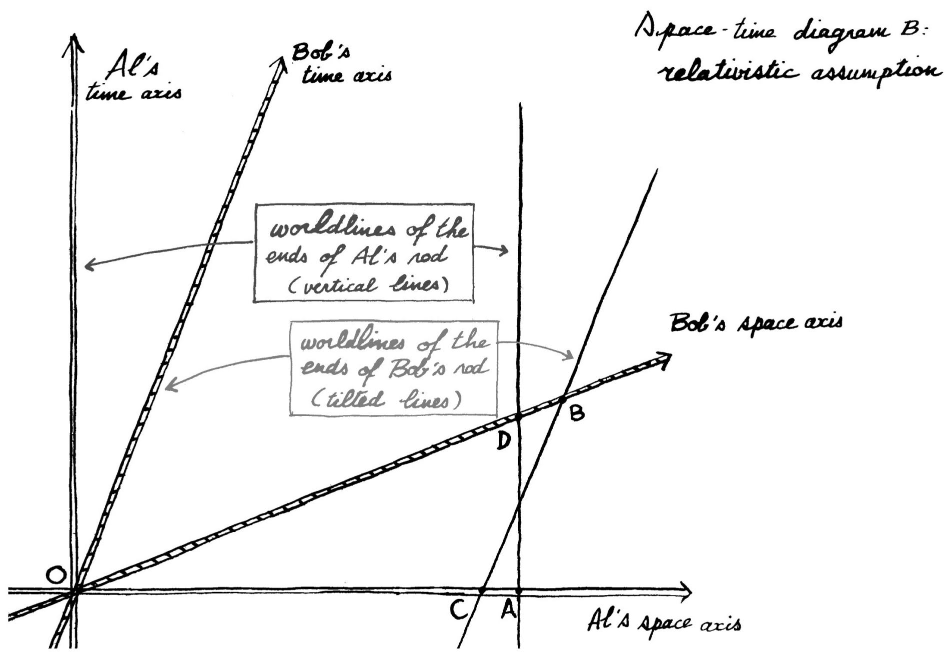

FIGURE 23. Moving rods contract.

The solution to the problem is to choose A and B in such a way that AD and BC intersect somewhere in between the space axes of Al and Bob, as shown in Fig. 23. In this case, Al and Bob both conclude that their own rod is longer than the other person's rod, Al because $OA > OC$, Bob because $OB > OD$. Length contraction thus restores the symmetry between Al and Bob.

Fig. 23 makes it clear how Al and Bob can both claim that the other person's rod is contracted. They agree that $OA > OC$ and that $OB > OD$ but they disagree about which line segments represent the lengths of the two rods, OA and OC on Al's space axis or OB and OD on Bob's.

Since the rods of Al and Bob are identical, the line segments representing the length of their own rods on their own space axes must be equally long, i.e., $OA = OB$. In Fig. 23, however, OB is longer than OA. This is the same complication that we encountered with time units in Fig. 21. One meter on a tilted space axis is represented by a larger line segment than one meter on a horizontal axis. The more tilted the axis, the larger the segment corresponding to one unit of distance.

As a result, we have to be careful comparing the length of segments of lines that are not either parallel or orthogonal to one another. The arguments in this section all turned on the comparison of the length of segments of the same line: (OP, OS) and (OQ, OR) in Figs. 20–21; (OA, OC) and (OB, OD) in Figs. 22–23. In Section 2.2, we already saw that the light postulate requires segments representing meters on a space axis and segments representing light meters on the time axis orthogonal to that space axis to be equally long.

2.5. **Euclidean geometry and the geometry of Minkowski space-time.** Fig. 24 shows, from the point of view of Al, a measuring rod of Bob passing by an identical measuring rod of Al at some constant velocity (cf. the drawing on the right in Fig. 10 in Section 1.4). The events 'shaded ends meet' and 'white ends meet' are labeled P and Q, respectively. On the right, these events are shown in a space-time diagram. (In Fig. 23, P would be the origin O and Q would be the point where AD and BC intersect.) Al's space and time axes are represented by horizontal and vertical lines, Bob's by tilted lines.

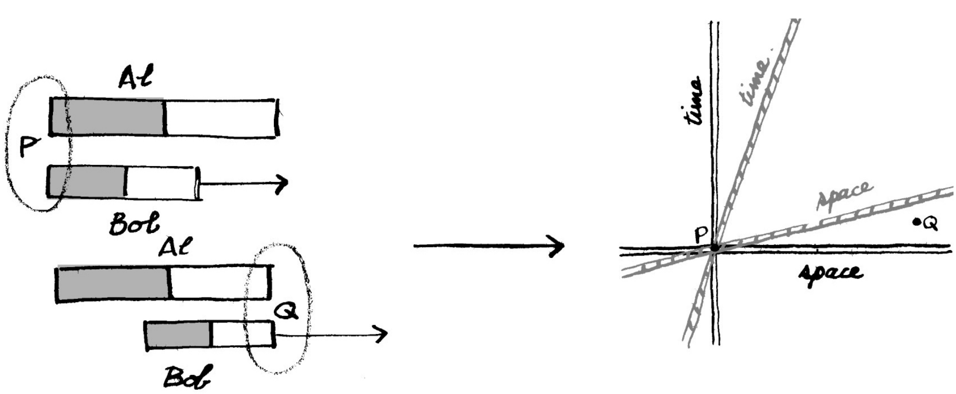

FIGURE 24. Different perspectives on two events in Minkowski space-time.

As we saw in Section 1.4, Al and Bob disagree about the order in which these two events take place. For Al, the shaded ends meet before the white ends meet; for Bob, it is just the other way around (cf. Fig. 10). This is brought out clearly in the space-time diagram on the right. In any space-time coordinate system, the space axis consists of all events simultaneous to the origin in that coordinate system. The event Q lies above Al's space axis but below Bob's. Hence, Q happens after P for Al, but before P for Bob.

This situation in space-time is closely analogous to situations in ordinary space such as the one illustrated in Fig. 25. Al and Bob are standing at the perimeter of a basin with a fountain at its center P and a duck at Q. Because they are looking at the basin from different angles, Al and Bob have a different sense of 'back-front' and 'left-right'. On the right, two sets of orthogonal 'front-back' and 'left-right' axes are shown, one for Al, one for Bob. They both choose the center of the basin as their zero point for 'front-back distance' and 'left-right distance'. They agree that the duck at Q is to the right of the fountain at P. Because of their different perspectives on the situation, however, they disagree about whether the duck is in front of the fountain or behind it. For Al, Q is in front of P; for Bob, Q is behind P. This is brought out in the coordinate systems on the right: Q is below Al's 'left-right axis' but above Bob's.

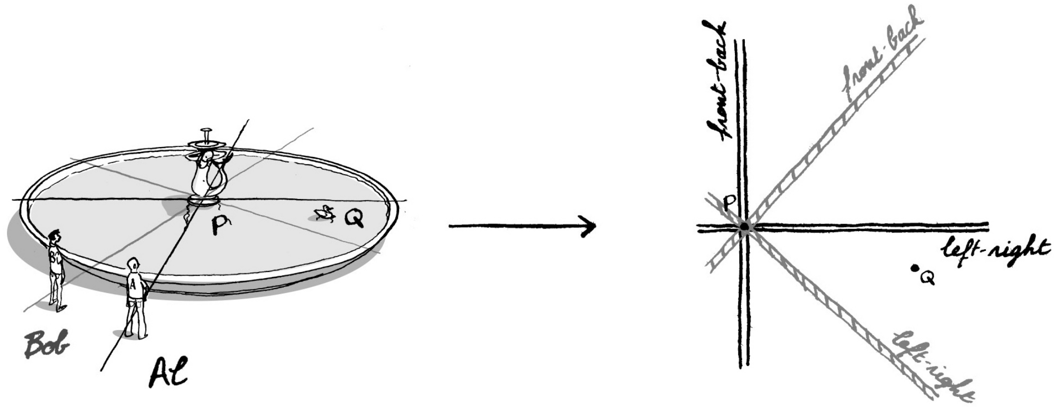

FIGURE 25. Different perspectives on two points in Euclidean space.

In Fig. 26, the diagrams on the right of Figs. 24 and 25 are shown next to each other. Consider the one on the left first. With the help of the various line segments with end-point Q perpendicular to the two sets of orthogonal axes, we can read off the 'front-back distance' and 'left-right distance' between P and Q for both Al and Bob. According to Al, Q is QR to the front and PR to the right of P; according to Bob, Q is QS to the back and PS to the right of P.

The total distance between P and Q is a combination of 'left-right distance' and 'front-back distance'. It can be computed from these two components with the help of the Pythagorean theorem. The rule for computing distances from these orthogonal components is:

$$(\text{total distance})^2 = (\text{front-back distance})^2 + (\text{left-right distance})^2.$$

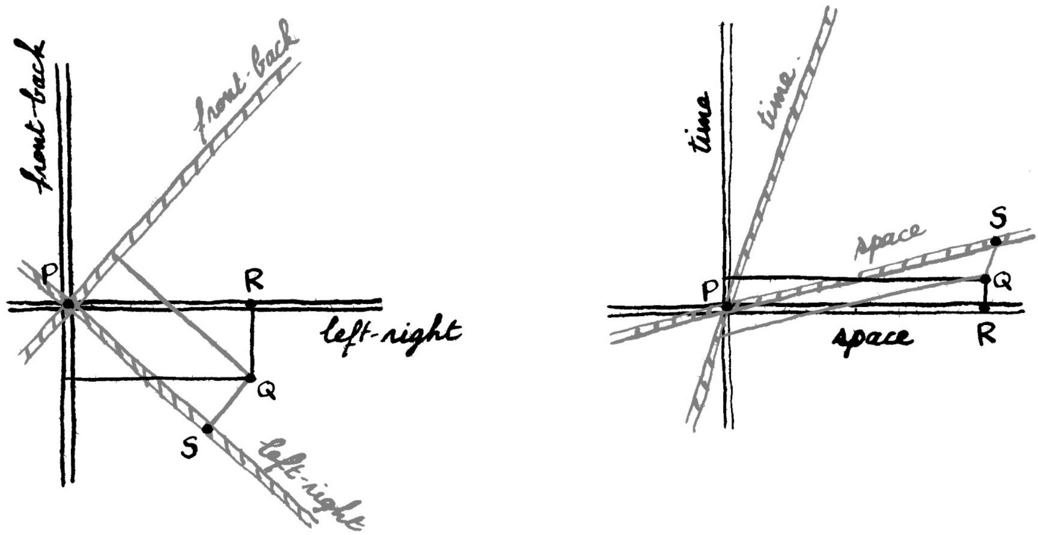

FIGURE 26. Distance in Euclidean space and in Minkowski space-time.

Al applies this rule to the right-angled triangle PRQ:

$$PQ^2 = RQ^2 + PR^2;$$

Bob applies it to the right-angled triangle PSQ:

$$PQ^2 = SQ^2 + PS^2.$$

These are two equivalent ways of computing the same distance PQ. Al and Bob agree on the total distance between P and Q; they only disagree about how the total distance breaks down into a 'front-back' and a 'left-right' component.

A similar result holds in Minkowski space-time, as is illustrated on the right in Fig. 26. As on the left, four line segments with end-point Q perpendicular to the two sets of orthogonal axes are drawn. In this case these are Al and Bob's space and time axes rather than their 'front-back' and 'left-right' axes. Al and Bob agree that the total 'spatio-temporal distance' between P and Q is the length of the line segment PQ. However, according to Al, PQ has a component QR in time and a component PR in space, whereas, according to Bob, it has a component QS in time and a component PS in space.

The question is how to compute the total spatio-temporal distance between P and Q from its spatial and temporal components given that the geometry of Minkowski space-time is not the same as Euclidean geometry. We have already encountered two features of space-time diagrams alerting us to the difference. First, the angle between the space and time axes of some inertial observer can be anywhere between 45° and 135° in a space-time diagram, but always represents a right angle in space-time (see Figs. 17–18). Second, the line segments representing one meter on a space axis and one light meter on the corresponding time axis in a space-time diagram get larger the more the angle between the space and time axes in the diagram deviates from 90° (see Figs. 21 and 23). Given these differences between Minkowski space-time and ordinary Euclidean space, it need not surprise us that we cannot use the Pythagorean theorem to compute distances in space-time. As we shall see in the next section, however, a theorem very similar to the Pythagorean theorem holds in Minkowski spacetime. As a consequence, the rule for computing distances in space and time is also very similar to the rule for computing distances in ordinary space:

$$(\text{distance in space-time})^2 = (\text{distance in time})^2 - (\text{distance in space})^2.$$

Note the minus sign. Also note that the subtraction on the right-hand side only makes sense if the quantities 'distance in time' and 'distance in space' have the same dimension, as they do when the former is measured in light meters and the latter in meters. Al applies the above rule to the right-angled triangle PRQ (in which $\angle PRQ$ is the right angle):

$$PQ^2 = RQ^2 - PR^2;$$

Bob applies it to the right-angled triangle PSQ (in which $\angle PSQ$ is the right angle):

$$PQ^2 = SQ^2 - PS^2.$$

The minus sign in the rule for computing spatio-temporal distances is responsible for all the differences between the geometry of Minkowski space-time and ordinary Euclidean geometry. Given the close similarity between the two geometries, the geometry of flat Minkowski space-time is sometimes called pseudo-Euclidean, and should be distinguished from the non-(pseudo-)Euclidean geometries of the curved space-times of general relativity.

Consider an arbitrary event O in space-time. Starting from this event, we can group all other events X in space-time in three different classes, characterized by the sign of the quantity Δs^2, the square of the spatio-temporal distance between O and X. The quantity Δs is called the *space-time interval* between O and X. Its square can be positive, negative, or zero:

$$\Delta s^2 > 0, \quad \Delta s^2 < 0, \quad \Delta s^2 = 0.$$

The corresponding regions of space-time are shown in the space-time diagram on the right in Fig. 27. Ignore the shading and the four hyperbolae for the time being, and focus on (a) the events labeled O, P, Q, R, S, and T; (b) the two sets of space and time axes, with coordinates (x, ct) and (x', ct'); and (c) the dashed lines bisecting the angle between the space axis and the time axis in both coordinate systems. These dashed lines are the world lines of light travelling in opposite directions. They form what is called the *light cone* of O. With two rather than one spatial dimension, these world lines do indeed form a double cone. The part before O is called the *past light cone*, the part after O is called the *future light cone*. Light cones play a central role in special relativity.

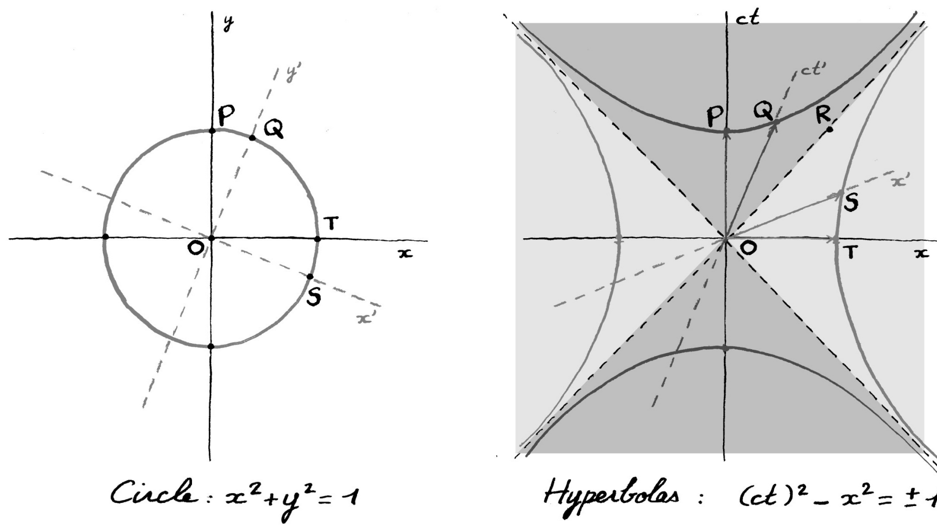

FIGURE 27. The analogy between circles in Euclidean space and hyperbolae in Minkowski space-time.

If the temporal distance between an arbitrary point X and O is *equal to* the spatial distance between X and O, then $\Delta s^2 = 0$. This is the case for all events on the light cone of O, such as the event R in Fig. 27. Events like O and R through which a world line of light can be drawn are called *lightlike connected.* The light cone of O is the set of all points that are lightlike connected to O.

If the temporal distance between X and O is *greater than* the spatial distance between X and O, then $\Delta s^2 > 0$ and the spatio-temporal distance between X and O is $\Delta s = \sqrt{\Delta s^2}$ light meters. This is the case for all events inside the light cone of O, the area with the dark shading in Fig. 27, such as the events P and Q. Pairs of events like (O, P) and (O, Q) are called *timelike connected.* For such pairs it is always possible to find an inertial observer for whom the two events happen at the same location but at different times. In the case of (O, P) this is the observer at rest in the (x, ct)-frame. In the case of (O, Q) it is the observer at rest in the (x', ct')-frame. The areas inside the past and future light cones of O are called the *absolute past* and the *absolute future* of O, respectively. Since nothing travels faster than light (see Section 2.3), no event outside the past light cone of O can have any influence on it and O can have no influence on any event outside of its future light cone. The collection of light cones of all space-time points thus determines what is called the *causal structure* of spacetime.

If the temporal distance between X and O is *less than* the spatial distance between X and O, then $\Delta s^2 < 0$ and the spatiotemporal distance between X and O is $\Delta s = \sqrt{-\Delta s^2}$ meters. This is the case for all points outside the light cone of O, the area with the light shading in Fig. 27, such as the events S and T. Pairs of points like (O, S) and (O, T) are called *spacelike connected.* For such pairs it is always possible to find an inertial observer for whom the two events happen at the same time but at different locations. In the case of (O, T) this is the observer at rest in the (x, ct)-frame. In the case of (O, S) it is the observer at rest in the (x', ct')-frame. The area outside the light cone of O is sometimes called the *absolute elsewhere* of O.

On the left in Fig. 27, a unit circle around the point O in a Euclidean plane is drawn. The center O is the origin of both the (x, y) and (x', y') coordinate systems shown in the figure. The unit circle with O as its center is the set of all point at unit distance from O. The coordinates of these points—such as P, Q, S, and T in the figure—satisfy the equation $x^2 + y^2 = x'^2 + y'^2 = 1$. The line segments OP, OQ, OS, and OT all have length 1, the radius of the unit circle.

The four hyperbolae in the space-time diagram on the right in Fig. 27 are the analogues of the unit circle on the left.

The space-time coordinates of all points at 1 light meter from O satisfy the equation $(ct)^2 - x^2 = (ct')^2 - x'^2 = 1$. These are the points on the two hyperbolae inside the light cone of O (such as P and Q).

The space-time coordinates of all points at 1 meter from O satisfy the equation $(ct)^2 - x^2 = (ct')^2 - x'^2 = -1$. These are the points on the two hyperbolae outside the light cone of O (such as S and T).

The line segments OP, OQ, OS, and OT are all of length 1 in spacetime, but in the space-time diagram in Fig. 27 OS and OQ (representing 1 meter and 1 light meter on the (x', ct')-axes, respectively) are longer than OT and OP (representing 1 meter and 1 light meter on the (x, ct)-axes, respectively). The formula for computing spatiotemporal distances introduced in this section thus confirms what we found in Section 2.4: units on tilted axes look bigger (see Figs. 21 and 23).[19]

2.6. **The analogue of the Pythagorean theorem in Minkowski space-time.** In this section, we adapt a simple geometrical construction used to prove the Pythagorean theorem in Euclidean space to prove the analogue of the theorem (with the plus sign replaced by a minus sign) in Minkowski space-time.[20]

Fig. 28 shows two right-angled triangles ABC, one in Euclidean space and one in Minkowski space-time. For the moment, focus on the Euclidean case on the left. The Pythagorean theorem says that the square of the hypotenuse equals the sum of the squares of the two sides: $AC^2 = AB^2 + BC^2$. To prove the theorem for this triangle, we proceed as follows. We first draw a line segment BD perpendicular to AC. This gives us three similar triangles: ABC, ADB, and BDC. The similarity of these three triangles can be seen upon inspection.

A formal proof of their similarity can be given with the help of the theorem that two triangles that have two angles in common are similar. We shall take this theorem to be intuitively obvious rather than derive it from the postulates of Euclidean geometry. By construction, the three triangles have a right angle in common. They also share the angle $\angle ABD = \angle ACB = \angle DCB$. The proof that the first angle is equal to the second and the third rests on another intuitively

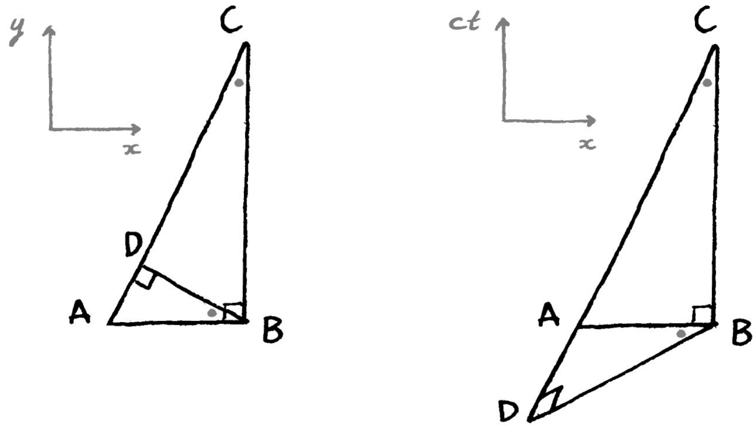

FIGURE 28. The Pythagorean theorem in Euclidean space and its analogue in Minkowski space-time.

obvious theorem of Euclidean geometry, namely that two angles are the same if the sides of these angles are mutually perpendicular. This is the case here. By construction, we have that $AB \perp BC$ and $DB \perp AC$ (where the symbol '$\perp$' stands for 'is perpendicular to').

From the similarity of the triangles ABC and ADB, it follows that

$$\frac{AB}{AC} = \frac{AD}{AB} \quad \Rightarrow \quad AB^2 = AC \cdot AD.$$

From the similarity of the triangles ABC and BDC, it likewise follows that

$$\frac{BC}{AC} = \frac{DC}{BC} \quad \Rightarrow \quad BC^2 = AC \cdot DC.$$

Adding these two relations, we find:

$$AB^2 + BC^2 = AC \cdot (AD + DC).$$

As can be seen from the figure, $AD + DC = AC$. Substituting this in the equation above, we find:

$$AB^2 + BC^2 = AC^2.$$

This proof of the Pythagorean theorem can easily be adapted to show that an analogous theorem holds for the triangle ABC in Minkowski space-time on the right of Fig. 28.

As in the Euclidean triangle on the right in Fig. 28, we construct a line segment BD perpendicular to AC. In this case, these two line segments certainly do not *look* perpendicular to one another, but that is because the geometry of the paper used to draw the diagram is different from the geometry of Minkowski space-time. In space-time, BD is in the direction of the space axis that goes with a time axis in the direction of AC. In other words, BD is tilted up from the horizontal by the same angle as AC is tilted to the right from the vertical. Hence, BD and AC

are indeed perpendicular to one another. As before, this gives us three similar triangles: ABC, ADB, and BDC. The similarity can no longer be seen upon inspection but it can be demonstrated in the exact same way as we proved the similarity of the corresponding three triangles in the Euclidean case. The proof rests on theorems that are valid both in Euclidean geometry and in the geometry of Minkowski space-time, namely that two triangles are similar if they have two angles in common and that two angles are equal if their sides are mutually perpendicular. On the basis of the similarity of the triangles ABC, ADB, and BDC, we can write down the exact same relations that we found in the Euclidean case:

$$AB^2 = AC \cdot AD, \quad BC^2 = AC \cdot DC.$$

For the Euclidean triangle ABC on the left in Fig. 28 we arrived at the Pythagorean theorem by *adding* these two relations and using that $AD + DC = AC$. For the triangle ABC on the right in Fig. 28, we arrive at the analogue of the Pythagorean theorem in Minkowski space-time by *subtracting* the first relation from the second and using that $DC - AD = AC$:

$$BC^2 - AB^2 = AC \cdot (DC - AD) = AC^2.$$

This is the basis for the rule for computing spatio-temporal distances in given in Section 2.5:[21] (distance in space-time)^2 = (distance in time)^2 − (distance in space)^2.

2.7. **The twin paradox.** Consider Figs. 29 and 30. For the time being, ignore the dashed lines and the points M and N in Fig. 29, and focus on the line segments SUR and SAR, the world lines of the identical twin sisters Suzy and Sara. Up to point S [for separation], the world lines of Suzy and Sara coincide. At S, the twins get separated. Suzy remains in the same state of inertial motion she has been in all along. Sara boards a spaceship that takes her away from her twin sister at about 75% the speed of light. At A [for acceleration], the spaceship turns around and starts moving back to Suzy, again at 75% the speed of light. At R [for reunion] the sisters are reunited. For Suzy, 9 years have passed (see the marks on Suzy's world line in Fig. 29 and the marks on her face in Fig. 30). For Sara, only 6 years have passed (see the marks on Sara's world line in Fig. 29 and her youthful appearance in Fig. 30).[22] Note that the line segments representing one year on the world line of Sara in Fig. 29 are larger than the line segments representing one year on the world line of Suzy. This is in accordance with what we found in Sections 2.4–2.5.

The scenario illustrated in Figs. 29–30 is known as the 'twin paradox'. As the phrase suggests, many people have been baffled by this prediction of special relativity (it has successfully been tested with identical atomic clocks instead of twins). Here is how people tend to get confused. They argue as follows: "Suzy and Sara are two equivalent observers who will both find that the clocks and metabolism of the other run slow, so that the other should age more slowly than she herself. So, according to Suzy, Sara should be younger when they are reunited,

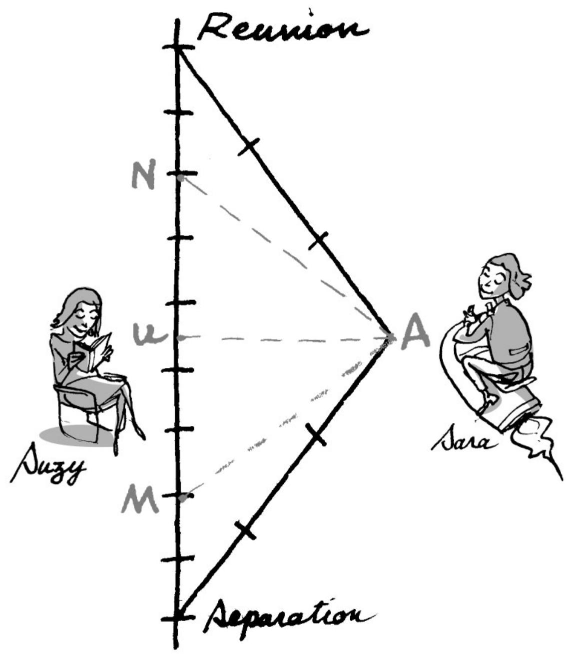

FIGURE 29. Space-time diagram for the twin paradox.

whereas, according to Sara, Suzy should be. Clearly, they cannot both be younger than the other. Hence, there must be some other factor that ensures that they are equally old when reunited." As we just saw, special relativity tells us that Sara will be younger than Suzy. So, what is wrong with the argument above? It is the very first step that is wrong. Suzy and Sara are not two equivalent observers. Suzy is an inertial observer, Sara is not. At A Sara switches from one inertial state of motion to another. She accelerates.

But, one might object, just as Sara accelerates with respect to Suzy, Suzy accelerates with respect to Sara. Why can they not both claim that the other is accelerating? The answer is that acceleration, unlike velocity, is an absolute, not a relative affair. Perhaps the easiest way to see this is to compare how Suzy and Sara experience the moment that they start to approach each other again. For Suzy there is nothing special about this moment. Sara, on the other hand, will feel it in her stomach. And if the spaceship were to turn around as abruptly as suggested by Fig. 29, the G-forces would kill her. So, Suzy and Sara are not equivalent observers. Sara accelerates and Suzy does not. Therefore, it need not surprise us that Suzy and Sara age a different number of years.

There has been a great deal of confusion about the role of the acceleration at A. Some have argued that Sara ages 6 years during the acceleration so that she ends up being as old as Suzy when the twins are reunited. Others have argued that the acceleration causes the difference in aging of the twins. Both claims are

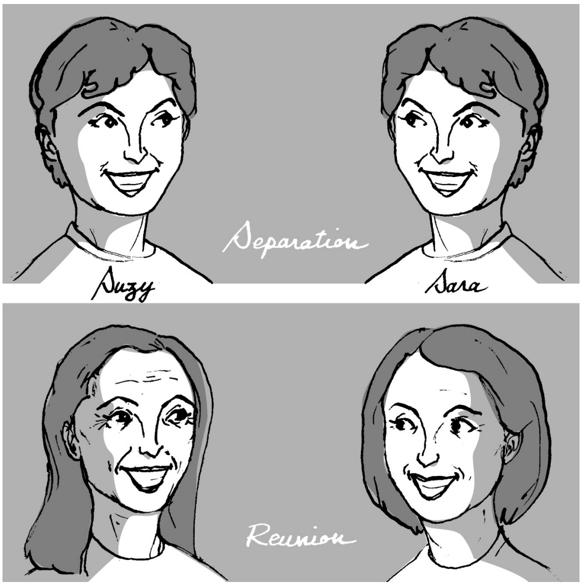

Figure 30. Suzy and Sara at separation and reunion.

erroneous. Here is a simple argument showing that the acceleration cannot be the cause of the difference in aging. Let Sara make a second journey. Again she takes off at 75% the speed of light, but now she keeps going for 6 rather than 3 years before reversing course. If it were the acceleration that caused the difference in aging, Sara should gain the same number of years on Suzy on both trips. After all, the acceleration part in both trips is exactly the same. According to special relativity, however, Sara gains 3 years during the first trip and 6 years during the second. Hence, the acceleration does not cause the difference in aging. But then again, what does?

The answer is that it is simply a matter of the space-time trajectory SAR between S and R being shorter than the trajectory SUR. The situation is analogous to the following situation in ordinary space illustrated in Fig. 31. Suppose Suzy and Sara drive from Pittsburgh to Minneapolis. The figure illustrates the two different routes they take. The first is an essentially straight highway from Pittsburgh to Minneapolis going through Chicago. This route is the analogue of the route SUR in space-time. The other is a terrible detour consisting of an essentially straight highway from Pittsburgh to Memphis and another essentially straight highway from Memphis to Minneapolis. This second route corresponds to the route SAR in space-time. The twin taking this route will not drive straight the whole way. She needs to make a sharp right turn in Memphis. This right turn is the analogue of the acceleration at A in Fig. 29. The second route is clearly longer than the first. This is a minor disanalogy with the situation in space-time: in ordinary space a straight line is the shortest route between two points, in space-time a straight line is the longest route between two points.

FIGURE 31. A spatial analogue of the twin paradox.

Suppose we do not know at first whether Suzy or Sara took the detour. Then we are told that Sara made a sharp right turn at one point whereas Suzy did not. This tells us immediately that Sara went from Pittsburgh to Minneapolis the long way. Similarly, knowing that Sara accelerated at some point in her space-time journey, whereas Suzy did not, tells us that Sara went from separation to reunion the short way. So the acceleration serves to pick out which one of the twins took the shortest trip in space-time, just as 'making a sharp right turn' serves to pick out which one of the twins took the longest trip in space. Nobody in his right mind, however, would claim that the right turn in Memphis causes the difference between Suzy and Sara's odometer readings upon arrival in Minneapolis. What causes the difference is that the route from Pittsburgh to Minneapolis through Memphis is longer than the one through Chicago. The acceleration at A likewise does not cause the age difference between Suzy and Sara. It is just that the space-time journey SAR is shorter than SUR.

To show that the acceleration really does not play a crucial role, we change the logistics of the experiment in such a way that we can establish the age difference between Suzy and Sara at R avoiding the acceleration at altogether. To this end we introduce the inertial observers Al and Bob. Al's world line is a straight line through S and A. It coincides with Sara's on the first half of her journey (SA). Bob's world line is a straight line through A and R and coincides with Sara's on the second half of her journey (AR). Al times the first half of the journey, Bob the second. Al and Bob both find that they accompanied Sara for 3 years. So, the space-time journey takes 6 years. Inertial observer Suzy already told us that the space-time journey takes 9 years. When the twins are reunited, Sara is 3 years younger than Suzy. Notice that we reached this conclusion using inertial observers only.

This maneuver of, in effect, replacing Sara with the inertial observers Al and Bob seems to leave us with another puzzle. Look at the first half of the journey. According to Suzy, Al ages only 3 years between S and A, which for her is $4\frac{1}{2}$ years. Al and Suzy are two fully equivalent inertial observers. If Suzy finds that Al ages at 2/3 her own rate, Al finds that Suzy ages at 2/3 his rate. So, according to Al, Suzy only ages 2 years between S and A, which for him is 3 years. A similar result is found when we compare Suzy and Bob on the second half of the journey. According to Bob, Suzy ages only 2 years between A and R, which for Bob is 3 years. Combining Al and Bob's conclusions, it looks as if Suzy only ages 4 years between S and R. However, we already know that Suzy ages 9 years. What is going on?

The dashed lines and the points marked M and N in Fig. 29 help us answer this question. Al and Bob fail to take into account 5 years of Suzy's life between S and R because of their disagreement about simultaneity. According to Al, A is simultaneous with M, whereas, according to Bob, A is simultaneous with N. This observation restores the consistency of the whole story. For Al, Suzy aged 2 years (between S and M) during the 3 years that he accompanied Sara from S to A. Likewise, for Bob, Suzy aged 2 years (between N and R) during the 3 years that he accompanied Sara from A to R. The 5 years that Suzy aged between M and N should be added to the 4 years counted by Al and Bob, giving a total of 9 years for Suzy's aging between S and R.

The upshot then is that there is nothing mysterious about the twin paradox. All there is to it is that the time that elapses for an observer between two points in spacetime depends on the space-time trajectory that takes her from one to the other, just as the distance she covers between two points in space depends on the route she takes to go from one to the other.

NOTES

[1] Given that the earth is rotating on its own axis, a frame at rest with respect to the earth is only approximately an inertial frame. A frame at rest with respect to the sun would be an even closer approximation to a true inertial frame.

[2] See the chapter "No success like failure ..." in this volume for a discussion of Einstein's unsuccessful attempts to generalize the relativity principle from uniform to arbitrary motion.

[3] In fact, Einstein made two more assumptions to the effect that space and time are the same everywhere. There are no special times and locations (space and time are it homogeneous) and no special directions (space and time are *isotropic*). Both assumptions are routinely granted.

[4] For a reconstruction of Einstein's path toward special relativity, see Norton's contribution to this volume.

[5] Note that Al is moving at about half the speed of light. To bring out the effects more clearly the speeds of the observers here and in subsequent figures have been grossly exaggerated.

[6]To bring out the effects more clearly, time differences here and in subsequent figures have been grossly exaggerated.

[7]'fast' is meant in the sense of 'out of sync' here, not in the sense of 'ticking at a higher rate'. This distinction becomes important in Section 1.3: relativity of simultaneity and time dilation are two different phenomena.

[8]As we shall see in Section 1.4, the analysis gets more complicated when the velocity is not perpendicular to L.

[9]According to the Pythagorean theorem (see Section 2.6 for a simple proof), the sum of the squares of the sides of a right-angled triangle is equal to the square of its hypotenuse. Consider the velocity diagram in Fig. 8. Since the net velocity is c and the horizontal velocity is v, the vertical velocity must be $c\sqrt{1-v^2/c^2}$ to satisfy the Pythagorean theorem: $v^2+c^2(1-v^2/c^2)=c^2$.

[10]To bring out the symmetry between Al and Bob more fully, we can attach clocks to the ends of Al's rod as well, one with a shaded front to the shaded end of his rod, one with a white front to the white end. Al makes sure these clocks are properly synchronized. Once the reading on one of Al's clocks in one of the four situations shown in Fig. 10 is fixed, all others can be inferred on the basis of the assumptions (a)–(c) made in constructing Fig. 10. If Al's white clock, like Bob's white clock, reads 12:00 when the white ends of the rods meet, the readings on Al's two clocks in the four drawings in Fig. 10 (upper-left, lower-left, upper-right, lower-right) will be as follows:

	left :		right :	
	shaded :	white :	shaded :	white :
upper :	11 : 35	12 : 00	11 : 45	11 : 45
lower :	11 : 45	12 : 10	12 : 00	12 : 00.

[11]In 1887, Michelson and Morley tried to detect the earth's presumed motion through the ether by checking whether the travel time of a light signal going back and forth between two mirrors depends on whether the line connecting these mirrors is parallel or perpendicular to the direction of motion. In accordance with the considerations above, they found that it does not. Independently of one another, FitzGerald and Lorentz suggested a few years later that this negative result can be accounted for by assuming that objects moving through the ether contract in the direction of motion by the same factor we just found (Brown 2005, Ch. 4). The Michelson-Morley experiment has been a staple of popular expositions of special relativity and has great pedagogical value, even though historians now agree that it played no role in Einstein's development of the theory (Holton 1969, Stachel 1982, Janssen 2002b, and Norton's contribution to this volume).

[12]Think of timing a runner on the 100 m by subtracting the time read on a clock at the finish from the time read on a clock at the start. If the clock at the start is fast compared to the clock at the finish, the difference needs to be added to the time found in this manner, otherwise average runners would seem to break the world record.

[13]Note that $u_{\text{corrected}}=-v$ for $u=-v$: Al and Bob agree on the velocity with which they are moving with respect to one another.

[14]It does not follow from the addition theorem that nothing can travel faster than the speed of light, although there are good reasons for believing this to be impossible (see Section 2.3 below).

[15]Inserting $u=c$ into the formula for $u_{\text{corrected}}$, we find:

$$u_{\text{corrected}}=\frac{c\left(1-v^2/c^2\right)}{1+v/c}=\frac{c(1+v/c)(1-v/c)}{1+v/c}=c-v.$$

.

[16]The light *meter* as a unit of *time* is fully analogous to the more familiar light *year*, the distance traveled by light in a year, as a unit of *distance* (Joan Baez thus made a category

mistake in the line "a couple of light years ago" of her song *Diamonds and Rust*). We could likewise introduce the 'light second', the distance traveled by light in one second, as our unit of distance. One light second is about $3 \cdot 10^8$ meters. The light second is thus a much larger unit of distance than the meter. To convert meters to light seconds, we need to divide by $c \approx 3 \cdot 10^8$ m/s. The result of dividing meters by meters per seconds is seconds. If distances are measured in light seconds and time in seconds, then these two quantities once again have the same dimension and can be added to one another. Moreover, the velocity of light in these units is once again equal to one. From the definition of a light second, it follows that light travels one light second per second.

[17]The world line OQ of the light signal from O to Q bisects the right angle between Bob's time and the space axes. For Bob, as for Al, the world lines of light are 45^{o} lines.

[18]For further discussion, see Janssen 2002b and Norton's contribution to this volume.

[19]The similarities between Euclidean space and Minkowski space-time led Minkowski to an important conclusion: that the laws of physics stay the same when we switch to a different set of orthogonal axes in space-time corresponding to a different state of inertial motion is no more surprising than that the laws of physics stay the same when we switch to a different set of orthogonal axes in space pointing in different directions. As Minkowski pointed out, in Newtonian mechanics the statement that uniform motion does not make a difference for the laws of physics is something completely different from the statement that orientation in space does not make a difference for the laws of physics. In special relativity, these two statements turn out to be intimately connected. Switching to a different inertial frame and switching to a rotated set of space axes are both nothing but a change of perspective in spacetime. Neither affects the laws of physics.

[20]The argument in this section is due to Jon Dorling.

[21]The triangle on the right in Fig. 28 is made up of the timelike line segments AC and BC and the spacelike line segment AB. A completely analogous argument can be given for the case of a right-angled triangle made up of one timelike line segment and two spacelike ones.

[22]Since Sara's velocity is $v = .75c$, she ages at $\sqrt{1 - v^2/c^2} \approx \frac{2}{3}$ the rate of Suzy.

References

Brown, Harvey R. (2005). *Physical Relativity. Space-time Structure from a Dynamical Perspective.* Oxford: Oxford University Press.

CPAE 2: Stachel, John , Cassidy, David C., Schulmann, Robert, and Renn, Jürgen, eds. (1989). *The Collected Papers of Albert Einstein.* Vol. 2. *The Swiss Years: Writings, 1900–1909.* Princeton: Princeton University Press.

Einstein, Albert (1905r). "Zur Elektrodynamik bewegter Körper." *Annalen der Physik*, 17, 891–921 (CPAE 2, Doc. 23). English translation in Einstein *et al.* (1952, 37–65).

Einstein, Albert, Lorentz, Hendrik Antoon, Minkowski, Hermann, and Weyl, Hermann (1952). *The Principle of Relativity.* New York: Dover.

Holton, Gerald (1969). "Einstein, Michelson, and the "Crucial" Experiment." *Isis* 60: 133–197. Reprinted as ch. 9 in Holton (1973) and as ch. 8 in Holton (1988).

Holton, Gerald (1973). *Thematic Origins of Scientific Thought. Kepler to Einstein.* Cambridge: Harvard University Press.

Holton, Gerald (1988). *Thematic Origins of Scientific Thought. Kepler to Einstein.* Rev. Ed. Cambridge: Harvard University Press.

Janssen, Michel (2002b). “Reconsidering a Scientific Revolution: the Case of Lorentz versus Einstein.” *Physics in Perspective* 4: 421–446.

Minkowski, Hermann (1909). “Raum und Zeit.” *Physikalische Zeitschrift* 20, 104–11. English translation in Einstein *et al.* (1952, 75–91).

Pais, Abraham (1982). *‘Subtle is the Lord ... ’ The Science and the Life of Albert Einstein.* Oxford: Oxford University Press.

Stachel, John (1982). “Einstein and Michelson: the Context of Discovery and the Context of Justification.” *Astronomische Nachrichten* 303 I: 47–53. Reprinted in Stachel (2002, 177–190).

Stachel, John, ed. (1998). *Einstein's Miraculous Year: Five Papers that Changed the Face of Physics.* Princeton: Princeton University Press.

Stachel, John (2002a). *Einstein from ‘B’ to ‘Z’* (*Einstein Studies*, Vol. 9). Boston: Birkhäuser.