1 Kinematics of Special Relativity

The special theory of relativity arose out of a conflict between the ideas of mechanics as developed by Galileo and Newton, and the ideas of electromagnetism. For this reason relativity is often discussed after electromagetism is developed. However, the the connections between relativity and electromagnetism are high-level and difficult to understand the first time around. Furthermore, special relativity is actually a generally valid extension to Newtonian mechanics that is needed when objects move at very high speeds, and is only coincidentally related to electromagnetism. For this reason I choose to discuss relativity before electromagnetism.

The only fact from electromagnetism that we need is introduced now: There is a maximum speed at which objects can travel. This is coincidentally equal to the speed of light in a vacuum, $c = 3 \times 10^8$ m s^{-1}. Furthermore, a measurement of the speed of a particular light beam yields the same answer regardless of the speed of the light source or the speed at which the measuring instrument is moving.

This rather bizarre experimental result is in contrast to what occurs in classical mechanics. If two cars pass a pedestrian standing on a curb, one at 20 m s^{-1} and the other at 50 m s^{-1}, the faster car appears to be moving at 30 m s^{-1} relative to the slower car. However, if a light beam moving at 3×10^8 m s^{-1} passes an interstellar spaceship moving at 2×10^8 m s^{-1}, then the light beam appears to be moving at 3×10^8 m s^{-1} to occupants of the spaceship, not 1×10^8 m s^{-1}. Furthermore, if the spaceship beams a light signal forward to its (stationary) destination planet, then the resulting beam appears to be moving at 3×10^8 m s^{-1} to instruments at the destination, not 5×10^8 m s^{-1}.

The fact that we are talking about light beams is only for convenience. Any other means of sending a signal at the maximum allowed speed would result in the same behavior. We therefore cannot seek the answer to this apparent paradox in the special properties of light. Instead we have to look to the basic nature of space and time.

1.1 Spacetime Thinking

In order to gain an understanding of special relativity it is important to begin thinking of space and time as being different dimensions of a four-

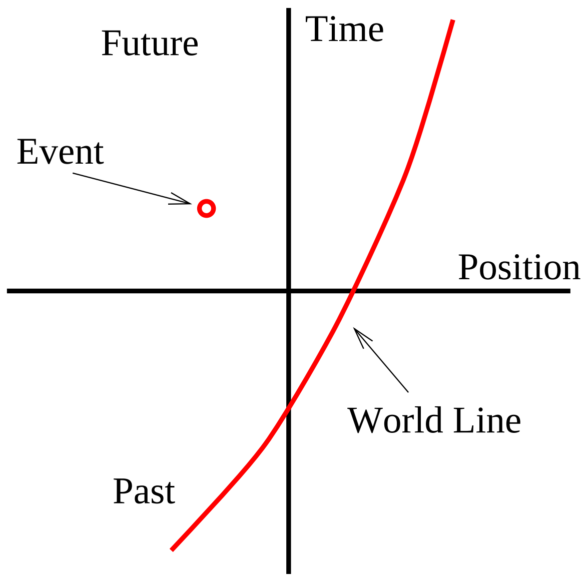

Figure 1: Spacetime diagram showing an *event* and a *world line*.

dimensional space called spacetime. Actually, since we can't visualize four dimensions very well, it is easiest to start with only one space dimension and the time dimension. Figure 1 shows a graph with time plotted on the vertical axis and the one space dimension plotted on the horizontal axis. An *event* is something that occurs at a particular time and a particular point in space. ("Julius X. wrecks his car in Lemitar, NM on 21 June at 6:17 PM.") A *world line* is a plot of the position of some object as a function of time (more properly, the time of the object as a function of position) on a spacetime diagram. Thus, a world line is really a line in spacetime, while an event is a point in spacetime.

In a spacetime diagram the slope of a world line has a special meaning. Notice that a vertical world line means that the object it represents does not move — the velocity is zero. If the object moves to the right, then the world line tilts to the right, and the faster it moves, the more the world line tilts. Quantitatively, we say that

$$velocity = \frac{1}{slope\ of\ world\ line}. \tag{1}$$

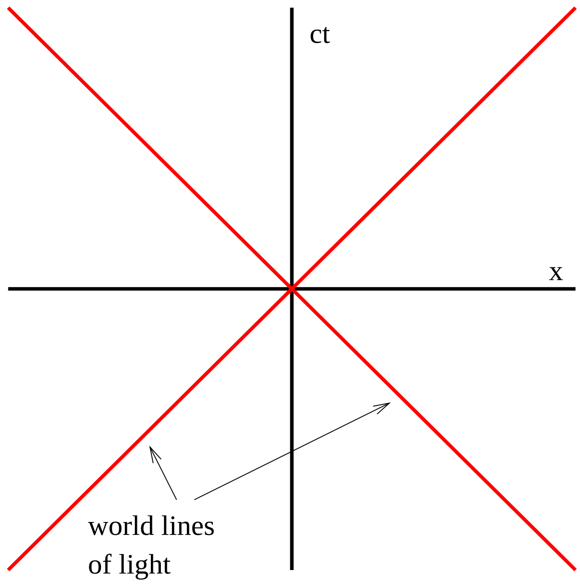

Figure 2: Scaled spacetime diagram showing world lines of light passing left and right through the origin.

Notice that this works for negative slopes and velocities as well as positive ones. If the object changes its velocity with time, then the world line is curved, and the instantaneous velocity at any time is the inverse of the slope of the tangent to the world line at that time.

The hardest thing to realize about spacetime diagrams is that they represent the past, present, and future all in one diagram. Thus, spacetime diagrams don't change with time — the evolution of physical systems is represented by looking at successive horizontal slices in the diagram at successive times. *Spacetime diagrams represent evolution, but they don't evolve themselves.*

If time and position are simply different dimensions of the same abstract space, then they should have the same units. The easiest way to arrange this is to multiply time by the maximum speed, c, resulting in the kind of spacetime diagram shown in figure 2. Notice that world lines of light have slope ± 1 when the time axis is scaled this way. Furthermore, (1) must be revised to read

$$v = \frac{c}{slope}. \tag{2}$$

1.2 The Principle of Relativity

The *principle of relativity* states that the laws of physics are the same in all *inertial* reference frames. An inertial reference frame is one that is not accelerated. Reference frames attached to a car at rest and to a car moving steadily down the freeway at 30 m s^{-1} are both inertial. A reference frame attached to a car accelerating away from a stop light is not inertial. Newton's second law, $F = ma$, works in both stationary and moving inertial frames, but not in accelerated frames. Newton's law can be fixed up in non-inertial frames by adding an additional *inertial force* equal to minus the mass times the acceleration of the reference frame to the left side of $F = ma$. However, this is just a fudge factor which is invoked to preserve appearances — the fact remains that $F = ma$ in unmodified form doesn't work in an accelerated reference frame.

The principle of relativity is not cast in stone, but is an educated guess or hypothesis based on extensive experience. If the principle of relativity weren't true, we would have to do all our calculations in some preferred reference frame in which we knew $F = ma$ to be true. This would be very annoying. However, the more fundamental problem is that we have no idea what the velocity of this preferred frame might be. Does it move with the earth? That would be very earth-centric. How about the velocity of the center of our galaxy or the mean velocity of all the galaxies? Rather than face the issue of a preferred reference frame, physicists have chosen to stick with the principle of relativity.

Notice that the constancy of the speed of light in all reference frames is consistent with the principle of relativity. (Why???) However, as noted above, it is inconsistent with our notions as to how velocities add, or alternatively, how we think the world should look from reference frames moving at different speeds. The classical way of understanding the view from different reference frames is called "Galilean relativity". The new way that reconciles the behavior of objects moving at very high speeds is called "Einsteinian relativity". Einstein's great contribution was to discover the laws that tell us how the world looks from reference frames moving at high speeds relative to

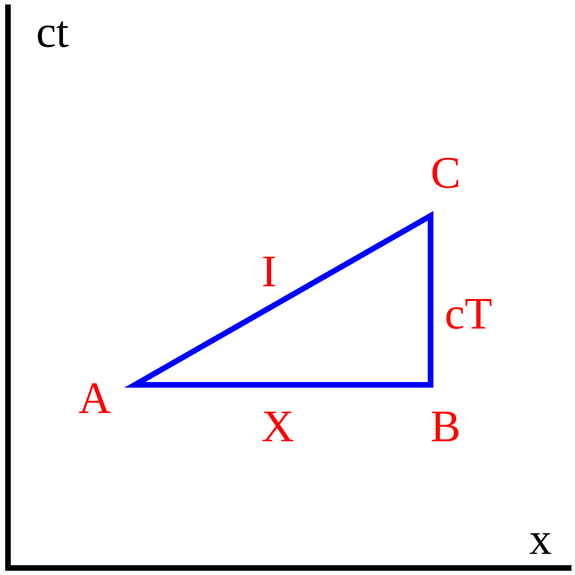

Figure 3: Triangle for Pythagorean theorem in spacetime.

each other. These laws constitute a *geometry of spacetime*, and from them all of special relativity can be derived.

1.3 Postulates of Special Relativity

All of the observed facts about spacetime can be derived from two postulates:

- Spacetime obeys a modified Pythagorean theorem, which gives the distance, I, in spacetime as

$$I^2 = X^2 - c^2T^2, \tag{3}$$

  where X, T, and I are defined in figure 3.

- Whether two events appear to be simultaneous depends on the coordinate system from which they are viewed.

Let us discuss these postulates in turn.

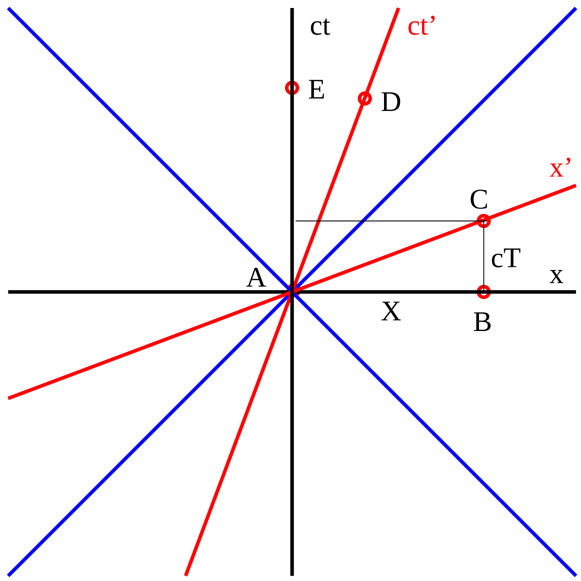

Figure 4: Sketch of coordinate axes for a moving reference frame, x', ct'. The meanings of the events A-E are discussed in the text. The lines tilted at $\pm 45°$ are the world lines of light passing through the origin.

The Pythagorean theorem of spacetime differs from the usual Pythagorean theorem in two ways. First, the vertical side of the triangle is multiplied by c. This is a trivial scale factor that gives time the same units as space. Second, the right side of (3) has a minus sign rather than a plus sign. This highlights a fundamental difference between spacetime and the ordinary xyz space in which we live. Spacetime is said to have a *non-Euclidean* geometry — in other words, the normal rules of geometry that we learn in high school don't always work for spacetime!

The main consequence of the minus sign in (3) is that I^2 can be negative and therefore I can be imaginary. Furthermore, in the special case where $X = \pm cT$, we actually have $I = 0$ even though $X,\ T \neq 0$ — i. e., the "distance" between two well-separated events can be zero. Clearly, spacetime has some weird properties!

The quantity I is usually called an *interval* in spacetime. If the two events defining the end points of the interval have the same t value, then the interval is the ordinary space distance between the two events. On the other hand, if they have the same x value, then the interval is equal to ic times the time interval between the events, where $i = (-1)^{1/2}$. This much is clear from figure 3. The interpretation of I in the more general case must await our discussion of the second postulate. However, generally speaking, if I is real, the interval is called *spacelike*, while an imaginary interval is called *timelike*.

The classical way of thinking about *simultaneity* is so ingrained in our everyday habits that we have a great deal of difficulty adjusting to what special relativity has to say about this subject. Indeed, understanding how relativity changes this concept is the single most difficult part of the theory — once you understand this, you are well on your way to mastering relativity!

Before tackling simultaneity, let us first think about *collocation.* Two events (such as A and E in figure 4) are collocated if they have the same x value. However, collocation is a concept that depends on reference frame. For instance, George is driving from Socorro to Albuquerque. Just as he passes Escondida he sneezes (event A in figure 4). As he drives by Belen, he sneezes again (event D). In the reference frame of the earth, these two sneezes are not collocated, since they are separated by about forty miles. However, in the reference frame of George's car, they occur in the same place — assuming that George hasn't left the driver's seat!

Notice that any two events separated by a timelike interval are collocated in *some* reference frame. The speed of the reference frame is given by (2), where the slope is simply the slope of the world line connecting the two events.

Classically, if two events are simultaneous, we consider them to be simultaneous in all reference frames. For instance, if two clocks, one in New York and one in Los Angeles, strike the hour at the same time in the earth reference frame, then these events also appear to be simultaneous to instruments in the space shuttle as it flies over the United States. However, if the space shuttle is moving from west to east, i. e., from Los Angeles toward New York, careful measurements will show that the clock in New York strikes the hour *before* the clock in Los Angeles!

Just as collocation depends on one's reference frame, this result shows that simultaneity also depends on reference frame. Figure 4 shows how this works. In figure 4 events A and B are simultaneous in the rest or unprimed reference frame. However, in the primed reference frame, events A and C are simultaneous, and event B occurs at an earlier time. If A and B correspond

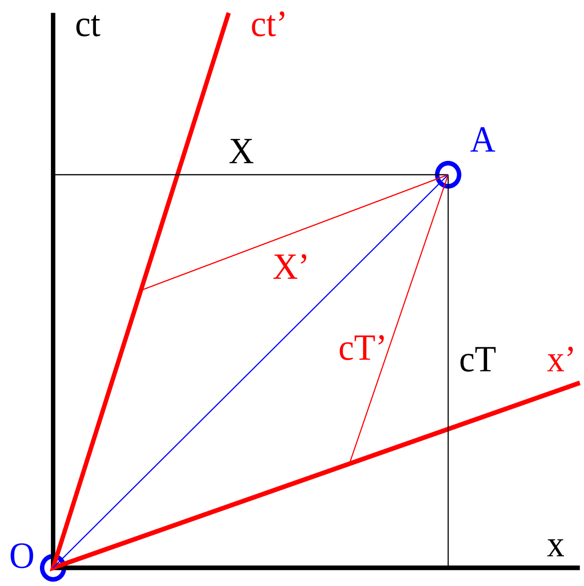

Figure 5: Resolution of a spacetime interval into different reference frames.

to the clocks striking in Los Angeles and New York respectively, then it is clear that B must occur at an earlier time in the primed frame if indeed A and C are simultaneous in that frame.

The tilted line passing through events A and C in figure 4 is called the *line of simultaneity* for the primed reference frame. Its slope is related to the speed, U, of the reference frame by

$$slope = U/c. \tag{4}$$

Notice that this is the inverse of the slope of the world line attached to the primed reference frame. There is thus a symmetry between the world line and the line of simultaneity of a moving reference frame — as the reference frame moves faster to the right, these two lines close like the blades of a pair of scissors on the 45° line.

1.4 Geometry of Spacetime

Figure 5 shows an interval I between the events O and A in spacetime resolved into space and time components in two different ways. In the unprimed

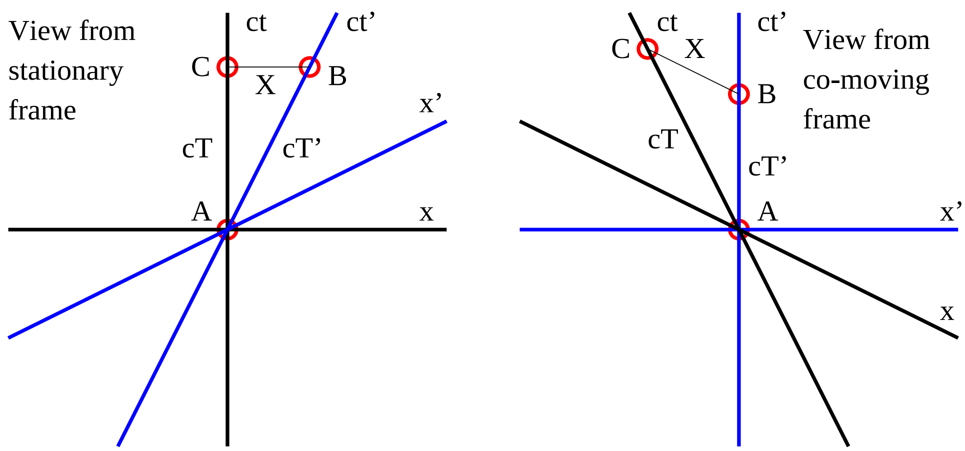

Figure 6: Two views of the relationship between three events, A, B, and C. The left panel shows the view from the unprimed reference frame, in which A and C are collocated, while the right panel shows the view from the primed frame, in which A and B are collocated.

reference frame the space and time components are X and cT, while in the primed frame they are X' and cT'. The spacetime interval is a quantity that doesn't depend on reference frame, just as the distance between objects in ordinary space (i. e., the number of miles between Albuquerque and Socorro) doesn't depend on the coordinate system used. Thus, from (3), we have

$$I^2 = X^2 - c^2T^2 = X'^2 - c^2T'^2. \tag{5}$$

The facts that 1) I is independent of coordinate system and 2) I has special meaning in certain cases can be used to attach a meaning to I in all circumstances. For example, figure 6 shows two views of three events. In the left panel the triangle A-B-C is a right triangle oriented with the coordinate axes, with the interval I being the hypotenuse. The Pythagorean theorem of spacetime thus applies, and

$$I = (X^2 - c^2T^2)^{1/2}, \tag{6}$$

where X and T are defined in the figure. The same triangle is shown in the right panel from the point of view of a reference frame that is moving

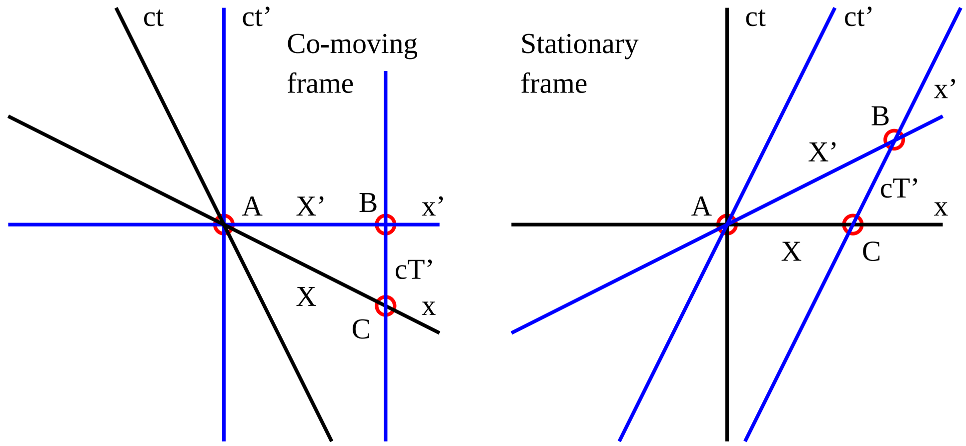

Figure 7: Definition sketch for understanding the Lorentz contraction. The parallel lines represent the world lines of the front and the rear of a moving object. The left panel shows a reference frame moving with the object, while the right panel shows a stationary reference frame.

to the right relative to the unprimed frame so as to make events A and B collocated. Notice that in this frame, the triangle A-B-C is no longer a right triangle. this is a consequence of the geometry of spacetime — angles are not preserved in transformations between reference frames in this geometry, and therefore have no particular physical meaning. However, since A and B are collocated in the primed frame, the interval between them is simply ic times the time between these two events, T'.

Rearranging (6) and using (2), which indicates that the velocity difference between the reference frames is $V = X/T$, we find

$$icT' = I = (X^2 - c^2T^2)^{1/2} = T(V^2 - c^2)^{1/2}, \tag{7}$$

which reduces to

$$T' = T[1 - (V/c)^2]^{1/2}. \tag{8}$$

This is the famous *time dilation* effect of special relativity — a moving clock appears to run more slowly than normal to a stationary observer.

A similar argument can be made to show how the postulates of relativity result in the *Lorentz contraction*. Figure 7 compares the length of a moving object measured in its own reference frame (left panel) with its length as measured in a stationary reference frame (right panel). The length of a moving object is measured by *simultaneously* measuring the positions of the front and the rear of the object and subtracting these two numbers. Events A and C correspond to these position measurements for the stationary reference frame since they are respectively on the rear and front world lines of the object. Thus, the interval AC, which is equal to X, is the length of the object as measured in the stationary frame.

In the left panel, X is the hypotenuse of a right triangle. Therefore, by the Pythagorean theorem of spacetime, we have

$$X = (X'^2 - c^2 T'^2)^{1/2}. \tag{9}$$

Now, the line passing through A and C in the left panel is the line of simultaneity of the stationary reference frame. The slope of this line is $-V/c$, where V is the speed of the object relative to the stationary reference frame. Geometrically in figure 7, the slope of this line is $-cT'/X'$, so we find by equating these two expressions for the slope that

$$T' = VX'/c^2. \tag{10}$$

Finally, combining equations (9) and (10) results in

$$X = X'(1 - V^2/c^2)^{1/2}. \tag{11}$$

This says that the length of a moving object as measured in a stationary reference frame (X) is *less* than the actual length of the object as measured in its own reference frame (X'). This apparent reduction in length is called the Lorentz contraction.