Idea 5: The Theory of Relativity

Mike Lee

October 9, 1999

1 Preamble: Theory of Relativity

The theory of relativity is enunciated in the popular assertion, *everything is relative*, and while that may or may not be true, the theory developed by Einstein is based on a good deal quantitative traditional physics. In fact, two theories are essential to the discovery of relativity - the laws of classical mechanics (Idea 2) and the laws of electricity and magnetism.

In 1900, almost everyone believed that they laws of mechanics and the laws of electricity were the crowning monuments of human understanding of the laws of nature. A few people, most notably Einstein, were troubled by small inconsistencies. These inconsistencies were to be the key to the discovery of the theory of relativity, and the now famous equation,

$$E = mc^2 .$$

To understand this famous equation only one simple fact needs to be recognized.

> *That one simple fact is that the laws of nature are the same for all persons, independent of their state of motion, as long as that motion is in a straight line at constant velocity.*

Peculiar? yes, but that was the great debate in the first decades of the twentieth century. More surprising than this debate is the discovery of relativity which results.

2 Frames of Reference

What is a reference frame? In physics, it is a matter of some precise definitions. It is not *perspective* or *viewpoint* or *where you're coming from*, but is a scientifically reliable method of providing a reference system within which scientific measurements and descriptions of physical phenomena may be unambiguously made. A frame of reference seeks to define nothing less fundamental than time and space.

> Reference frame and frame of reference are essentially the same thing. It will not be necessary for us to take care of making sure we get the subtle differences correctly. But, just for completeness we can say that a reference frame is a convention that permits the specification of the position of objects and, when needed, the time that something happens at a specific location. Since there can be many possible choices for a reference frame, we sometimes need to say that one particular one is being used. In that case we speak of one particular one as the frame of reference. That means it is the particular reference frame is the one being used to record or report a particular set of observations. So, as you can see, blurring this distinction isn't going to cause us any trouble.

2.1 Reference Frame

To establish the position of something in space, one needs a **reference frame.** Fig. 1 shows a typical example of a reference frame. In Fig. 1 there are three orthogonal axes which serve as rulers. In the figure is an arrow labeled r which points from the origin (where the three rules begin) to a particular point in space that we may be wanting to refer to. So, we can refer to it very specifically by specifying its position with three numbers corresponding to the distances that you would have to travel in the three direction specified by the three rulers labeled x, y, z .

The reference frames that we use in everyday life are not labeled with x, y and z. It is more commonly North, East and Up. We also have their opposites, i.e. South, West and Down, but for most reference systems giving negative coordinate values means going in the opposite direction. For instance, we could say that a town named Ravenna is five miles east and a town named Akron is negative twelve miles east instead of saying it is west.

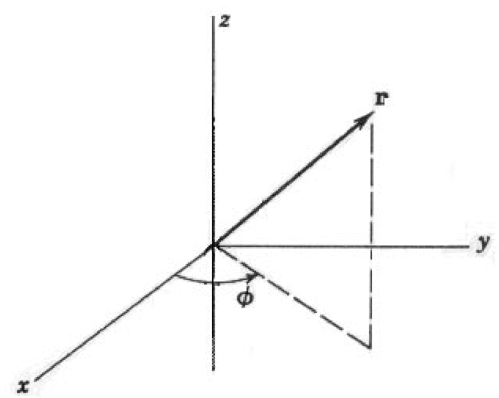

Figure 1: A Coordinate System Constructed from Three Orthogonal Axes

In such a discussion, we would tacitly be assuming that our coordinate system started where we were located. If we were writing a book, we would have to say where these places are located relative to Kent, Ohio because the reader would not necessarily be where we are or know where we were when this discussion occurred.

The proper way to locate the top of terminal tower in Cleveland would be first to state the our origin of our coordinate system is in Kent, Ohio on the first floor of the physics building. Then we can say the top of terminal tower is negative 12 miles east, 25 miles north and .1 miles up. While that is not very accurate, it illustrates the use of three numbers which we might as well have called the x,y,z coordinates of the top of terminal tower in our coordinate system.

2.2 Time as a Dimension

Not only is it important to know where something happened, it is essential to know when it happened. All of the laws of physics involve space and time. Thus, it what follows, we will describe events by their location and time. There are three dimensions to space and if we add to everything we describe the time, then we can treat time as a dimension too. This turns out to be essential and is the reason why we sometimes say that we live in a space-time continuum which is four dimensional. It is not necessary that you accept this statement that we are in a four dimensional existence - the laws of physics can be understood without this convention. You may, as we go along, decide that it is a lot more convenient to describe space-time as four dimensional.

3 Galilean Relativity

Before Einstein, there was Galileo. Galileo would have told us that the transformation laws between observers in different inertial reference frames were very simple. If two observers such as shown in Fig. 2, wanted to exchange information about positions and times of things they saw, they would only have to know their relative velocities and agree upon a origin.

A figure showing the two inertial Galilean observers moving along in two reference frames.

Figure 2: A Galilean View of Relative Motion

Fig. 2 shows the case when two observers move at a relative velocity v in the x-direction and set their clocks to agree that time starts at $t = 0$ when the origins of their coordinate systems pass. If we call the one on the left, Observer L, and the one on the right, Observer R, then.

$$t_R = t_L$$

$$x_L = x_R + vt_R$$

$$y_L = y_R$$

$$z_L = z_R$$

These are the equations that we use in everyday life. We can translate any physical quantity from one Galilean reference frame to another quite readily. For instance velocity. If the observer in reference frame L throws a volley ball in the x direction with a velocity u, then as observed by the R person the velocity is actually $u + v$. We say then that the velocity transformation laws are

$$u_R = u_L + v\,.$$

As simple as this seems, it is wrong. Indeed, all these transformation laws are wrong. We don't notice it in every day life, but if the velocities were large, say near the speed of light, then we would observe very striking changes. For that reason. Einstein developed the theory of special relativity (Sec. 5).

Before leaving Galilean relativity, we must appreciate that the errors in our transformation laws are something that we as individuals will never see. Indeed, as we will discover errors in the classical laws of mechanics as Newton formulated them, we will also never see a deviation from Newtonian mechanics in our normal everyday experiences because the discrepancies are so small. They are not small on the large scale of astrophysics and they are not small on the small scale of what happens to subatomic particles or electrons in large atoms like lead. We are perhaps fortunate to live in a domain where the simple classical laws of Newton are totally adequate. As we will see int he following sections, we would have to develop a new view of space and time if we had to deal with relativistic physics in our daily lives.

4 Electricity and Magnetism

Figure 3: James Clerk Maxwell

After Newton, all was right with the world, at least the scientific world, in that there was a real sense that Mother Nature was sharing her secrets and because of the outstanding success of Newtonian Mechanics, we had every right to expect similar fundamental and immutable laws to be discovered in other areas. Among of these other areas were the subjects of electricity and magnetism. Two hundred years after Newton, about 1865, James Clerk Maxwell was to write done the fundamental laws of electromagnetic theory, just like Newton wrote down the fundamental laws of mechanics. Unfortunately (or perhaps fortunately), before the end of the 19th century a crisis of sorts developed when it was discovered that these two fundamental theories of Mother Nature could not both be correct!

To understand the dilemma faced at the dawn of the 20th century, we need to understand a little about electricity and magnetism. It is a curious story and it led Einstein to the discovery of Relativity.

4.1 Electrostatics

Humankind's experience with the everyday phenomena of static electricity is, naturally, thousands of years old. Electric sparks, whether the results of lightning or rubbing a wool cloth when the humidity is very low, have long been a part of human experience. Both of these phenomena (lightning and sparks from static electricity) are fundamentally they same. They are made by charges moving through the air.

Positive and Negative Charge

In class we demonstrated that objects such as fur and rubber could be charged when subbed together. We showed how this charge could be transferred to pith balls. We observed that there are to be two kinds of charge. Ben Franklin is credited with giving the two types the names positive and negative. We observed, as did Ben, that when two pith balls have the same charge, they repel and when they have opposite charges, they attract.

Over the years, people learned how to measure the amount of charge. One can imagine that this started with observing that the more one rubbed the rubber rod against the fur, the greater the force between the pith balls. Must more sophisticated methods were developed when led to a determination of how the force relates to the amount of charge. If an object has an amount of charge, saw q_1, and another object has an amount of charge, say q_2, then the force between these two objects is given by **Coulombs Law**,

$$F = k\frac{q_1 q_2}{r^2}\,. \tag{1}$$

In this law, the distance between he objects is labeled r. We see that this is an example of an **inverse square law**, meaning that the dependence of the force in inverse to the square of the distance between the objects. This is very much like the force law for gravitation discovered by Newton. Then, the gravitation force was,

$$F = G\frac{M_1 M_2}{r^2}\,. \tag{2}$$

Both Eq. 1 and Eq. 2 show the same type of dependence on the product of the charge or mass of the two bodies and the decreasing of the magnitude of the force with the square of distance.

One of the most significance differences between the Coulomb's law and Newton's law of gravity is that gravitation forces are always attractive. Charges come in two varieties, positive and negative, and so charges can attract or repel. Mass is always positive, so the produce of the two masses ($M_1 M_2$) is necessarily positive. Scientists wonder what change in our understanding of Nature would be required if we ever discovered negative mass. Much more than Newton's law of gravity would have to be rethought!

North and South Poles and Magnets

5 Special Relativity

5.1 Time Dilation

5.2 Length Contraction

5.3 Simultaneity and Lorentz Transformations

We have learned that things moving quickly look shorter (Sec. 5.2) and clocks moving near the speed of light will be seen to slow down (Sec. 5.1) considerably. It is important to understand that this is not simply appearances, it is quite real and it is true for all observers.

One of the things that makes this confusing is how can two sets of observers both be seeing each other having slower clocks and shorter measuring sticks? And how can this be happening simultaneously? Well, simultaneity is something that is lost when learning relativity. We find that things we thought occurred at the same time may or may not actually be simultaneous - depending on who is doing the observing.

> An **event** is specified by a specific location and time in a particular reference system.

An event may be observed from more than one reference frame in which case it may have a different location and time in that different reference frame. To obtain the location and time of an event in a new reference frame from the location in time in some known reference frame, we need a mathematical formula known as a **transformation law.**

A transformation law is a rule that gives the location and time of an event in one reference frame in terms of the location and time in another reference frame. These are also called **transformation equations.**

If the coordinates (location and time) of an event is given in an *inertial reference systems*, then a transformation rule that gives it in another *inertial reference system* in a **Lorentz transformation**.

Two events are said to be **simultaneous** if they are have the same time. This is reference system dependent, so we must say that they are *simultaneous in a particular reference frame.*

5.4 Defining Synchronization and Distance

To determine the space time coordinates of two events one needs a distance measure and a time measure. In face, one needs to establish distance measures and time measure throughout the entirety of one's coordinate system. In Fig. 1 we showed a spatial coordinate system and we may suppose that we have taken a meter stick and gone along all the x,y, and z axes and marked off one meter intervals.

Marking the coordinate axes with meter sticks fixes the distance measure, but we can't put a clock at each point where we have marked too. Well, yes we can - at least hypothetically. We will suppose that our coordinate system has a clock at each location that is marked on the coordinate axis. Then, any time we see an event happen, we can mark down its location by reading what the label is on the location and reading what the clock says.

The only problem with the clocks is to find someone to check them regularly to see if they are keeping good time. In fact, they all have to exactly read the **same time** at all location at all times. Instead of paying someone to walk around and check the times, we can hire someone who has taken a physics class and pay them to do it. A person who has taken a physics class knows that one of the laws of physics is that light travels in a straight line at a constant speed. The speed, $c = 186,000$ miles/second, means that if we take a picture of a bunch of clocks at night using a flash bulb on a camera, the clocks further away will read later times because it takes longer for the light from the camera to reach the clocks. In fact, if we take the picture at exactly midnight, which we would call zero time, the time, t, on each clock should be related to the distance, x, to the clock by

$$t = \frac{|x|}{c}.$$

This should work in all x, y and z directions and, if we are a little more sophisticated and define distance by

$$d = \sqrt{x^2 + y^2 + z^2}$$

then $t = d/c$, should work everywhere.

> We define the measurements of time in an inertial coordinate system by defining **simultaneous** to mean that two clocks **synchronized** using light traveling between them will read the same time when two simultaneous events occur at the locations of the clocks.

Well, that took a long time, but know we have a fool proof way of measuring distance and time defining the time interval between events. At least in the inertial coordinate system that we have established with our clocks and meter sticks.

6 Lorentz Transformation Laws

In Sec. 5.4, we decided how to make a fool proof reference frame with meter sticks and clocks so that no matter where something happened and when it happened, we could just look at it and see the position and time of the event. Now, we need to learn how to translate these positions and times between observers in different inertial coordinate systems.

6.1 Defining the Origin

If two inertial observers were travelling along minding their own business, they would have established coordinate systems which were not necessarily close to each other, they might be oriented at weird angles, and their clocks would be set entirely differently. This disorganized state of affairs makes writing the transformation equations difficult and messy. Since we are trying to discover the laws of nature, we inherently believe that nature does no care how we set up coordinate systems, so we might as well set them up simply.

Two different inertial observers can set up reference systems that make the transformation equations simpler if the agree on a few things. First, they must locate the origin of there coordinate systems so that they actually coincide as the two observers pass each other. It will be even simpler if they choose the direction of relative motion and name that the X direction. Finally it will be most convenient if the set their clocks so that the clocks at the origin read zero just at the instant that the two reference frame pass each other.

A figure showing the simultaneous illumination of clocks at the two ends of a rocket ship.

Figure 4: A figure showing the simultaneous illumination of clocks at the two ends of a rocket ship.

6.2 Velocity Addition

6.3 Einstein's Famous Formula