The Revelations of Einstein
Version A: Full Math

W. Blaine Dowler

April 21, 2012

Contents

1 Einstein's Revelation

In the early 1900s, physicists were confused. Electricity and magnetism, as described by Maxwell's equations, seemed incompatible with Newton's work in motion. The incompatibility arose when two different observers studied the same phenomenon from two different frames of reference. On top of that, they had a hard time studying the ether they had proposed as a fundamental part of their theory of light, as they could not produce the interference effects one would expect when the speed of light changed with respect to the medium the light was traveling through.

In 1905, Einstein solved both of these problems with one revelation.[1] Scientists were thrilled with Maxwell's calculation of the speed of light, which was derived solely from the values of a few physical constants. Einstein was the first to recognize that a fundamental piece of information appeared to be missing from the calculation. Maxwell's calculation of the speed of light was completely unrelated to the observer's frame of reference.

In other words, two different observers looking at the same light wave would agree on the speed the light was moving, regardless of how they were moving with respect to each other. They do not necessarily agree on the *velocity*, as that includes direction, but they will agree on the speed.

The manner in which this explains the Michelson-Morley result is relatively clear: if you cannot produce a difference in the speed of light along the two paths, you cannot produce different interference patterns by moving the apparatus. The manner in which this reconciles with Maxwell is far more involved; we will need to lay some groundwork before we can cover that.

2 Postulates

In mathematics, when one is formulating a new theory or idea, one begins with "axioms," or ideas that are assumed to be true. The equivalent statements in the experimental sciences are called "postulates." A postulate is something that seems reasonable and is assumed to be true based on daily experiences or scientific observations. A properly formed physics theory begins with postulates that are thoughts formed in some spoken language, which is then translated into mathematics and the implications are followed.

[1] The third problem mentioned in the previous lesson, dealing with Mercury's orbit, would also be solved by Einstein, but not for another ten years.

2.1 Newton's Postulates

When Newton did his work in motion, he proposed several postulates, including three laws and six corollaries. He also made a few assumptions that he did not formally propose as a part of his system. The laws and his key assumptions are listed here:

1. Law 1: Unless an outside force is applied, a stationary object (object "at rest") will remain stationary and a moving object will continue moving at the same speed and in the same direction.[2]

2. Law 2: If one does apply a force to an object, that object will accelerate in the direction the force is applied. Greater forces create greater changes in speeds. These changes are proportionate to both the force and the "quantity of matter" in an object. Newton originally proposed that the "quantity of matter" was the inertia of an object, and then suggested that the mass of an object is that inertia, so that mass and inertia were identical, which is why inertia is discussed conceptually in class and then mass is used in the algebra.

3. Law 3: If body A pushes against body B with a given force in a given direction, body A also feels a force of equal magnitude and opposite direction coming from body B.

4. Unstated assumption 1: In all of Newton's work, though he didn't formally state it, he assumed that two different individuals measuring the elapsed time between two events would always measure the same interval of time. In other words, if some poor kid's scoop of ice cream falls off its cone at a fair, the time it takes that ice cream scoop to fall to the ground would be the same amount of time whether measured by that child's parent or by someone riding the ferris wheel.

5. Unstated assumption 2: This, technically, was unstated in the *Principia Mathematica* in which he detailed his laws of motion, but it was written in Newton's work in theology. Newton recognized the possibility of nonequivalent reference frames, and was one of the first to propose that a "preferred" reference frame could be found that would indicate the location of the Christian Heaven.

[2] Prior to Newton, it was believed amongst the Christian cultures that moving objects required continuous outside forces to stay in motion. Objects moving along the ground stopped not due to friction but due to the lack of a continuous outside driving force. The stars, planets, moons and other heavenly bodies remained in motion because, it was believed, God had tasked angels with the job of pushing them constantly around in the sky for the amusement of humans. The author is unaware of any suggested reason that God disliked these angels so much that He doomed them so such a monotonous eternity, nor is he aware of alternate theories proposed by the non-Christian cultures.

The corollaries are clarifications of how the math behind the three laws works, and will not be discussed here in detail.

2.2 Einstein's Postulates

Einstein proposed two postulates in 1905 that turned the physics world on its head, eventually. The initial reaction to Einstein's crazy ideas was utter disbelief in response to some of the mathematical implications.

1. All "inertial" reference frames are valid and equivalent. An "inertial" reference frame is one in which the origin of the coordinate system is not being accelerated. Thus, you can measure relative to a stationary object or to an object moving at a constant speed, but not relative to someone riding a ferris wheel, since that person's direction is changing. (Direction counts for acceleration.) In 1915, Einstein would formulate a new theory that would allow the use of any reference frame, treating them all as equivalent, but that addition will come in a later lesson.
2. The speed of light is "invariant," meaning all observers in all valid reference frames measure exactly the same *value* for the speed light is moving at.

Einstein felt he was on the right track when he used these postulates to derive new ways to adjust the math for changing reference frames, and arrived at a result *identical* to the one that H. A. Lorentz found when he tried to reconcile Maxwell's equations with Newton's laws. That seemed to be a very compelling coincidence.

The complete implications of these postulates are not obvious, particularly if one is trying to see the connection between these ideas and the famous $E = mc^2$ equation associated with relativity. We will examine some of the implications of these postulates from now until the end of lesson six. In lesson seven, we will add the postulates that allow reference frames to experience acceleration, and deal with those implications for the rest of this series of lessons. These early lessons deal with the special theory of relativity, since they relate to the "special case" in which gravity and other outside fields and forces do not apply. The latter lessons deal with the general theory of relativity, in which gravity and accelerations are included, so that we can work with "general" cases of physical situations.

3 Implications for Distances

Let us begin by examining the way distances are measured by moving observers under Newtonian mechanics and reference frames.

Imagine we have two observers, named Hal and Barry. Hal is standing still, while Barry is running at a constant speed. Both are observing a car driving in the same direction Barry is running and at a higher speed than Barry is running with. When Hal measures the distance the car has traveled, he measures relative to himself. Barry also measures the distance the car has traveled relative to himself, but he is *between* Hal and the car, so he concludes that the distance traveled by the car is *less* than the distance Hal measures.

"But wait!" one might cry. "Why not measure relative to the road to eliminate this problem?" This is what is commonly done in labs for consistency. However, the important thing to recognize is that there *is no* problem here. If all inertial frames are equivalent, then it doesn't matter *who* does the measurement, provided the frame of reference is described accurately in the results. This helps in cases such as the motion of astronomical bodies, in which case there is no clear "road" to use as a point of reference.

On a conceptual level, there is no difference between the Newtonian view and the Einsteinian view. The observer moving in the same direction as the car will measure a smaller distance traveled than the observer who is stationary relative to the road. Once the math is examined in detail, one finds that the difference between the two measurements is more dramatic in the Einsteinian view, but the basic concept is the same. The difference is due to the invariance of the speed of light; there are no "instantaneous" measurements, since information travels no faster than the speed of light.[3] In short, Hal and Barry don't measure the position of the car "now," but instead measure the position of the car according to where it was when it reflected the light their eyes are detecting "now." This lag causes the difference between the Newtonian and Einsteinian measurements observed here.

This has further implications. A car has a volume; what if they measure the distances to both the front *and* rear of the car? The time lag due to the transit time of light is more pronounced from the front (further) side of the vehicle than it is from the rear (closer) side. Thus, if you ask both Hal and Barry to measure the *length* of the car, they would provide two different answers! In relativity, objects moving relative to the observer contract or shrink, and do so in greater degree as the differences in speed become more pronounced. This is known as *length contraction.*

[3] These lessons do not deal with some quantum mechanical phenomena such as entanglement or recent neutrino results which may indicate exceptions to this rule. Look for separate lessons on those topics when they are better understood.

4 Implications for Time Intervals

Under Newtonian mechanics, Hal and Barry would disagree about the distance the car traveled in our previous example, but they *would* agree on the amount of time the car spent in motion. The invariance of the speed of light throws a monkey wrench into that perspective.

Imagine the same situation as above. This time, however, the car's headlights are on. Instead of measuring the distance and time elapsed of the car's travel, Hal and Barry measure those quantities for the light waves emitted by the car's headlights.

Just as before, they measure the distance traveled by that light, and come up with different results. In Newtonian mechanics, they would each measure the speed of that light by measuring the elapsed time, and then dividing the distance traveled by the elapsed time. As they would agree on that time interval, they would arrive at two different numbers for the speed of light. In Einstein's relativity, this is impossible; all observers measure light in a vacuum to be traveling at the same, invariant speed. What implications does that have?

We can flip the math around. Both observers measure the distance traveled, though they disagree on the exact value of this distance. They also measure the speed of light, and both arrive at the same value of $c = 299\ 792\ 458\ \frac{\mathrm{m}}{\mathrm{s}}$. The time elapsed can be calculated by taking the distance traveled and dividing by the speed of light. Hal and Barry each do this calculation, using *different* numbers for the distance traveled, but the *same* numbers for the speed of light: they arrive at *different* numbers for the time elapsed, each of which is correct in that observer's reference frame!

This was a staggering thought. Time depends upon the speed of the observer. This phenomenon, known as *time dilation*, has since been verified by many, many experiments. This is a real and undeniable effect. If someone observes a moving clock, he or she would conclude that the clock in motion is running more slowly than an accurate clock which is considered to be "at rest."

5 Implications for Velocities

Let us return to Hal, Barry and the car. This time, instead of running in the same direction as the car, Barry runs in the opposite direction. Let us also assume that, as Hal measures things, both Barry and the car are moving at more than half the speed of light; for the sake of argument, let us choose three quarters of the speed of light as the speed they travel.

In the Newtonian view, if Barry were to measure the speed of the car, he would calculate the speed that is simply the sum of his speed and the car's speed, as measured by Hal. In this case, that would result in an answer that is one and a half times the speed of light. In the Einsteinian view, this changes dramatically.

For the sake of argument, let us assume the car's headlights are still on. When Barry measures the speed of the light coming from those headlights, he concludes that it travels at the speed of light $c = 299\ 792\ 458\ \frac{\text{m}}{\text{s}}$, as he must. When he measures the speed of the car itself, he finds that it is *slower* than the speed of light; the light wave from the headlights are further away from the car, and getting further away all the time.

In other words, the Einsteinian view allows for agreement about *relative* speeds of bodies, but not *absolute* speeds. Simple comparisons such as "faster" and "slower" are maintained, but the actual numbers vary considerably. This is because the observers don't just disagree on the distances traveled, but also on the time elapsed. The truly stunning implication is this: *nothing* can travel faster than the speed of light. Any moving object can "turn on its headlights" or not, and it will not affect how others measure the speed of that body. Everybody will, however, agree that the hypothetical light from those headlights travels faster than the source (the moving object) and that the light is traveling at the speed of light. Therefore, all observers agree that the source of the light is moving slower than light.

6 Einstein and Maxwell

Now that these simple implications have been discussed, we can return to our paradox from electricity and magnetism. Recall the situation with two infinitely long and identically electrically charged rods. From a reference frame in which they are stationary, the rods will repel each other due to the like charges on the rods. The more dense the electrical charges are on those rods (i.e. the more charge is present in any given length of the rod) the stronger this repulsion becomes. From a moving reference frame, these charges form a current, which results in a magnetic attraction. If one moves quickly enough, the laws of nature described by Newton and Maxwell would predict that they would eventually attract each other more strongly than they would repel, creating a paradox: the rods cannot repel from Hal's perspective but collide from Barry's perspective.

The Einsteinian model resolves this paradox in a somewhat surprising way. The charge density on the two rods plays just as strong a role in generating the electrical force as it does the generating the magnetic force. As a result, every situation changes from a repulsive force to an attractive force at the exact moment the moving observer reaches the speed of light, which is not possible. (If Barry has a flashlight in hand, he will never outrun the light that comes out of it, so he must always be traveling at less than the speed of light.) True, Barry will observe a stronger force than Hal observes, but he also measures less elapsed time than Hal, so the overall effect results in the same predicted motion of the two rods. This completes the conceptual content for this lesson. Now, we will see the math that backs this up.

7 The Lorentz Transformations

The method of derivation used by Albert Einstein[4] will be followed here. That used by Lorentz is significantly longer and will be omitted.

We begin by examining a photon in two different reference frames, which we label S and S'. Each frame contains an observer who considers himself or herself to be stationary, or at rest, with respect to the other frame. The origins of the two frames match up perfectly, so that when $x = y = z = ct = 0$ then $x' = y' = z' = ct' = 0$, and the axes align. (i.e. at time $t = t' = 0$, x is parallel to x' through the common origin, y is parallel to y' through the common origin, and z is parallel to z' through the common origin.) From the perspective of an observer in frame S, the frame S' moves in the positive x direction with speed v.

Assume the photon being observed passes through both origins and is also traveling in the positive x (and positive x') direction. By Einstein's postulate, it travels with speed c in both reference frames, so it satisfies $c = \frac{x}{t} = \frac{x'}{t'}$ automatically. It will be convenient to rearrange these formulae as $x - ct = 0$ and $x' - ct' = 0$ instead. By logic, we assume that there is a function λ which depends solely on v such that the relationship

$$(x - ct) = \lambda (x' - ct')$$

is satisfied. From the S' frame, it can be argued that frame S is moving in the negative x' (x) direction with speed $-v$, so we also have the similar relation

$$(x' + ct') = \mu (x + ct)$$

Adding and subtracting the two transforms them into the relations

$$x' = \gamma x - \delta ct \tag{1}$$

[4]See appendix 1 of *Relativity: The Special and the General Theory*, Albert Einstein, 1905-1961, available from a variety of publishers.

and

$$ct' = \gamma ct - \delta x \tag{2}$$

where

$$\gamma = \frac{\lambda + \mu}{2}$$

and

$$\delta = \frac{\lambda - \mu}{2}$$

If we observe the origin of S' now, instead of the photon, we have $x' = 0$ at all times. We can rearrange equation 1 on the preceding page to obtain

$$\frac{x}{t} = v = \frac{\delta c}{\gamma}$$

Now, let us combine equations 1 on the previous page and 2 by first isolating t in equation 1 on the previous page:

$$t = \frac{\gamma x - x'}{\delta c}$$

and substituting it into equation 2:

$$ct' = \frac{x}{\delta}\left(\gamma^2 - \delta^2\right) - \frac{\gamma}{\delta}x'$$

This general expression should hold for any object, including one traveling along the x (and x') axis in such a way that one end passes through the origin of both coordinate systems while the other end is on the positive x' (and x) axis. If we measure the length of this object at this moment, then $t' = 0$ and we have

$$x' = \gamma x \left(1 - \frac{v^2}{c^2}\right)$$

We can apply the same logic in either reference frame, and arrive at the overall formula

$$\gamma^2 = \frac{1}{1 - \frac{v^2}{c^2}}$$

or (selecting the positive root)

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{3}$$

Complete substitution results in the Lorentz transformations:

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{4}$$

$$y' = y \tag{5}$$

$$z' = z \tag{6}$$

$$t' = \frac{t - \frac{v}{c^2}x}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{7}$$

It is common to use equation 3 on the previous page to express these as

$$\begin{aligned} x' &= \gamma\,(x - vt) \\ y' &= y \\ z' &= z \\ t' &= \gamma\left(t - \frac{v}{c^2}x\right) \end{aligned}$$

and we will do so for most of these lessons.

We can also examine frame S from the frame S' and find the similar equations

$$x = \gamma\,(x' + vt') \tag{8}$$

$$y = y' \tag{9}$$

$$z = z' \tag{10}$$

$$t = \gamma\left(t' + \frac{v}{c^2}x'\right) \tag{11}$$

8 Length Contraction and Time Dilation

Once the Lorentz transformations have been established, it becomes a straightforward matter to quantify the length contraction. We begin with equation 4:

$$x' = \gamma\,(x - vt)$$

Now, if we wish to determine the length of an object moving in frame S' as measured by an observer in frame S, we measure the positions of the two ends of the object and subtract them so that $l = x_2 - x_1$. We have the freedom to choose our origins of the coordinate systems as before, with the "trailing" edge of the moving object passing through the origins of both coordinate systems. With that definition, $x_1 = 0$ and we can drop the subscripts. We also measure both ends simultaneously[5] as the trailing edge passes through the origin, so that $t = 0$. This planned definition of variables makes the calculation relatively simple:

$$x' = \gamma x$$

Remembering that the object is at rest in frame S', we find that

$$l = l'\sqrt{1 - \frac{v^2}{c^2}}$$

When $v < c$, we find that $l < l'$ (as $\sqrt{1 - \frac{v^2}{c^2}} < 1$), and the length is contracted.

Time dilation is calculated by virtually identical math to length contraction. When an observer in S watches a stopwatch positioned at the origin in frame S', he or she finds that

$$t = t'\sqrt{1 - \frac{v^2}{c^2}}$$

9 Velocity Transformations

The velocity transformation is fairly easy from a conceptual standpoint. From a notational standpoint, we already use the symbol v as the speed of the frame S' relative to frame S, so we need a different symbol. Therefore. let the velocity we are concerned with be denoted as

$$\mathbf{u} = \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix}$$

or

$$\mathbf{u} = \frac{d\mathbf{x}}{dt}$$

Thus, to calculate the velocity as measured in frame S', we compute

$$\mathbf{u}' = \frac{d\mathbf{x}'}{dt'}$$

In the case when the reference frames move with speed v along the x/x' axes only, then we have $u'_y = u_y$ and $u'_z = u_z$. To calculate u'_x, we use the total differentials. We treat v as a constant, so we have

$$dx' = \gamma\left(dx - vdt\right)$$

[5]More on this word later.

and

$$dt' = \gamma\left(dt - \frac{v}{c^2}dx\right)$$

Taking the ratio, we have

$$\begin{aligned}
u'_x &= \frac{dx'}{dt'} \\
&= \frac{\cancel{\gamma}\left(dx - vdt\right)}{\cancel{\gamma}\left(dt - \frac{v}{c^2}dx\right)} \\
&= \frac{\frac{dx}{dt} - v}{1 - \frac{vdx}{c^2dt}} \\
u'_x &= \frac{u_x - v}{1 - \frac{u_xv}{c^2}}
\end{aligned}$$

which is the desired result.