Comments on Theories of Relativity

John Smigel, 31 Jan 2018, Updated 25 October 2018

1. Introduction

This paper shows the steps in a derivation of Einstein's Special Theory of Relativity. This has been done many times, but I will try to add enough detail so that I can follow it. The principles that it is based on are very simple:

1) The laws of physics do not change between all inertial (uniform velocity, not accelerating) frames of reference (coordinate systems).
2) The speed of light, *c*, is constant in a uniform medium through which it is propagating (does not depend on motion of the light source).

It would be very disturbing if 1) were not true, considering we are on a rotating planet that is also orbiting around the sun; we are flying through space. Reactions to 2) vary from 'it is obvious and already covered by 1)' to 'it is clearly false or not proven.' The speed of light is different from the speed of most things we observe, partly because it has no mass and partly just because it is different. We are taught that light can travel more slowly through some materials relative to free space, which is how a lens works. Electromagnetic theory states that electromagnetic energy, including light, propagates at a constant speed through a material and the speed depends on the material properties of

$\mu_0 = 4\pi \times 10^{-7}\ \mathrm{N{\cdot}A^{-2}}$ (or H m^{-1})= magnetic permeability of free space (constant; henries per meter or newtons per ampere squared)

ε_0 = vacuum permittivity (or permittivity of free space or electric constant) = $8.854187817\times10^{-12}\ \mathrm{F{\cdot}m^{-1}}$ (farads per meter)

$$c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}} = 2.99792458 \times 10^{8}\ \mathrm{m/s} \qquad (1\text{-}1)$$

I probably should call this c_0, but this will make the notation messy. The speed of light through various materials is often characterized by the index of refraction defined as light speed in a vacuum divided by speed in the material.

Material	**Index of Refraction**	
Vacuum	1.0000	<--lowest optical density
Air	1.0003	
Ice	1.31	
Water	1.333	
Ethyl Alcohol	1.36	
Plexiglas	1.51	

Crown Glass	1.52	
Light Flint Glass	1.58	
Dense Flint Glass	1.66	
Zircon	1.923	
Diamond	2.417	
Rutile	2.907	
Gallium phosphide	3.50	<--highest optical density

2. Special Relativity Derivation

The special theory of relativity has to do with how objects defined in a moving coordinate system relate to the same objects defined in a fixed coordinate system. We generally define something's location in space by three spatial coordinates with axes at right angles. Here I define the fixed spatial coordinate axes as x, y, and z. There is also some time, t, which goes along with this coordinate system. Space and time are linked together by speed whether we like it or not since speed is a spatial distance moved divided by a time and speed is fundamentally important here. For convenience, and since we know the answer already, we form a 4-dimensional coordinate system called space-time that includes the 3 spatial coordinates and a time-related coordinate. The time coordinate is chosen to be scaled by the constant, c, the speed of light to make all the coordinate units the same. The time coordinate is then converted to the distance light would move in that amount of time.

The first (fixed) coordinate system space-time vector is then given by

$$C_1 = \begin{bmatrix} x \\ y \\ z \\ ct \end{bmatrix} \text{ = fixed or reference coordinate system} \qquad (2\text{-}1)$$

We define a second coordinate system that is moving with a constant speed, v , relative to the first coordinate system. For simplicity, this speed is defined as being in the x direction, (velocity vector $\mathbf{v} = v\hat{x}$ where $\hat{x}$ is a unit vector in the x direction) but it could be in any direction, with the coordinates rotated appropriately. Something not moving in the second coordinate system will be moving with speed, v, relative to the first coordinates.

A space-time location in the moving coordinate system is represented by

$$C_2 = \begin{bmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \\ c\tau \end{bmatrix} \text{ = coordinate vector for system moving with speed, } v. \qquad (2\text{-}2)$$

The question is: How do the coordinates of something measured in C_2 relate to the coordinates measured in C_1? In other words, what is the best mapping or transformation between C_1 and C_2? We would like to know if we measure something in C_2, what is the corresponding measurement in C_1 and vice versa. By 'best' transformation, we mean that we want the measurements to agree with what we would actually measure in the two different coordinate systems. Einstein takes great care to define precisely how distance and time measurements would be made and what it means to have clocks synchronous or not. Maybe by the time I finish this document I will be able to explain it and the associated importance. For now, just note the second coordinates have a different time variable, τ , measured local to the second coordinate system.

One thing is clear to me if (since) the speed of light is constant: there is a potential problem measuring time and/or distance using light if you are on a potentially moving frame of reference (anywhere on Earth, for example).

The first question in solving this problem is: What exactly does it mean that the laws of physics do not change between the different reference frames? I don't know about you, but if I was doing this analysis and I came up with the answer that time was not constant, but depended on speed, I would assume that this was a change to the laws of physics. Perhaps the principle should be more accurately stated that "the current laws of physics do not change, unless they are wrong."

If we want the laws of physics to be the same, a translation of the coordinates (or moving linearly), must not change the equations that define the laws, translational symmetry. In other words, the laws do not depend on the absolute location, but only on relative distances between points. For this to be true, the transformation between coordinates must be linear; able to be expressed in the form of a matrix times the coordinate vector plus a constant

$$\begin{bmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \\ c\tau \end{bmatrix} = \mathbf{\Lambda}(v\hat{x}) \begin{bmatrix} x \\ y \\ z \\ ct \end{bmatrix} + \mathbf{k} \ , \qquad (2\text{-}3)$$

where

$$\mathbf{\Lambda}(v\hat{x}) = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} & \Lambda_{13} & \Lambda_{14} \\ \Lambda_{21} & \Lambda_{22} & \Lambda_{23} & \Lambda_{24} \\ \Lambda_{31} & \Lambda_{32} & \Lambda_{33} & \Lambda_{34} \\ \Lambda_{41} & \Lambda_{42} & \Lambda_{43} & \Lambda_{44} \end{bmatrix} = \text{coordinate transformation matrix} \qquad (2\text{-}4)$$

and $\mathbf{k}$ is a constant vector.

I try to use bold symbols for a matrix or vector, with lower case for a vector and upper case for a matrix. It is conventional and convenient to use the notation

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ z \\ ct \end{bmatrix},\ \tilde{\mathbf{x}} = \begin{bmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \\ c\tau \end{bmatrix}$$

, where bold **x** represents a space-time coordinate vector (don't confuse with x coordinate only). (2-5)

The coordinate transformation can be written as

$$\tilde{\mathbf{x}} = \mathbf{\Lambda}\left(v\hat{x}\right)\mathbf{x} + \mathbf{k}\,. \tag{2-6}$$

We can define the coordinates (including time) so that the constant, $\mathbf{k}$, can be taken as zero for simplicity. It remains to find physical constraints to determine the 16 terms in the 4x4 transformation matrix.

We now have

$$\tilde{\mathbf{x}} = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} & \Lambda_{13} & \Lambda_{14} \\ \Lambda_{21} & \Lambda_{22} & \Lambda_{23} & \Lambda_{24} \\ \Lambda_{31} & \Lambda_{32} & \Lambda_{33} & \Lambda_{34} \\ \Lambda_{41} & \Lambda_{42} & \Lambda_{43} & \Lambda_{44} \end{bmatrix}\mathbf{x}\,. \tag{2-7}$$

Many Λ's can be determined by the physical constraints associated with how the problem is defined. First, motion is only in the x direction so any coordinate component in the y or z dimensions will not depend on x or ct. Also, z will not depend on y and y will not depend on z since there is only x translation and the coordinate systems are orthogonal. The second and third rows of $\mathbf{\Lambda}$ determine the transformed values of $\tilde{y}$ and $\tilde{z}$. All terms in these rows will be zero except the diagonal terms, Λ_{22} and Λ_{33} will be equal to 1. So far this gives

$$\tilde{\mathbf{x}} = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} & \Lambda_{13} & \Lambda_{14} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \Lambda_{41} & \Lambda_{42} & \Lambda_{43} & \Lambda_{44} \end{bmatrix}\mathbf{x}\,. \tag{2-8}$$

Also, the new x component can't depend on y or z because that would not preserve a translation as just a translation or a rotation as just a rotation. This means that Λ_{12} and Λ_{13} are zero,

$$\tilde{\mathbf{x}} = \begin{bmatrix} \Lambda_{11} & 0 & 0 & \Lambda_{14} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \Lambda_{41} & \Lambda_{42} & \Lambda_{43} & \Lambda_{44} \end{bmatrix} \mathbf{x}\,. \tag{2-9}$$

By definition of the relative motion of the two coordinates, something located at $\tilde{x} = 0$ in the second coordinate will remain at zero if it is moving at speed, $x = vt$ in the first coordinate system. The corresponding equations are then

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ c\tau \end{bmatrix} = \begin{bmatrix} \Lambda_{11} & 0 & 0 & \Lambda_{14} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \Lambda_{41} & \Lambda_{42} & \Lambda_{43} & \Lambda_{44} \end{bmatrix} \begin{bmatrix} vt \\ 0 \\ 0 \\ ct \end{bmatrix} = \begin{bmatrix} \Lambda_{11} vt + \Lambda_{14} ct \\ 0 \\ 0 \\ \Lambda_{41} vt + \Lambda_{44} ct \end{bmatrix}. \tag{2-10}$$

The top equation gives

$$\Lambda_{41} = -\beta \Lambda_{11}\ , \tag{2-11}$$

where

$$\beta \equiv \frac{v}{c} \text{ and} \tag{2-12}$$

$$\tilde{\mathbf{x}} = \begin{bmatrix} \Lambda_{11} & 0 & 0 & -\beta\Lambda_{11} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \Lambda_{41} & \Lambda_{42} & \Lambda_{43} & \Lambda_{44} \end{bmatrix} \mathbf{x}\,. \tag{2-13}$$

If I did not already know the answer, I would not have said that time in the second coordinate system depends on x. So, I am going out on a limb saying that $c\tau$ should not depend on y or z; since there is no motion in the y or z directions. I'll make the claim, $\Lambda_{42} = \Lambda_{43} = 0$, or in words that the transformed time component does not depend on y or z,

$$\tilde{\mathbf{x}} = \begin{bmatrix} \Lambda_{11} & 0 & 0 & -\beta\Lambda_{11} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \Lambda_{41} & 0 & 0 & \Lambda_{44} \end{bmatrix} \mathbf{x}\,. \tag{2-14}$$

The remaining three variables are determined from the principle that the speed of light is the same in both coordinate systems.

For light propagating along the x axis in the two coordinate systems, the location is given by

$$x = ct \tag{2-15}$$

and

$\tilde{x} = c\tau$, and then (2-16)

$$\begin{bmatrix} c\tau \\ 0 \\ 0 \\ c\tau \end{bmatrix} = \begin{bmatrix} \Lambda_{11} & 0 & 0 & -\beta\Lambda_{11} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \Lambda_{41} & 0 & 0 & \Lambda_{44} \end{bmatrix} \begin{bmatrix} ct \\ 0 \\ 0 \\ ct \end{bmatrix} = \begin{bmatrix} \Lambda_{11}(1-\beta)ct \\ 0 \\ 0 \\ (\Lambda_{41}+\Lambda_{44})ct \end{bmatrix} . \tag{2-17}$$

This doesn't get me there (more unknowns than equations), so I will try equations for propagation at speed c in a general 3-D direction,

$$x^2 + y^2 + z^2 = (ct)^2 \tag{2-18}$$

and

$$\tilde{x}^2 + \tilde{y}^2 + \tilde{z}^2 = (c\tau)^2 \ . \tag{2-19}$$

From the transformation so far, we have

$$\tilde{x} = \Lambda_{11}x - \beta\Lambda_{11}ct \tag{2-20}$$

$$\tilde{y} = y \tag{2-21}$$

$\tilde{z} = z$ and (2-22)

$$c\tau = \Lambda_{41}x + \Lambda_{44}ct \ . \tag{2-23}$$

Inserting (2-20) to (2-23) in (2-19) yields

$$(\Lambda_{11}x - \beta\Lambda_{11}ct)^2 + y^2 + z^2 = (\Lambda_{41}x + \Lambda_{44}ct)^2 \ . \tag{2-24}$$

Expanding the squared terms in (2-24) gives

$$\Lambda_{11}^2x^2 - 2\beta\Lambda_{11}^2xct + \beta^2\Lambda_{11}^2(ct)^2 + y^2 + z^2 = \Lambda_{41}^2x^2 + 2\Lambda_{41}\Lambda_{44}xct + \Lambda_{44}^2(ct)^2 \ . \tag{2-25}$$

Collecting terms multiplying the same coordinate variable

$$(\Lambda_{11}^2 - \Lambda_{41}^2)x^2 + y^2 + z^2 - (2\beta\Lambda_{11}^2 + 2\Lambda_{41}\Lambda_{44})xct = (\Lambda_{44}^2 - \beta^2\Lambda_{11}^2)(ct)^2 \ . \tag{2-26}$$

For (2-18) and (2-19) to both represent the same light wave propagating at speed *c*, the coefficients of x^2, y^2, z^2, *xct, and* $(ct)^2$ in (2-26) and (2-18) must match giving

$$\Lambda_{11}^2 - \Lambda_{41}^2 = 1 \tag{2-27}$$

$$2\beta\Lambda_{11}^2 + 2\Lambda_{41}\Lambda_{44} = 0 \text{ and} \quad (2\text{-}28)$$

$$\Lambda_{44}^2 - \beta^2\Lambda_{11}^2 = 1 . \quad (2\text{-}29)$$

There are three equations in 3 unknowns so we should be good to go. One way is to solve both (2-27) and (2-29) in terms of Λ_{11}^2, insert the results into (2-28) and then solve for Λ_{11}. Solving for Λ_{41} in terms of Λ_{11} from (2-27) gives

$$\Lambda_{41}^2 = \Lambda_{11}^2 - 1 \ . \quad (2\text{-}30)$$

Dividing (2-28) by 2 and then subtracting $\Lambda_{41}\Lambda_{44}$ from both sides gives

$$\beta\Lambda_{11}^2 = -\Lambda_{41}\Lambda_{44} . \quad (2\text{-}31)$$

Squaring both sides of (2-31) gives

$$\beta^2\Lambda_{11}^4 = \Lambda_{41}^2\Lambda_{44}^2 . \quad (2\text{-}32)$$

Solving for Λ_{44}^2 in terms of Λ_{11}^2 in (2-29)

$$\Lambda_{44}^2 = 1 + \beta^2\Lambda_{11}^2 . \quad (2\text{-}33)$$

Now inserting (2-30) and (2-33) in (2-32) yields

$$\beta^2\Lambda_{11}^4 = \left(\Lambda_{11}^2 - 1\right)\left(1 + \beta^2\Lambda_{11}^2\right) = \beta^2\Lambda_{11}^4 + \left(1-\beta^2\right)\Lambda_{11}^2 - 1. \quad (2\text{-}34)$$

Simplifying (2-34) gives

$$\left(1-\beta^2\right)\Lambda_{11}^2 = 1 . \quad (2\text{-}35)$$

Solving for Λ_{11} in (2-35),

$$\Lambda_{11} = \pm\frac{1}{\sqrt{1-\beta^2}} \ . \quad (2\text{-}36)$$

We use the positive solution and define

$$\gamma \equiv \frac{1}{\sqrt{1-\beta^2}} . \quad (2\text{-}37)$$

Inserting (2-36) and (2-37) in (2-30) gives

$$\Lambda_{41}^2 = \frac{1}{1-\beta^2} - 1 = \frac{1}{1-\beta^2} - \frac{1-\beta^2}{1-\beta^2} = \frac{\beta^2}{1-\beta^2} = \beta^2\gamma^2 . \quad (2\text{-}38)$$

Taking the square root of both sides results in

$$\Lambda_{41} = \pm\beta\gamma \ . \tag{2-39}$$

Inserting (2-36) and (2-37) in (2-33) gives

$$\Lambda_{44}^2 = 1 + \frac{\beta^2}{1-\beta^2} = \frac{1-\beta^2}{1-\beta^2} + \frac{\beta^2}{1-\beta^2} = \frac{1}{1-\beta^2} \ . \tag{2-40}$$

Taking the square root of both sides

$$\Lambda_{44} = \pm\frac{1}{\sqrt{1-\beta^2}} = \pm\gamma \ . \tag{2-41}$$

We choose the solutions with the signs such that

$$\Lambda_{11} = \Lambda_{44} = \gamma \tag{2-42}$$

and

$$\Lambda_{14} = \Lambda_{41} = -\gamma\beta \ . \tag{2-43}$$

The resulting transformation is called the Lorentz transformation given by

$$\begin{bmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \\ c\tau \end{bmatrix} = \begin{bmatrix} \gamma & 0 & 0 & -\gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma\beta & 0 & 0 & \gamma \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ ct \end{bmatrix} = \begin{bmatrix} \gamma\left(x-\beta ct\right) \\ y \\ z \\ \gamma\left(ct-\beta x\right) \end{bmatrix} . \tag{2-44}$$

This set of equations is the conclusion in Einstein's Special Theory of Relativity. So, what does it mean? The most shocking things are that, since γ is greater than 1 for non-zero speeds, time and distance in a moving coordinate system will be scaled relative to the fixed coordinate system. Time passes more slowly and lengths are shorter in a moving system.

2.1 Discussion

While I am still fascinated by quantum physics and relativity, I am disappointed in my inability to fully understand what is going on. I was hoping that if I studied and thought about these things more, I would understand them better. Unfortunately, the opposite seems true: the more I learn and study, the more disturbed and confused I get. Part may be due to being so old. At first, I naively thought that the speed of light being constant in all inertial reference frames made sense, but the more I think about it and the ramifications, the more disturbing it is. In addition to time scaling, Einstein's principle of relativity of simultaneity that follows from special relativity is disturbing. One classic example is light pulses on a moving train. If two light pulses are sent from the center of a moving train in opposite directions at the same time, an observer on the train in the moving frame of reference sees them hit the train front and back walls at the same time. The observer on the train can't see that he or she is moving and the light is constrained to move at c in this local reference frame; therefore, hitting the front and back walls simultaneously because the distance is the same. However, a stationary observer off, but near the same train sees the train is moving and can see that the back of the train moves closer to where the light pulse originates as it propagates (and the front moves farther away), so the light is observed to hit the back of the train before the front. From this point of view, the same light pulses do not hit the walls at the same time, not simultaneous. In both reference frames the light is observed to travel at the same speed of light, c. In the moving train case, the light actually travels a little farther in the fixed frame than it appears in the moving frame because the frame is moving. That is what causes the time to have to be a little different for the moving reference frame in order to keep the light speed in that frame constant at the value, c.

This kind of bizarre behavior seems standard for quantum physics. How can the light behave differently just because it is being observed differently or measured in a different reference frame? In other similar quantum physics puzzles, the currently-accepted answer is generally along the lines of: multiple realities are all happening simultaneously and are only known with certain probabilities until a specific observation is made. The observation then changes the physics of what happens or happened. This seems like physicists punting to me. We are defining a solution that is forced to match the observations by definition, but giving little or no insight into what is going on. Still fascinating, but frustrating to understand.

In the next section, I'll show the derivation of $E=mc^2$ that follows directly from the section 2 special relativity result. If I really need more punishment, then I'll move on to general relativity (eventually).

3.0 Equivalence of Mass and Energy

This section shows a derivation of Einstein's famous equation, $E = mc^2$, describing the equivalence between mass and energy. The derivation uses Einstein's special theory of relativity. The question to be investigated is "if a body of mass, *m*, emits energy amount, ΔE, then how much does the mass of the body change?" Einstein set up the problem in two coordinate systems, one in which the body is at rest (moving with the body) and a second moving at a small linear velocity, *v*, with respect to the body. The body simultaneously emits two photons in opposite directions, with total energy of ΔE. Having the two photons in opposite directions simplifies the problem because the velocity of the body does not change due to the emission of the photons in opposite directions.

Some things we know from Special Relativity are that the speed of light (photons) is constant, *c*, and that the laws of physics have to be the same in both the coordinate systems that differ only by a linear velocity, *v*. The relevant law of physics for this analysis is conservation of total energy. This means the total energy must be the same before and after emission of the photons in both coordinate systems.

Before writing the equations defining the photon energy emission, it is a good idea to perform a quick dimensional analysis. Energy is the potential to do work, which can be expressed as a force operating over a distance

$$energy = (force)(distance). \tag{3-1}$$

So energy units are force units times a distance unit. We want a relationship between energy and mass, so the force can be written in terms of mass times acceleration,

$$force = (mass)(acceleration) \tag{3-2}$$

or inserting (3-2) in (3-1)

$$energy = (mass)(acceleration)(distance)\ . \tag{3-3}$$

Acceleration is a rate of change of velocity, and velocity is a rate of change in distance. So acceleration has units of distance divided by time squared. The resulting energy units must be given by

$$energy = mass\left(\frac{\text{distance}}{\text{time}^2}\right)\text{distance}=\left(\frac{\text{distance}}{\text{time}}\right)^2 = mass(velocity)^2 \tag{3-4}$$

Just by dimensional analysis, we know that energy should be related to mass times a velocity constant squared (to within some unitless constant scale factor not shown here).

The velocity constant could also be written as some acceleration constant times a distance constant (equivalent to some velocity constant squared).

Now back to the analysis if a body emitting photons. The body is at rest in the body-fixed reference frame, F, with mass, *m*, and total energy, E_b. After the body emits the two photons, the total energy will be split between the photons and the body (mass),

$$E_b = \left(E_b - \Delta E\right) + E_{p1} + E_{p2} \tag{3-5}$$

or

$$\Delta E = E_{p1} + E_{p2} \tag{3-6}$$

where

E_{p1} = energy of photon 1 in body-fixed coordinates = $\frac{\Delta E}{2}$

E_{p2} = energy of photon 2 in body-fixed coordinates = $\frac{\Delta E}{2}$

ΔE = total energy emitted by the body.

The energy of a photon is given by Planck's law,

$$E_p = h\nu \text{ ,} \tag{3-7}$$

where

$$\nu = \text{temporal frequency} = \frac{c}{2\pi}k \text{ , (do not confuse this with velocity, } v\text{)} \tag{3-8}$$

k = wave number (or radian spatial frequency),

and *h* = Planck's constant. Note that the temporal frequency units are cycles per second and the wave number units are radians per meter. For a sinusoidal wave propagating at speed, *c*, the wave amplitude is

$$A\sin(k(x-ct)+\theta) \text{ .} \tag{3-9}$$

For a given time, the wave number is the number of cycles across space for a unit distance.

The body's mass changes by an amount, Δm, after emitting the photons.

The total energy expressions for this photon emission event can also be given in the moving coordinate system, F_v.

Both the energy of the mass and the photons are slightly different in the moving coordinate system than the body-fixed coordinates. From Newtonian mechanics, a body of mass, *m*, moving at velocity, *v*, obtains kinetic energy of

$$E_k = \frac{1}{2}mv^2 \text{ . (See note 2 below for details)} \tag{3-8}$$

To get the correct expressions for the energy of the photons in the moving coordinate system, we must use the special relativity transformation given by (2-44) above, repeated here

$$\begin{bmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \\ c\tau \end{bmatrix} = \begin{bmatrix} \gamma & 0 & 0 & -\gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma\beta & 0 & 0 & \gamma \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ ct \end{bmatrix} = \begin{bmatrix} \gamma\left(x-\beta ct\right) \\ y \\ z \\ \gamma\left(ct-\beta x\right) \end{bmatrix} . \tag{2-44}$$

The photon energy depends on the temporal frequency or wave number. The frequency is shifted by a Doppler shift in the moving coordinate system, F_v. One of the photons is taken to be in the same direction as *v* and the other is in the opposite direction. Therefore, one photon is Doppler compressed and the other photon is Doppler expanded. It is counter intuitive to me that the photons would have different energy when expressed relative to a moving coordinate system compared with a fixed coordinate system. But it is clear why one would measure a different frequency when moving, compared with not moving.

We desire the transformation of the spatial frequency from the fixed coordinate frame to the moving frame, with propagating wave amplitude,

$$A\sin(\tilde{k}(\tilde{x}-c\tau)+\theta)\,. \tag{3-9}$$

I am using the tilde overbar ($\tilde{k}$ for example) to distinguish moving coordinate variables from fixed coordinates. The temporal frequency has units of reciprocal time and the spatial frequency has units of radians per unit distance. From the time transformation (4th equation) in (2-44),

$$c\tau=\gamma\left(ct-\beta x\right), \tag{3-10}$$

where

$$\beta=\frac{v}{c} \tag{3-11}$$

$$\gamma=\frac{1}{\sqrt{1-\beta^2}}=\frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}}\ . \tag{3-12}$$

Let's insert for $\tilde{x}$ and $c\tau$ from (2-44) in (3-9) and solve for $\tilde{k}$ in terms of k.

$$A\sin(\tilde{k}(\gamma\left(x-\beta ct\right)-\gamma\left(ct-\beta x\right))+\theta)=A\sin\left(\tilde{k}\gamma\left[\left(1-\beta\right)x-c(1+\beta)t\right]\right) \tag{3-13}$$

or

$$A\sin\left(\tilde{k}\gamma\left(1-\beta\right)\left[x-c\frac{(1+\beta)}{\left(1-\beta\right)}t\right]\right). \tag{3-14}$$

The coefficient of x in (3-14) is the wave number, k, in the fixed coordinates. Solving for k in terms of $\tilde{k}$ gives

$$k=\gamma\left(1-\beta\right)\tilde{k}\,. \tag{3-15}$$

Since the energy is proportional to wave number, the energy in the moving coordinates scales according to

$$\tilde{E}_p=\gamma\left(1-\beta\right)E_p \tag{3-16}$$

or

$$\tilde{E}_p=\frac{1-\frac{v}{c}}{\sqrt{1-\left(\frac{v}{c}\right)^2}}E_p\ . \tag{3-17}$$

The above scale factor can be approximated by a series expansion in ${}^{v}\!/\!{}_{c}$ taken only to second order (See note 1 below). This is a good approximation because *c* is large and *v* can be taken to be arbitrarily small. Since the photon energy scales directly with frequency (*v* is negative for photon 2, p2)

$$\tilde{E}_{p1} \cong \left(1 - {}^{v}\!/\!{}_{c} + {}^{v^2}\!/\!{}_{2c^2}\right) E_{p1} = \left(1 - {}^{v}\!/\!{}_{c} + {}^{v^2}\!/\!{}_{2c^2}\right) \frac{\Delta E}{2} \tag{3-18}$$

$$\tilde{E}_{p2} \cong \left(1 + {}^{v}\!/\!{}_{c} + {}^{v^2}\!/\!{}_{2c^2}\right) E_{p2} = \left(1 + {}^{v}\!/\!{}_{c} + {}^{v^2}\!/\!{}_{2c^2}\right) \frac{\Delta E}{2}. \tag{3-19}$$

In the moving coordinate system, the conservation of energy equation (total energy before emitting photons = total energy after emitting photons) is

$$E_b + \frac{1}{2} m v^2 = \left(E_b - \Delta E\right) + \frac{1}{2}\left(m - \Delta m\right) v^2 + \tilde{E}_{p1} + \tilde{E}_{p2} \ . \tag{3-20}$$

Inserting for the photon energies in the moving coordinate system

$$E_b + \frac{1}{2} m v^2 = \left(E_b - \Delta E\right) + \frac{1}{2}\left(m - \Delta m\right) v^2 + \left(1 - {}^{v}\!/\!{}_{c} + {}^{v^2}\!/\!{}_{2c^2}\right) \frac{\Delta E}{2} + \left(1 + {}^{v}\!/\!{}_{c} + {}^{v^2}\!/\!{}_{2c^2}\right) \frac{\Delta E}{2}. \tag{3-21}$$

Simplifying and rearranging

$$E_b + \frac{1}{2} m v^2 = \left(E_b - \Delta E\right) + \frac{1}{2}\left(m - \Delta m\right) v^2 + \left(2 + {}^{v^2}\!/\!{}_{c^2}\right) \frac{\Delta E}{2} \tag{3-22}$$

$$0 = -\frac{1}{2}\left(\Delta m\right) v^2 + \left({}^{v^2}\!/\!{}_{2c^2}\right) \Delta E \tag{3-23}$$

$$\frac{1}{2}\left(\Delta m\right) v^2 = \frac{1}{2}\left({}^{\Delta E}\!/\!{}_{c^2}\right) v^2. \tag{3-24}$$

From the above, the final result is obtained

$$\Delta m = \frac{\Delta E}{c^2} \tag{3-25}$$

or

$$\Delta E = \Delta m c^2 \ . \tag{3-26}$$

In words, the energy emitted by the body is equal to the change in mass times the speed of light squared.

Notes:

1. Doppler Shift Derivation

What is often approximated by a shift in frequency at high frequencies is actually a compression or expansion of the time scale. We desire an approximation for the time compression (or expansion) factor,

$$s = \frac{1 - \frac{v}{c}}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}, \quad \text{(N1-1)}$$

for small values of $\frac{v}{c}$.

A Taylor series of a function, $f(x)$, to second order about a is defined by

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)(x-a)^2}{2!} + \ldots \quad \text{(N1-2)}$$

For the denominator with

$$f(x) = \left(1 - x^2\right)^{-\frac{1}{2}} \quad \text{(N1-3)}$$

$$f'(x) = -\frac{1}{2}\left(1 - x^2\right)^{-\frac{3}{2}}(-2x) = x\left(1 - x^2\right)^{-\frac{3}{2}} \quad \text{(N1-4)}$$

$$\frac{1}{\sqrt{1 - x^2}} = \sum_{k=0}^{\infty} \frac{\prod_{i=1}^{k}(2k-1)}{2^k k!} x^{2k} \ . \quad \text{(N1-5)}$$

With $x = \frac{v}{c}$ then to second order

$$\frac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} \cong 1 + \frac{v^2}{2c^2} \ . \quad \text{(N1-6)}$$

The energy scale factor, *s*, to second order in *v/c* is then

$$s \cong \left(1 - \frac{v}{c}\right)\left(1 + \frac{v^2}{2c^2}\right) = 1 - \frac{v}{c} + \frac{v^2}{2c^2} + \text{ higher order terms.} \quad \text{(N1-7)}$$

2. Kinetic Energy Derivation

Kinetic energy is defined as work so using this and Newton's second law

$$\Delta K = W = F\Delta x = ma\Delta x \ , \qquad \text{(N2-1)}$$

where

ΔK = kinetic energy

W = work

F = force

Δx = distance moved

m = mass

a = acceleration.

Key kinematic equations are summarized below:

a) $v_f = v_f + a\Delta t$ (N2-2)

b) $\Delta x = \frac{1}{2}\left(v_i + v_f\right)\Delta t$ (N2-3)

c) $\Delta x = v_i\Delta t + \frac{1}{2}a\left(\Delta t\right)^2$ (N2-4)

d) $v_f^2 = v_i^2 + 2a\Delta x$ (N2-5)

where

$\Delta t = t_f - t_i$ = delta time (N2-6)

$\Delta x = x_f - x_i$ = delta position (N2-7)

v = velocity (really speed since above are not written as vectors).

Equations a) to c) above come directly from the definitions of position, velocity, and acceleration. Equation d) is obtained by combining a) to c) to eliminate dependence on time. It is equation d) that can be used to obtain the kinetic energy equation directly. Solving d) for $a\Delta x$ gives

$$a\Delta x = \frac{v_f^2 - v_i^2}{2} \ . \qquad \text{(N2-8)}$$

Inserting (N2-8) in (N2-1) yields

$$\Delta K = m\left(\frac{v_f^2 - v_i^2}{2}\right). \qquad \text{(N2-9)}$$

Assuming the initial velocity is zero and the final velocity is just, v, the kinetic energy is

$$E_k = K = \frac{1}{2}mv^2 \ . \qquad \text{(N2-10)}$$

References

1. Einstein, A., 1920. *Relativity: The Special and General Theory*
2. Einstein, A., June 30, 1905. *On the Electrodynamics of Moving Bodies*
3. Tegmark, M., 2006. Relativity. MIT course 8.033 lecture notes. MITOPENCOURSEWARE
4. Einstein, A., September 27, 1905, “Does the Inertia of a Body Depend Upon its Energy-Content?”