Chapter 2. Special Relativity

Notes:

- *Some material presented in this chapter is taken "The Feynman Lectures on Physics, Vol. I" by R. P. Feynman, R. B. Leighton, and M. Sands, Chap. 15 (1963, Addison-Wesley).*

2.1 The Ether and the Michelson-Morley Experiment

As we saw at the end of Chapter 1, the prediction and experimental verification of the propagation of electromagnetic waves in space led physicists to the conclusion that there should exist a medium, the *ether*, permeating all of space on which these waves travelled. Of course, the wave equation we derived from Maxwell's equations (see equation (1.24) in Chapter 1) does not at all require that. It does, in fact, imply that such waves could propagate in vacuum. But this notion of wave propagation without a medium was so counterintuitive to physicists at the time that they elected to postulate the existence of the ether. On the other hand, this hypothesis had the advantage of being testable through experiments.

This is exactly what American physicists **Albert Michelson** (1852-1931) and **Edward Morley** (1838-1923) sought out to do in a famous experiment (i.e., the Michelson-Morley experiment) in 1887. To do so they used an apparatus (now called a Michelson interferometer) as shown in the schematic of **Figure 1**. In a nutshell, the experiment consists of sending a light signal (from Source A in the figure) and splitting it over two mutually orthogonal paths (at the plate B), each propagating through a distance L to a mirror (mirrors C and E in the up-down and left-right directions, respectively) where they are reflected back and recombined ("below" the plate B).

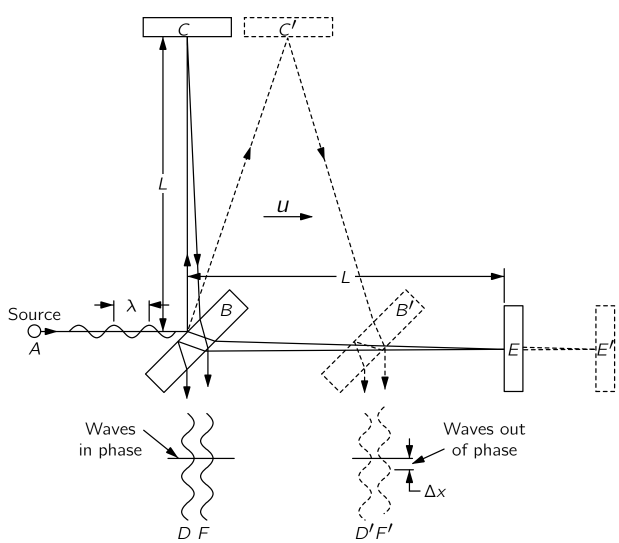

Figure 1 - Schematic diagram of the Michelson-Morley experiment (from *The Feynman Lectures on Physics, Vol. I*).

Michelson and Morley surmised that if the earth is moving relative to the ether at a speed u in the B-E direction, for example, then the electromagnetic wave from the light source travelling along that direction should make its round trip from plate B to mirror E and back to plate B in a different amount of time than the wave propagating along the B-C-B path. More precisely, since the apparatus is moving in the B-E direction the corresponding velocity of the light wave relative to the interferometer should be (according the Newtonian mechanics) equal to $c-u$. It follows that the time needed for the light wave to go from B to E is

$$t_1 = \frac{L}{c-u}. \tag{2.1}$$

Once the wave reflects on mirror E, its velocity relative to the apparatus becomes equal to $c+u$ and the time needed to travel from E to B is

$$t_2 = \frac{L}{c+u}. \tag{2.2}$$

We then find that the time necessary for the round trip is

$$\begin{aligned} t_{\mathrm{E}} &= t_1 + t_2 \\ &= \frac{2L/c}{\left(1-u^2/c^2\right)}. \end{aligned} \tag{2.3}$$

For the light wave going from plate B to mirror C and back the speed relative to the interferometer is simply c, since its velocity is perpendicular to that of the apparatus. The distance the light goes through when going from B to C is given by

$$d_3 = \sqrt{\left(ut_3\right)^2 + L^2}, \tag{2.4}$$

where t_3 is the time needed for the wave to travel that distance. But since we also have $t_3 = d_3/c$, we can write

$$\left(ct_3\right)^2 = \left(ut_3\right)^2 + L^2, \tag{2.5}$$

or

$$t_3 = \frac{L/c}{\sqrt{1-u^2/c^2}}. \tag{2.6}$$

It follows that the round trip B-C-B is

$$t_{\mathrm{C}} = \frac{2L/c}{\sqrt{1-u^2/c^2}}. \tag{2.7}$$

According to the ether hypothesis, the time difference between the two paths should have been

$$\begin{aligned}\Delta t &= t_{\mathrm{C}} - t_{\mathrm{E}} \\ &= \frac{2L/c}{\sqrt{1-u^2/c^2}}\left(1 - 1\Big/\sqrt{1-u^2/c^2}\right) \\ &\neq 0.\end{aligned} \tag{2.8}$$

This non-zero time difference Δt predicted that Michelson and Morley would observe that, when recombined, the two light waves should *not* be in phase with each other. That is, if when they left the plate B at the initial time t_0 the waves had the same amplitude

$$E(t_0) = A\cos(\phi_0), \tag{2.9}$$

then, when recombined, the total amplitude of the signal should had been

$$\begin{aligned}E_{\mathrm{T}}(t) &= A\cos(\omega t_{\mathrm{C}} + \phi_0) + A\cos(\omega t_{\mathrm{E}} + \phi_0) \\ &= A\left[\cos(\omega t_{\mathrm{C}} + \phi_0) + \cos(\omega t_{\mathrm{C}} - \omega\Delta t + \phi_0)\right] \\ &= 2A\cos\left(\frac{\omega\Delta t}{2}\right)\cos\left(\omega t_{\mathrm{C}} + \phi_0 - \frac{\omega\Delta t}{2}\right),\end{aligned} \tag{2.10}$$

where we used the identity $\cos(\theta_1)\cos(\theta_2) = \left[\cos(\theta_1 - \theta_2) + \cos(\theta_1 + \theta_2)\right]\big/2$.

However, the signal Michelson and Morley detected corresponded to

$$E_{\mathrm{T}}(t) = 2A\cos(\omega t_{\mathrm{C}} + \phi_0), \tag{2.11}$$

which implied that $\Delta t = 0$! In other words, the time taken by the light waves to travel their respective paths was the same, just as if the speed of light was the same in both orientations. This result was directly at odds with the idea that electromagnetic waves propagated on the hypothetical ether.

It interesting to note the observation of Lorentz who suggested that the result could be explained if bodies (like the B-E leg of the Michelson interferometer) contracted according to

$$L_{\mathrm{E}} = L\sqrt{1-u^2/c^2}, \tag{2.12}$$

in the direction of their velocity u. It could indeed verified from equation (2.8) that $\Delta t = 0$ in that case even with the existence of the hypothetical ether. Although this suggestion was too artificial to be satisfying to physicists at the time, we will soon see that it incorporated part of the solution to the problem exposed by the Michelson-Morley experiment.

2.2 The Invariance of Maxwell's Equations

Let us now revisit the earlier statement made in Section 1.2.2 of Chapter 1 that Maxwell's equations were not *invariant* under a *Galilean transformation*. More precisely, let us start with the wave equation we derived from Maxwell's equations in Chapter 1 (i.e., equation (1.24))

$$\frac{\partial^2 E_z}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 E_z}{\partial t^2} = 0 \tag{2.13}$$

and let us see how it transforms when going to a reference frame moving at a relative speed u in the positive x-direction such that

$$\begin{aligned} t' &= t \\ x' &= x - ut \\ y' &= y \\ z' &= z. \end{aligned} \tag{2.14}$$

We now use the chain rule to write

$$\begin{aligned} \frac{\partial}{\partial x} &= \frac{\partial x'}{\partial x}\frac{\partial}{\partial x'} + \frac{\partial t'}{\partial x}\frac{\partial}{\partial t'} \\ &= \frac{\partial}{\partial x'} \\ \frac{\partial^2}{\partial x^2} &= \frac{\partial^2}{\partial x'^2} \end{aligned} \tag{2.15}$$

and

$$\begin{aligned} \frac{\partial}{\partial t} &= \frac{\partial t'}{\partial t}\frac{\partial}{\partial t'} + \frac{\partial x'}{\partial t}\frac{\partial}{\partial x'} \\ &= \frac{\partial}{\partial t'} - u\frac{\partial}{\partial x'} \\ \frac{\partial^2}{\partial t^2} &= \left(\frac{\partial}{\partial t'} - u\frac{\partial}{\partial x'}\right)\left(\frac{\partial}{\partial t'} - u\frac{\partial}{\partial x'}\right) \\ &= \frac{\partial^2}{\partial t'^2} - 2u\frac{\partial^2}{\partial x'\partial t'} + u^2\frac{\partial^2}{\partial x'^2}. \end{aligned} \tag{2.16}$$

Inserting the last of equations (2.15) and (2.16) into equation (2.13) yields[1]

$$\left(1-\frac{u^2}{c^2}\right)\frac{\partial^2 E_z}{\partial x'^2}-\frac{1}{c^2}\frac{\partial^2 E_z}{\partial t'^2}+\frac{2u}{c^2}\frac{\partial^2 E_z}{\partial x'\partial t'}=0. \tag{2.17}$$

Because physicists believed then, as we still do now, that the laws of physics should be the same in any and all inertial frames, it follows that the fact that equation (2.17) does not "look" the same as equation (2.13) was a major problem. By "looking" the same we mean that if we replace x and t with x' and t' in equation (2.13) we do not recover equation (2.17), and vice-versa. Indeed, we can see that the two equations are similar only when $u \ll c$.

It was once again Lorentz who pointed out a fact that would eventually be central to resolving this problem. More precisely, he discovered that if one instead used the following coordinate transformation between the two inertial frames

$$\begin{aligned} t' &= \frac{t - xu/c^2}{\sqrt{1-u^2/c^2}} \\ x' &= \frac{x-ut}{\sqrt{1-u^2/c^2}} \\ y' &= y \\ z' &= z, \end{aligned} \tag{2.18}$$

then Maxwell's equations, and equation (2.13), are invariant (i.e., look the same) when going from one to the other (or any other) inertial frames. The transformation given by equations (2.18) is the so-called *Lorentz transformation*, which Fitzgerald had previously and independently derived. We can also see that we recover the Galilean transformation of equations (2.14) when $u \ll c$, implying that the Lorentz transformation may be more general in scope and application.

Physicists had two sets of laws that appeared to work extremely well in predicting the outcome of experiments: Newton's Laws and Maxwell's equations. But to make things even more confusing, the first set of laws were invariant under a Galilean transformation but not under a Lorentz transformation, while the opposite was true for the second set of equations… Because of the success and long standing status of Newton's Laws, many physicists believed that something had to be wrong with Maxwell's equations.

2.3 Einstein and Special Relativity

---

[1] In the transformation of equation (2.17) we should also account for the corresponding transformation of the electric field components. But we will neglect this for the purpose of the present discussion.

In one of his groundbreaking papers of 1905 Einstein showed that Newton's Laws only applied in non-relativistic situations, i.e., when $u \ll c$, and needed to be corrected otherwise. He based his analysis on two fundamental postulates:

I. The laws of Physics are the same in all inertial frames. This is the *principle of relativity*.
II. The speed of light in the vacuum is the same in all inertial frames, irrespective of the velocity of the source responsible for the corresponding radiation.

To reach his conclusions, Einstein considered two inertial frames in uniform relative motion, at the speed u, along the x-direction, as displayed in Figure 2. One main results was that *the Lorentz transformation of equations (2.18) is the correct relationship between the coordinates of the two inertial frames*, not the Galilean transformation. The constancy of the speed of light can easily be verified by considering the distance travelled by a ray of light emitted at an arbitrary direction at $t' = 0$ in Moe's frame $\sqrt{x'^2 + y'^2 + z'^2} = ct'$, which we can rewrite as

$$c^2 t'^2 - x'^2 - y'^2 - z'^2 = 0. \tag{2.19}$$

Let us now insert equations (2.18) into (2.19), with $\gamma = \left(1 - u^2/c^2\right)^{-1/2} \geq 1$,

$$\begin{aligned} c^2 t'^2 - x'^2 - y'^2 - z'^2 &= c^2\gamma^2\left(t - xu/c^2\right)^2 - \gamma^2\left(x - ut\right)^2 - y^2 - z^2 \\ &= \gamma^2\left[t^2\left(c^2 - u^2\right) - x^2\left(1 - u^2/c^2\right) - 2tx\left(u - u\right)\right] - y^2 - z^2 \\ &= \gamma^2\left[\gamma^{-2}\left(c^2 t^2 - x^2\right)\right] - y^2 - z^2 \\ &= c^2 t^2 - x^2 - y^2 - z^2 \\ &= 0, \end{aligned} \tag{2.20}$$

and we thus verify that the Lorentz transformation ensures the constancy of the speed of

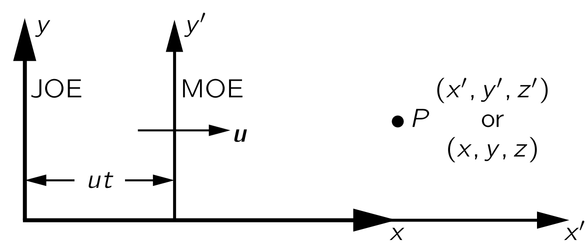

Figure 2 – Two inertial frames. The primed coordinate system (i.e., Moe's) is moving with a velocity u relative to the unprimed system (i.e., Joe's) in the x-direction (from *The Feynman Lectures on Physics, Vol. I*).

light in the two inertial frames. A simplified derivation of the Lorentz transformation will be provided in the Appendix at the end of the chapter, but Einstein also showed, and as can be seen in equations (2.20), the quantity $c^2t^2 - \left(x^2 + y^2 + z^2\right)$ is not only invariant for light, but also for any other system or conditions. That is, invariance also holds when $c^2t^2 - \left(x^2 + y^2 + z^2\right) \neq 0$.

We also note that the first postulate implies that Maxwell's equations are good laws of physics since they are invariant to the Lorentz transformation (therefore Newton's Laws are not). We now look at a few of the important implications stemming from Einstein's theory of Special Relativity and the Lorentz transformation.

2.3.1 Time Dilation and Length Contraction

Let us consider a time interval $\Delta t = t_2 - t_1$ measured with a clock at rest in Joe's inertial frame of Figure 2 (i.e., the unprimed coordinates), and inquire about the corresponding time interval $\Delta t' = t_2' - t_1'$ as measured by a similar clock at rest in Moe's reference frame (the primed coordinates). Using the Lorentz transformation we write

$$\begin{aligned} t_2' - t_1' &= \gamma\left(t_2 - x_2 u/c^2\right) - \gamma\left(t_1 - x_1 u/c^2\right) \\ &= \gamma\left[\left(t_2 - t_1\right) - \left(x_2 - x_1\right)u/c^2\right] \\ &= \gamma\left(t_2 - t_1\right), \end{aligned} \tag{2.21}$$

since $x_1 = x_2$ because the clock is at rest in Joe's frame. We therefore find that

$$\begin{aligned} \Delta t' &= \frac{\Delta t}{\sqrt{1 - u^2/c^2}} \\ &> \Delta t. \end{aligned} \tag{2.22}$$

It thus appears that the clock at rest in Joe's frame *runs slower* than the clock at rest in Moe's frame (according to Moe)! Let us now proceed in the opposite way. But to calculate Δt as a function of $\Delta t'$ we must first determine the "inverse" of the Lorentz transformation given in equations (2.18). This is easily done, as we only need to change the sign of the relative velocity. That is, if Moe's is moving at the velocity u relative to Joe, then Joe moves at $-u$ relative to Moe. The corresponding Lorentz transform is therefore

$$\begin{aligned} t &= \gamma\left(t' + x'u/c^2\right) \\ x &= \gamma\left(x' + ut'\right) \\ y &= y' \\ z &= z'. \end{aligned} \tag{2.23}$$

We can now calculate

$$\begin{aligned} t_2 - t_1 &= \gamma\left(t_2' + x_2' u/c^2\right) - \gamma\left(t_1' + x_1' u/c^2\right) \\ &= \gamma\left[\left(t_2' - t_1'\right) + \left(x_2' - x_1'\right) u/c^2\right] \\ &= \gamma\left(t_2' - t_1'\right), \end{aligned} \tag{2.24}$$

with the last step resulting from the fact that $x_1' = x_2'$ for the clock at rest in Moe's frame. We are then left with

$$\begin{aligned} \Delta t &= \frac{\Delta t'}{\sqrt{1 - u^2/c^2}} \\ &> \Delta t'. \end{aligned} \tag{2.25}$$

How could this be? That is, how can the clock at rest relative to Joe appears to be running slower according to Moe (from equation (2.22)), while the clock at rest relative to Moe also appears to run slower (not faster) relative to Joe (from equation (2.25))? The answer is that, according to the principle of relativity, we should not have been expecting anything else. More precisely, a clock at rest in an inertial frame moving at some uniform velocity relative to another inertial frame will always appear to run slower according to observers at rest in the latter. That is, *moving clocks run slower…*

Let us now consider the ends of a rod of length $L_0 = x_2 - x_1$ laying along the x-axis in Joe's frame, and inquire about its length $L' = x_2' - x_1'$ as measured by Moe. We therefore write

$$\begin{aligned} x_2 - x_1 &= \gamma\left(x_2' + u t_2'\right) - \gamma\left(x_1' + u t_1'\right) \\ &= \gamma\left[\left(x_2' - x_1'\right) + u\left(t_2' - t_1'\right)\right] \\ &= \gamma\left(x_2' - x_1'\right), \end{aligned} \tag{2.26}$$

since to measure the length of the rod in his frame Moe must measure the position of its ends at the same time $t_1' = t_2'$ according to *his clock*. We thus find that

$$\begin{aligned} L' &= L_0\sqrt{1 - u^2/c^2} \\ &\le L_0. \end{aligned} \tag{2.27}$$

We therefore recover the result anticipated by Lorentz that *moving rods appear to contract* (i.e., Joe would also measure a rod at rest relative to Moe to be shorter…). This apparently peculiar result is just a consequence of the fact that, in special relativity, events that are simultaneous (i.e., happen at the same time) in one reference frame will not be in another (if the two frames are moving relative to one another). *The concept of simultaneity must be abandoned in special relativity.*

2.3.2 Addition of Velocities

It is straightforward to extend the Lorentz transformation given in equations (2.23) to infinitesimal quantities the dt, dx, etc., with

$$\begin{aligned} dt &= \gamma\left(dt' + dx'u/c^2\right) \\ dx &= \gamma\left(dx' + udt'\right) \\ dy &= dy' \\ dz &= dz'. \end{aligned} \tag{2.28}$$

We can now inquire of the velocity of a particle moving at a velocity $\mathbf{v}'$ relative to Moe's inertial frame, as seen by Joe. We thus consider the following

$$\begin{aligned} v_x &= \frac{dx}{dt} \\ &= \frac{\gamma\left(dx' + udt'\right)}{\gamma\left(dt' + dx'u/c^2\right)} \\ &= \frac{u + v'_x}{1 + v'_x u/c^2} \\ v_y &= \frac{dy}{dt} \\ &= \frac{dy'}{\gamma\left(dt' + dx'u/c^2\right)} \\ &= \frac{v'_y}{\gamma\left(1 + v'_x u/c^2\right)} \\ v_z &= \frac{dz}{dt} \\ &= \frac{dz'}{\gamma\left(dt' + dx'u/c^2\right)} \\ &= \frac{v'_z}{\gamma\left(1 + v'_x u/c^2\right)}. \end{aligned} \tag{2.29}$$

Although these relation are more complicated than the simple addition of velocity law obtained with the Galilean transformation in Newtonian physics, we do recover the same relations when $u, |\mathbf{v}'| \ll c$ (i.e., in the non-relativistic case). It is also straightforward to show that

$$
\begin{aligned}
v'_x &= \frac{v_x - u}{1 - v_x u/c^2} \\
v'_y &= \frac{v_y}{\gamma\left(1 - v_x u/c^2\right)} \\
v'_z &= \frac{v_z}{\gamma\left(1 - v_x u/c^2\right)}.
\end{aligned}
\tag{2.30}
$$

It is also clear from equations (2.29) and (2.30) that speed of the particle cannot exceed the speed of light in any inertial frame.

2.3.3 Relativistic Kinematics

Let us consider a simple example where Moe, in the primed inertial frame, holds up a ball that he measures to have a mass m_1 (when at rest in that frame), and drops it on a similar ball, let us call it m_2, at rest in Joe's inertial frame (the unprimed frame). The situation is still as shown in Figure 2 with Moe's frame moving at the uniform velocity $\mathbf{u}$ in the x'- (and x-) direction. At the moment when the two balls are about to be aligned vertically, Moe throws his ball m_1 downward at a vertical speed v'_{1y} (measured in Moe's frame) such that it elastically strikes Joe's m_2 ball head-on. After the collision, m_1 comes to a stop (as seen in Moe's frame) while m_2 starts moving downwards at a speed v_{2y} as measured in Joe's frame. We know from Newtonian mechanics that such a case of complete transfer of linear momentum from one mass to another during an elastic collision happens when the colliding partners have the same mass, i.e., $m_1 = m_2$. Let us now inquire whether Newton's definition of linear momentum, i.e., $\mathbf{p} = m\mathbf{v}$, applies to relativistic motions.

Although we know from equations (2.29) and (2.30) that Moe and Joe will not agree on the speeds of the two balls, they will however agree that before the collision m_1 is moving vertically and m_2 is not while the opposite is true after the collision. And they should therefore also agree on the fact that linear momentum should be conserved in the process. In Joe's frame the equality of the before- and after-collision momenta is written as

$$
m_1 \frac{v'_{1y}}{\gamma_u} = m_2 v_{2y} \tag{2.31}
$$

or

$$
m_1 v'_{1y} = m_2 \gamma_u v_{2y}, \tag{2.32}
$$

where we used equations (2.29) ($v'_{1x} = 0$) with $\gamma_u = \left(1 - u^2/c^2\right)^{-1/2}$. In Moe's frame we have

$$m_1 v'_{1y} = m_2 \frac{v_{2y}}{\gamma_u}, \tag{2.33}$$

from equations (2.30) since $v_{2x} = 0$. Equating equations (2.32) and (2.33) we are faced with the impossibility

$$m_2 \gamma_u v_{2y} = m_2 \frac{v_{2y}}{\gamma_u}, \tag{2.34}$$

which is clearly erroneous. We note that the mass m_2 on the left-hand side is that obtained from Joe's measurement in the unprimed inertial frame, while the one on the right-hand side is from Moe's frame. The only conclusion we can draw from this is that, if the conservation of momentum is to hold, then, *the measured mass is not the same the two inertial frames*!

Let us further calculate the following for the speed of m_1 as seen by Joe

$$\begin{aligned} v_1^2 &= u^2 + \frac{v_{1y}'^2}{\gamma_u^2} \\ &= u^2 + \gamma_u^{-2} v_1'^2, \end{aligned} \tag{2.35}$$

since $v'_1 = v'_{1y}$ in Moe's frame (i.e., $v'_{1x} = 0$), and therefore

$$\begin{aligned} \gamma_{v_1}^{-2} &= 1 - \frac{v_1^2}{c^2} \\ &= 1 - \frac{u^2}{c^2} - \gamma_u^{-2} \frac{v_1'^2}{c^2} \\ &= \gamma_u^{-2} \left(1 - \frac{v_1'^2}{c^2} \right) \end{aligned} \tag{2.36}$$

or

$$\gamma_{v_1} = \gamma_{v'_1} \gamma_u. \tag{2.37}$$

We also find through a similar exercise that[2]

[2] Note that equations (2.37) and (2.38) are not valid in general but only apply to the problem at hand. More general considerations show that $\gamma_{v_1} = \gamma_{v'_1} \gamma_u \left(1 + \mathbf{u} \cdot \mathbf{v}'_1 / c^2 \right)$ and $\gamma_{v'_2} = \gamma_{v_2} \gamma_u \left(1 - \mathbf{u} \cdot \mathbf{v}_2 / c^2 \right)$, respectively.

$$\gamma_{v_2'} = \gamma_{v_2}\gamma_u. \tag{2.38}$$

Since both masses are assumed to be the same when measured in their rest frame, we write $m_1 = m_2 = m_0$, and attempt a new definition for the linear momentum

$$\mathbf{p} = \gamma_v m_0 \mathbf{v}. \tag{2.39}$$

Coming back to problem, we now have for the conservation of (vertical) linear momentum in Joe's frame

$$\gamma_{v_1} m_0 \frac{v_{1y}'}{\gamma_u} = \gamma_{v_2} m_0 v_{2y}, \tag{2.40}$$

or with equation (2.37)

$$\gamma_{v_1'} m_0 v_{1y}' = \gamma_{v_2} m_0 v_{2y}. \tag{2.41}$$

Likewise, in Moe's frame we have

$$\gamma_{v_1'} m_0 v_{1y}' = \gamma_{v_2'} m_0 \frac{v_{2y}}{\gamma_u}, \tag{2.42}$$

or from equation (2.38)

$$\gamma_{v_1'} m_0 v_{1y}' = \gamma_{v_2} m_0 v_{2y}. \tag{2.43}$$

The fact that equations (2.41) and (2.43) are the same shows that the new definition for the linear momentum (i.e., equation (2.39)) ensures that Joe and Moe agree on its conservation in a consistent manner, as expected. Incidentally, we note that equation (2.43) also implies that $v_{1y}' = v_{2y}$, which we could have intuitively deduced. Equation (2.39) is the correct definition for the linear momentum.

It is often the case that the special relativistic linear momentum is written in a way similar to the Newtonian form as

$$\mathbf{p} = m\mathbf{v}, \tag{2.44}$$

with the *relativistic mass* given by

$$\begin{aligned} m &= \gamma_v m_0 \\ &= \frac{m_0}{\sqrt{1 - v^2/c^2}}. \end{aligned} \tag{2.45}$$

A distinction is then made with the *rest mass* m_0, which is the same for all inertial frames (i.e., it is the mass when the corresponding object is not moving relative to the observer). We therefore see that *the relativistic mass increases as it gains speed*, since $\gamma_v > 1$ when $v > 0$.

What about the energy of the particle? Here we use the same definition as in Newtonian physics while retaining equation (2.39) (or (2.44) and (2.45)) for the linear momentum. That is, the force acting on a particle is still given by

$$\begin{aligned}\mathbf{F} &= \frac{d\mathbf{p}}{dt} \\ &= \frac{d\left(\gamma_v m_0 \mathbf{v}\right)}{dt} \\ &= m_0 \frac{d\left(\gamma_v \mathbf{v}\right)}{dt},\end{aligned} \tag{2.46}$$

while it acquires the following kinetic energy when acted upon by the force starting from rest at $\mathbf{x} = 0$ to some final position $\mathbf{x}_\mathrm{f}$

$$K = \int_0^{\mathbf{x}_\mathrm{f}} \mathbf{F} \cdot d\mathbf{x}. \tag{2.47}$$

Inserting equation (2.46) into equation (2.47) yields

$$\begin{aligned}K &= m_0 \int_0^{\mathbf{x}_\mathrm{f}} \frac{d\left(\gamma_v \mathbf{v}\right)}{dt} \cdot d\mathbf{x} \\ &= m_0 \int_0^{(\gamma_v \mathbf{v})_\mathrm{f}} \mathbf{v} \cdot d\left(\gamma_v \mathbf{v}\right) \\ &= m_0 \int_0^{(\gamma_v v)_\mathrm{f}} v\, d\left(\gamma_v v\right),\end{aligned} \tag{2.48}$$

since $d\mathbf{x} = \mathbf{v}dt$ and where at $\mathbf{x}_\mathrm{f}$ the particle has acquired the final linear momentum per unit mass $\mathbf{p}_\mathrm{f}/m_0 = \left(\gamma_v \mathbf{v}\right)_\mathrm{f}$. To solve this integral we first calculate

$$\begin{aligned}\frac{d\left(\gamma_v v\right)}{dv} &= \frac{d}{dv}\left(\frac{v}{\sqrt{1 - v^2/c^2}}\right) \\ &= \frac{1}{\sqrt{1 - v^2/c^2}} + \frac{v^2/c^2}{\left(1 - v^2/c^2\right)^{3/2}} \\ &= \frac{1}{\left(1 - v^2/c^2\right)^{3/2}},\end{aligned} \tag{2.49}$$

which we then insert in equation (2.48) to get

$$\begin{aligned} K &= m_0 \int_0^{v_\mathrm{f}} \frac{v\,dv}{\left(1 - v^2/c^2\right)^{3/2}} \\ &= m_0 \left. \frac{c^2}{\sqrt{1 - v^2/c^2}} \right|_0^{v_\mathrm{f}} \\ &= m_0 c^2 \left(\gamma_{v_\mathrm{f}} - 1\right). \end{aligned} \tag{2.50}$$

This equation for the relativistic kinetic energy is obviously different than the Newtonian version. However, in the limit when $v_\mathrm{f} \ll c$ we have

$$\begin{aligned} \gamma_{v_\mathrm{f}} &= \frac{1}{\sqrt{1 - v_\mathrm{f}^2/c^2}} \\ &\simeq 1 + \frac{1}{2}\frac{v_\mathrm{f}^2}{c^2}, \end{aligned} \tag{2.51}$$

and therefore

$$\lim_{v_\mathrm{f} \ll c} K = \frac{1}{2} m_0 v_\mathrm{f}^2. \tag{2.52}$$

We thus recover the Newtonian result in the non-relativistic limit.

The final form of equation (2.50) led Einstein to suggest that the *total* energy of the particle is given by

$$\begin{aligned} E &= m_0 c^2 + K \\ &= \gamma_v m_0 c^2, \end{aligned} \tag{2.53}$$

where we have dropped the ‘f’ subscript for convenience. This implies that a particle at rest, i.e., with $K = 0$, has a latent *rest energy*

$$E_0 = m_0 c^2. \tag{2.54}$$

This rest energy is extremely large for a non-relativistic particle since we can verify that $K/E_0 = v^2/2c^2 \ll 1$ when $v \ll c$, and its existence will have very important ramifications when discussing nuclear physics later on. When using equation (2.45), the energy equation can also be written as what is perhaps the most famous equation in all of physics

$$E = mc^2, \tag{2.55}$$

stating the *equivalence between mass and energy*.

Finally, we note that we can combine the energy and momentum as follows

$$\begin{aligned} E^2 - p^2c^2 &= \gamma_v^2 m_0^2 c^4 - \gamma_v^2 m_0^2 v^2 c^2 \\ &= \gamma_v^2 m_0^2 c^4 \left(1 - v^2/c^2\right) \end{aligned} \tag{2.56}$$

or, using $\gamma_v^{-2} = 1 - v^2/c^2$,

$$E^2 = m_0^2 c^4 + p^2 c^2. \tag{2.57}$$

Appendix – The Lorentz Transformation

We can provide a mathematical derivation of the Lorentz transformation for the system shown in Figure 2 as follows. Because of the homogeneity of space-time, we will assume that the different coordinates of the two frames are linked by a set of linear relations. For example, we write

$$\begin{aligned} ct' &= Act + Bx + Cy + Dz \\ x' &= Ect + Fx + Gy + Hz, \end{aligned} \tag{2.58}$$

and similar equations for y' and z'. However, since the two inertial frames exhibit a relative motion only along the x-axis, we will further assume that the directions perpendicular to the direction of motion are the same for both systems with

$$\begin{aligned} y' &= y \\ z' &= z. \end{aligned} \tag{2.59}$$

Furthermore, because we consider that these perpendicular directions should be unchanged by the relative motion, and that at low velocity (i.e., when $u \ll c$) we must have

$$x' = x - ut, \tag{2.60}$$

we will also assume that the transformations do not "mix" the parallel and perpendicular components. That is, we set $C = D = G = H = 0$ and simplify equations (2.58) to

$$\begin{aligned} ct' &= Act + Bx \\ x' &= Ect + Fx. \end{aligned} \tag{2.61}$$

Therefore, we only need to solve for the relationship between (ct', x') and (ct, x). To do so, we first consider a particle at rest at the origin of Joe's referential such that $x = 0$ and its velocity as seen by an observer at rest in Moe's frame is $-u$. Using equations (2.61) we find that

$$\frac{x'}{ct'} = -\frac{u}{c} = \frac{E}{A}. \tag{2.62}$$

Second, we consider a particle at rest at the origin of Moe's referential such that now $x' = 0$ and its velocity as seen by Joe is u. This time we find from the last of equations (2.61) that

$$\frac{x}{ct} = \frac{u}{c} = -\frac{E}{F}, \tag{2.63}$$

and the combination of equations (2.62) and (2.63) shows that $A = F$; we rewrite equations (2.61) as

$$\begin{aligned} ct' &= A\left(ct + \frac{B}{A}x\right) \\ x' &= A\left(x - \frac{u}{c}ct\right). \end{aligned} \tag{2.64}$$

Third, we note that from Postulate II the propagation of a light pulse must happen at the speed of light in *both* inertial frames. We then set $ct' = x'$ and $ct = x$ in equations (2.64) to find that

$$\frac{B}{A} = -\frac{u}{c}, \tag{2.65}$$

and

$$\begin{aligned} ct' &= A\left(ct - \frac{u}{c}x\right) \\ x' &= A\left(x - \frac{u}{c}ct\right). \end{aligned} \tag{2.66}$$

Evidently, we could have instead proceeded by first expressing the unprimed coordinates as a function of the primed coordinates

$$\begin{aligned} ct &= A'ct' + B'x' \\ x &= E'ct' + F'x', \end{aligned} \tag{2.67}$$

from which, going through the same process as above, we would have found that

$$
\begin{aligned}
ct &= A'\left(ct' + \frac{u}{c}x'\right) \\
x &= A'\left(x' + \frac{u}{c}ct'\right).
\end{aligned}
\tag{2.68}
$$

Not surprisingly, equations (2.68) are similar in form to equations (2.66) with u replaced by $-u$. The first postulate of special relativity tells us, however, that the laws of physics must be independent of the inertial frame. This implies that

$$
A = A' \tag{2.69}
$$

(this result is verified by inserting equations (2.66) and (2.68) into $c^2t'^2 - x'^2 = c^2t^2 - x^2$, as this will yield $A^2 = A'^2$). If we insert equations (2.68) into equations (2.66) we find that

$$
A = \left[1 - \left(\frac{u}{c}\right)^2\right]^{-1/2}. \tag{2.70}
$$

We can finally write the Lorentz transformation, in its usual form for the problem at hand, as

$$
\begin{aligned}
ct' &= \gamma_u\left(ct - \beta x\right) \\
x' &= \gamma_u\left(x - \beta ct\right) \\
y' &= y \\
z' &= z,
\end{aligned}
\tag{2.71}
$$

with

$$
\begin{aligned}
\boldsymbol{\beta} &= \frac{\mathbf{u}}{c} \\
\beta &= \left|\boldsymbol{\beta}\right| \\
\gamma_u &= \left(1 - \beta^2\right)^{-1/2}.
\end{aligned}
\tag{2.72}
$$

The inverse transformation is easily found by swapping the two sets of coordinates, and by changing the sign of the velocity. We then get

$$
\begin{aligned}
ct &= \gamma_u\left(ct' + \beta x'\right) \\
x_1 &= \gamma_u\left(x' + \beta ct'\right) \\
y &= y' \\
z &= z'.
\end{aligned}
\tag{2.73}
$$