

Relativity

N.G. Schultheiss

translated and adapted by K. Schadenberg

1 Introduction

In this module will try to explain what happens when two observers look at the same event. Every observer has his own coordinate system and clock ('frame of reference') with which the event is observed. These frames of references are not independent of each other. In this text we will deduce the relation between the two.

Some quantities are *relative*, they depend on the observer, or more specifically the coordinate system of the observer. Other quantities are independent of the coordinate system, these are called *invariant*. With this starting point we end up with Einstein's famous equation $E = mc^2$. In this text the entire derivation of this formula is written down. Study it carefully so you can make similar derivations on your own in the future.

2 Lorentz and Poincaré

Hendrik Antoon Lorentz is one of many Dutch Nobel prize laureates, he won the prize together with Pieter Zeeman. You might know Lorentz from the (equation for) Lorentz force $\mathbf{F}_l = l(\mathbf{I} \cdot \mathbf{B})$. Lorentz won the Nobel prize for the discovery and theoretical explanation of the Zeeman effect.¹

Lorentz was one of the first scientists to work on a theory of relativity (although it was not yet called relativity). To compare the observations of two observers we need to relate their frames of reference. This can be done with the use of the Galilean transformation:

$$\begin{aligned}x &= x' + vt' \\ y &= y' \\ z &= z' \\ t &= t'\end{aligned}\tag{2.1}$$

Quantities without a prime mark are observed by the observer in the coordinate system without a prime mark. Quantities with a prime mark ' are observed by the second observer' in the coordinate system' with prime mark.² In this example only the

¹The Zeeman effect describes the interaction between magnetic fields and the spin of particles. See module 'Fluorescence' for more details.

²The prime mark is also used to denote the derivative of a function. When using the notation of equation 2.1 it is then unclear which derivative is meant, time or a specific space? To avoid confusion it is best to denote the derivative to a specific variable in the following manner: The derivative of the function $f(x, y, z, t)$ with respect to the time t can be written as: $\frac{df(x, y, z, t)}{dt}$.

x-coordinate changes. This means that the second observer' moved away from the first observer in the x direction with a certain speed v .

Galileo Galilei assumed that time is the same in all frames of reference. However, when one approaches the speed of light this is no longer the case. The speed of light is a constant in the universe (invariant). Using the Galilean transformation we break this rule. We therefore need a different transformation method in which the speed of light remains constant. In this text we will only be looking at changes along the x-axis but all equations below can be expanded to also be valid for changes in other directions:

$$\begin{aligned} c &= \frac{\Delta x}{\Delta t} \\ c &= \frac{\Delta x'}{\Delta t'} \end{aligned} \tag{2.2}$$

Lets look at a beam of light originating from the origin at $t = 0$ travelling along the x-axis. We can use the following equations to calculate when this beam of light is observed at a certain position:

$$\begin{aligned} x &= ct \\ x' &= ct' \end{aligned} \tag{2.3}$$

$$\begin{aligned} x - ct &= 0 \\ x' - ct' &= 0 \end{aligned} \tag{2.4}$$

The length of time can differ between the two frames, this allows us the write:

$$x' - ct' = \lambda(x - ct) \tag{2.5}$$

For a beam of light travelling in the opposite direction:

$$x' + ct' = \mu(x + ct) \tag{2.6}$$

Both equations are valid at the same time. Lets see if we can solve them.

- The sum of equations 2.5 and 2.6 yields:

$$x' - ct' + x' + ct' = \lambda(x - ct) + \mu(x + ct) \tag{2.7}$$

$$2x' = (\lambda + \mu)x - (\lambda - \mu)ct \tag{2.8}$$

$$x' = \frac{\lambda + \mu}{2}x - \frac{\lambda - \mu}{2}ct \tag{2.9}$$

- The difference of equations 2.5 and 2.6 yields:

$$x' - ct' - (x' + ct') = \lambda(x - ct) - \mu(x + ct) \tag{2.10}$$

$$-2ct' = (\lambda - \mu)x - (\lambda + \mu)ct \tag{2.11}$$

$$ct' = -\frac{\lambda - \mu}{2}x + \frac{\lambda + \mu}{2}ct \tag{2.12}$$

By introducing two new constants we can write the above equation in a more concise form:

$$a = \frac{\lambda + \mu}{2} \quad (2.13)$$

$$b = \frac{\lambda - \mu}{2} \quad (2.14)$$

$$x' = ax - bct \quad (2.15)$$

$$ct' = -bx + act \quad (2.16)$$

Lets look at how fast the (coordinate) system' is moving with respect to the system. We set $x' = 0$:

$$0 = ax - bct \quad (2.17)$$

$$v = \frac{x}{t} = \frac{bc}{a} \quad (2.18)$$

$$\frac{v}{c} = \frac{b}{a} \quad (2.19)$$

We can use this to expand equation 2.15:

$$x' = a \left(x - \frac{b}{a} ct \right) \quad (2.20)$$

$$x' = a \left(x - \frac{v}{c} ct \right) \quad (2.21)$$

In a similar fashion we can rearrange equation 2.16:

$$ct = \frac{ct'}{a} + \frac{bx}{a} \quad (2.22)$$

$$ct = \frac{ct'}{a} + \frac{vx}{a} \quad (2.23)$$

Substitution yields:

$$x' = a \left(x - \frac{v}{c} \left(\frac{ct'}{a} + \frac{vx}{a} \right) \right) \quad (2.24)$$

$$x' = a \left(1 - \left(\frac{v}{c} \right)^2 \right) x - a \frac{v}{c} \frac{ct'}{a} \quad (2.25)$$

$$x' = a \left(1 - \left(\frac{v}{c} \right)^2 \right) x - vt' \quad (2.26)$$

We set our clocks in such a way that $t = t' = 0$. For symmetry reasons a standard length which moves along with one coordinates system will always be observed in the same way from a second system'. From equation 2.15 follows:

$$x' = ax \quad (2.27)$$

This means that a rod with a length of 1 (metre) in system' will appear to have a length

$$x = \frac{1}{a} \quad (2.28)$$

in system.

Because we set both times at zero, equation 2.26 becomes:

$$x' = a \left(1 - \left(\frac{v}{c} \right)^2 \right) x \quad (2.29)$$

When one observes a rod of length 1 in system from system', the rod will appear to have length:

$$x' = a \left(1 - \left(\frac{v}{c} \right)^2 \right) \quad (2.30)$$

This transformation works both ways:

$$x = x' \quad (2.31)$$

$$\frac{1}{a} = a \left(1 - \left(\frac{v}{c} \right)^2 \right) \quad (2.32)$$

$$a = \frac{1}{\sqrt{1 - \left(\frac{v}{c} \right)^2}} \quad (2.33)$$

Using equation 2.19:

$$b = \frac{\frac{v}{c}}{\sqrt{1 - \left(\frac{v}{c} \right)^2}} \quad (2.34)$$

Substituting these formula in equations 2.15 and 2.16 yields:

$$x' = \frac{x - vt}{\sqrt{1 - \left(\frac{v}{c} \right)^2}} \quad (2.35)$$

$$t' = \frac{t - \frac{v}{c^2}x}{\sqrt{1 - \left(\frac{v}{c} \right)^2}} \quad (2.36)$$

This transformation, the Lorentz transformation, is different from the Galilean transformation. There is a new constant involved in the transformation, the Lorentz factor:

$$\gamma = \frac{1}{\sqrt{1 - \left(\frac{v}{c} \right)^2}} \quad (2.37)$$

Lorentz was a supporter of the 'aether' theory and therefore had some trouble accepting his own equations. Jules Henri Poincaré had contact with Lorentz and together they improved the new theory of what would become Lorentz transformations and ultimately

relativity. One of Poincaré's suggestions was to imagine space as four dimensional (with time as one of the dimensions). The Lorentz transformations then describe this four dimensional world. If we define time as being imaginary in this space¹, vectors always keep their length when transforming from one frame to another:

$$x^2 + y^2 + z^2 - (ct)^2 = x'^2 + y'^2 + z'^2 - (ct')^2 \quad (2.38)$$

3 Invariants

In the previous sections we saw that quantities can change when we change our frame of reference. Not all quantities do so, those who do not are called invariant. Michelson and Morley showed with their experiment² that the speed of light is invariant.

We can describe the state of an object using a position vector and time. Remember that most quantities are relative such as these position vector and time. A more useful way of describing an object is by stating its momentum and energy. These two quantities can (and are) also used when describing collisions. Both momentum and location consist of three directions: x , y , and z . Time, according to Poincaré and Minkowski, is a fourth direction. When talking about momentum, energy can be seen as the fourth dimension.

The state of an object can thus be described as a four dimensional vector; 4-vector. The four components of a momentum/energy vector can be transformed from one system to another system' which moves along its x -axis using the following equations:

$$\begin{aligned} p'_x &= \gamma \left(p_x - \beta \frac{E}{c} \right) \\ p'_y &= p_y \\ p'_z &= p_z \\ \frac{E'}{c} &= \gamma \left(\frac{E}{c} - \beta p_x \right) \end{aligned} \quad (3.1)$$

In these equations p denotes momentum and E denotes energy. β is a new variable, it describes how fast the systems are moving with respect to each other $\beta = \frac{v}{c}$ (compare this to the earlier variables a and b). If we want to convert from the second system' to the first β will be negative. Using our new variable we can rewrite the Lorentz factor as $\gamma = \frac{1}{\sqrt{1-\beta^2}}$.

The length of the 4-vector can be determined using Pythagoras' theorem. If we take time or energy to be imaginary our 4-vector is invariant. Squaring the time or energy component will yield a negative number. Apparently the following equality holds:

$$p_{x'}^2 + p_{y'}^2 + p_{z'}^2 - \left(\frac{E'}{c} \right)^2 = p_x^2 + p_y^2 + p_z^2 - \left(\frac{E}{c} \right)^2 \quad (3.2)$$

We can verify if this is indeed true by substituting the equations from 3.1 into 3.2:

$$\left(\gamma \left(p_x - \beta \frac{E}{c} \right) \right)^2 + p_y^2 + p_z^2 - \left(\gamma \left(\frac{E}{c} - \beta p_x \right) \right)^2 = p_x^2 + p_y^2 + p_z^2 - \left(\frac{E}{c} \right)^2 \implies (3.3)$$

¹Imaginary numbers are denoted by the symbol i , i is defined as $i = \sqrt{-1}$.

²See module 'Michelson Morley' for a description of the experiment and their results.

$$\left(\gamma\left(p_x - \beta\frac{E}{c}\right)\right)^2 - \left(\gamma\left(\frac{E}{c} - \beta p_x\right)\right)^2 = p_x^2 - \left(\frac{E}{c}\right)^2 \implies \quad (3.4)$$

$$\gamma^2\left(p_x^2 - 2p_x\beta^2\left(\frac{E}{c}\right)^2\right) - \gamma^2\left(\left(\frac{E}{c}\right)^2 - 2p_x\beta\frac{E}{c} + \beta^2 p_x^2\right) = p_x^2 - \left(\frac{E}{c}\right)^2 \implies \quad (3.5)$$

$$\gamma^2 p_x^2 - 2\gamma^2 p_x \beta \frac{E}{c} + \gamma^2 \beta^2 \left(\frac{E}{c}\right)^2 - \gamma^2 \left(\frac{E}{c}\right)^2 + 2\gamma^2 p_x \beta \frac{E}{c} - \gamma^2 \beta^2 p_x^2 = p_x^2 - \left(\frac{E}{c}\right)^2 \implies \quad (3.6)$$

$$\gamma^2 p_x^2 - \gamma^2 \beta^2 p_x^2 - 2\gamma^2 p_x \beta \frac{E}{c} + 2\gamma^2 p_x \beta \frac{E}{c} + \gamma^2 \beta^2 \left(\frac{E}{c}\right)^2 - \gamma^2 \left(\frac{E}{c}\right)^2 = p_x^2 - \left(\frac{E}{c}\right)^2 \implies \quad (3.7)$$

$$\gamma^2(1 - \beta^2)p_x^2 + \gamma^2(\beta^2 - 1)\left(\frac{E}{c}\right)^2 = p_x^2 - \left(\frac{E}{c}\right)^2 \implies \quad (3.8)$$

$$\frac{1 - \beta^2}{1 - \beta^2} p_x^2 + \frac{\beta^2 - 1}{1 - \beta^2} \left(\frac{E}{c}\right)^2 = p_x^2 - \left(\frac{E}{c}\right)^2 \implies \quad (3.9)$$

$$p_x^2 - \left(\frac{E}{c}\right)^2 = p_x^2 - \left(\frac{E}{c}\right)^2 \quad (3.10)$$

Which proves that our 4-vector is invariant.

The length of this 4-vector can be related to another invariant; the (rest) mass. The next section will describe how Einstein deduced that when the momentum equals $p = \gamma mv$ the energy becomes $E = \gamma mc^2$.

$$p^2 - \left(\frac{E}{c}\right)^2 = (m\gamma v)^2 - \left(\frac{m\gamma c^2}{c}\right)^2 = m^2 c^2 \gamma^2 \left(\left(\frac{v}{c}\right)^2 - 1\right) = m^2 c^2 \frac{\beta^2 - 1}{1 - \beta^2} \quad (3.11)$$

$$m^2 c^2 = \left(\frac{E}{c}\right)^2 - p^2 \quad (3.12)$$

Using this equation we can calculate the momentum of a photon. Recall that a photon does not have a (rest) mass and that therefore $p_{\text{photon}} = \frac{E_{\text{photon}}}{c}$, or according to Planck $p_{\text{photon}} = \frac{h\nu}{c}$.

If we choose our units in such a way that all quantities are expressed in terms of light speed and seconds¹, equation 3.12 can be written as:

$$m = \sqrt{E^2 - p^2} \quad (3.13)$$

Changing units is a method often used in physics and mathematics to make problems more clear. A different method is to make quantities dimensionless or ‘without units’. The variables β and γ in this text are examples of dimensionless quantities.

¹This is known as the Natural Units (NU) system. There are actually two NU systems; Lorentz-Heaviside and Gaussian. In both the speed of light is equal to 1 because they define $\varepsilon_0 = 1$ and $\mu_0 = 1$ ($c = \frac{1}{\varepsilon_0 \mu_0}$). Looking at the equations in this text it becomes clear why it is advantageous to set c at 1, it greatly reduces the number of calculations.

4 Einstein and Minkowski

Albert Einstein studied the recoil of light in a thought experiment and with this he came to his now famous equation $E = mc^2$. He imagined a sealed box (in rest) with a certain length l and mass m_{box} . On the left side of the box we send a photon to the right, after a certain time this photon will be absorbed by the right side of the box. We already mentioned how we can describe the momentum of a photon:

$$p_{photon} = \frac{E_{photon}}{c} \quad (4.1)$$

Furthermore:

$$p_{box} = m_{box}v_{box} \quad (4.2)$$

The speed of the box can be calculated by dividing the displacement of the box by the time the photon exists:

$$p_{box} = m_{box} \frac{\Delta x_{box}}{\Delta t} \quad (4.3)$$

Of course we must not forget the law of conservation of momentum. The box was stationary before the photon was emitted, therefore:

$$m_{box} \frac{\Delta x_{box}}{\Delta t} + \frac{E_{photon}}{c} = 0 \quad (4.4)$$

We know that photons travel at the speed of light, thus the box will move $\Delta t = \frac{l}{c}$ seconds.

$$m_{box} \frac{\Delta x_{box}}{l} + \frac{E_{photon}}{c^2} = 0 \quad (4.5)$$

If we know the displacement of the box we can solve this equality. We will now introduce a new quantity; m_{photon} . The law of conservation of momentum still holds¹:

$$m_{photon}c + m_{box}v_{box} = 0 \quad (4.6)$$

$$m_{photon}c\Delta t + m_{box}v_{box}\Delta t = 0 \quad (4.7)$$

$$m_{photon}l + m_{box}\Delta x = 0 \quad (4.8)$$

Substitution in equation 4.5 yields:

$$m_{photon} = \frac{E_{photon}}{c^2} \quad (4.9)$$

¹Here something strange happens, the photon will not traverse the entire length of the box if the box is moving in the opposite direction at the same moment. The assumption that the photon does travel the entire length only holds when the box has a far greater mass than the photon, can you explain why?. And on a side note; what happens with Δt ?

Generalising the equation above results in the famous equation:

$$E = mc^2 \quad (4.10)$$

Einstein's thought experiment tells us that when we put energy into an object, its mass will increase. In the previous section we saw that the length of the 4-vector is invariant. This means that the (rest) mass m is also invariant. We now have two different masses, the rest mass and the apparent or observed mass. We will denote the observed mass as M :

$$E = Mc^2 = \alpha mc^2 \quad (4.11)$$

The observed mass M is a factor α larger than the rest mass m ¹. Calculating the momentum is now done using:

$$p = Mv = \alpha mv \quad (4.12)$$

We can substitute this in equation 3.12:

$$m^2 c^2 = \left(\frac{\alpha mc^2}{c} \right)^2 - (\alpha mv)^2 \quad (4.13)$$

$$m^2 = \alpha^2 m^2 \left(1 - \left(\frac{v}{c} \right)^2 \right) \quad (4.14)$$

$$\alpha^2 = \frac{1}{1 - \left(\frac{v}{c} \right)^2} \quad (4.15)$$

$$\alpha = \frac{1}{\sqrt{1 - \left(\frac{v}{c} \right)^2}} \quad (4.16)$$

Compare this α to the Lorentz factor γ , they are the same. The observed mass can thus also be written as $M = \gamma m$ and energy as:

$$E = \gamma mc^2 \quad (4.17)$$

What this equation tells us is that an object, even when it is not moving, has a certain amount of energy. The classical equation for kinetic energy thus denotes a change in energy; $\Delta E = \frac{1}{2}mv^2$. In table 4.1 and figure ?? the importance of this distinction at difference speeds can be seen. At low speeds both models give the same value (if we ignore the 'mass energy' term). The classical model was constructed by Galilei, Huygens, Kepler, and Newton. They did not observe phenomena happening at the speed of light. So it is no surprise that their model could be in error at higher speeds. At 0.1 c the difference between the classical model and Einstein's model is less than 1%. One tenth of the speed of light is 30000 km/s, compare this to the speed of weight falling from the Leaning Tower of Pisa in Galilei's experiments or the speed of the celestial bodies observed by Huygens and Kepler.

¹In most literature the rest mass is denoted as m_0 , the observed mass is then written as m .

v	$E_{\text{Einstein}} \text{ (J)}$	$\Delta E_{\text{classical}} \text{ (J)}$
0c	mc^2	0
0.1c	$mc^2 + 0.005mc^2$	$0.005mc^2$
0.2c	$mc^2 + 0.021mc^2$	$0.020mc^2$
0.3c	$mc^2 + 0.048mc^2$	$0.045mc^2$
0.4c	$mc^2 + 0.091mc^2$	$0.080mc^2$
0.5c	$mc^2 + 0.155mc^2$	$0.125mc^2$
0.6c	$mc^2 + 0.250mc^2$	$0.180mc^2$
0.7c	$mc^2 + 0.400mc^2$	$0.245mc^2$
0.8c	$mc^2 + 0.667mc^2$	$0.320mc^2$
0.9c	$mc^2 + 1.294mc^2$	$0.405mc^2$
1.0c	∞	$0.500mc^2$

Table 4.1: Difference between the Einstein and classical model.

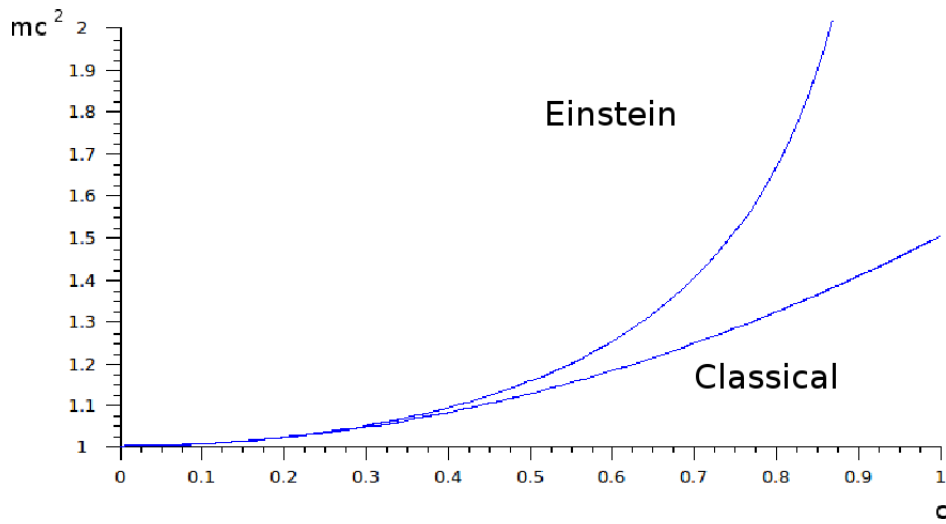


Figure 4.1: The energy as function of the speed (in terms of c) according to Einstein and the classical model.

At low speeds the classical model is a good approximation reality. However, at high speeds the error becomes unacceptably large. The speed of an object which has mass can never reach the speed of light. As you can see in table 4.1 such an object would have an infinite amount of energy. Near the speed of light most of the energy you put into an object is 'transformed' into mass. A direct consequence of this is that all particles travelling at the speed of light must have no rest mass.

With the laws for conservation of energy and momentum we should be able to do some calculations involving particles colliding at speeds near the speed of light.

While in Zurich, Albert Einstein was a student of Hermann Minkowski. During that time Minkowski was working at multidimensional mathematical problems. He took the ideas of Poincaré and expanded them into what we now call Minkowski space or Minkowski spacetime.

In 1915 Einstein expanded his special theory of relativity with the general theory of relativity. What was missing in the special theory was gravity. In the general theory gravity is explained as a curving of space(-time). Just at travelling at high speeds, gravity has an effect on time.

Exercise 1 : In the sections 'Invariants' and 'Einstein and Minkowski' we run the risks of circle reasoning. Where does this risk occur and is this a problem?