Contents

1. Theory of relativity

1.1. Introduction

The two most successful physical theories of the 19th century, (classical) mechanics and electromagnetism, stood the test of time and work almost seamlessly hand in hand. There is however one troublesome aspect that, when taken into account, has the two theories completely at odds: light.

Maxwell's Equations for electromagnetism, in accord to the experimental results, predict that light should travel with speed:

$$c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} = \frac{1}{\sqrt{(8.854 \times 10^{-12} F/m)(1.256 \times 10^{-6} N/A^2)}} =$$

$$= \frac{1}{\sqrt{11.120 \times 10^{-18}(s^{\cancel{4}} \cdot A^{\cancel{2}} \cdot \cancel{kg} \cdot \cancel{m} \cdot s^{\cancel{-2}}/\cancel{kg} \cdot m^2 \cdot \cancel{m} \cdot \cancel{A^2})}} =$$

$$= \frac{1}{\sqrt{11.120 \times 10^{-18} s^2/m^2}} = \frac{1}{3.334 \times 10^{-9} s/m} \approx 3.000 \times 10^8 m/s$$

That is, light apparently moves at speed c in any circumstance. This is problematic, because mechanics postulates that there is no such thing as an "absolute velocity": all velocities depend on the reference frame from which they are observed, as the principle of Galilean relativity states.

Consider two inertial reference frames, S and S', where S' is moving at constant speed v (along the x axis) with respect to S. Suppose that a beam of light were to be travelling at speed c (again, along the x axis) with respect to S; then, applying Galilean relativity, one would expect the light beam to travel at velocity $c - v$ with respect to S'. However, when measuring the speed of the light beam with respect to S' one still finds that it moves with speed c.
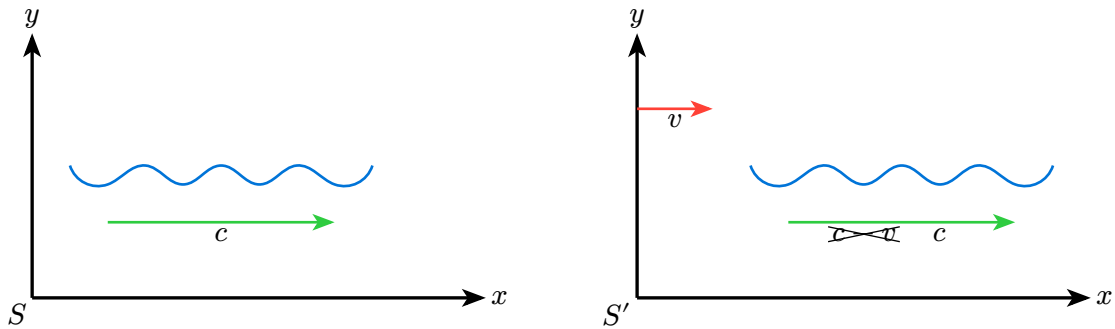

Figure 1: With respect to an inertial frame of reference S, a light beam (in blue) moves with velocity c. With respect to another inertial frame of reference S', that is moving with respect to S at constant velocity v, it would be assumed that the light beam would travel with velocity $v - c$ with respect to S'. However, this is not the case: the light beam moves at velocity c both with respect to S and with respect to S'.

Also, despite its accomplishments, Maxwell Equations left a question opened: how could light travel in the absence of a medium? Light was understood to be a wave, and each wave needed a medium through which they propagate (like air is the medium through which sound propagates, or water is the medium through which water waves propagate). However, how can light travel, for example, from the Sun to the Earth with nothing but empty space in between?

One hypothesis that was advanced early on to reconcile the two theories and, at the same time, to solve this conundrum, was to postulate that the entire Universe is permeated by a peculiar substance called **luminiferous ether**, or just **ether**. This ether would be the medium through which light propagates,

and the speed of light is constant when measured with respect to the "special" frame of reference of the ether. Stated otherwise, the value of c that arises from Maxwell equations does not refer to the speed of light "per se", but the value measured with respect to this priviledged frame of reference.

The ether hypothesis had some ground, since Maxwell's Equations invalidated Galilean coordinate transformations. That is, when observing the same experiment involving light from different frames of reference, the results would be different, even if the frames are inertial. This means that it's possible to rely on Maxwell's Equations to tell apart which system is moving with respect to the other in an "absolute" sense, and "absolute" motion is explicitly forbidden in Newtonian/Galilean mechanics.

If this hypothesis were to be true, the contradiction would be solved, even though this would elevate the ether's reference frame as a "special" reference frame, in which motion is absolute. Given an inertial frame of reference S, let v be the velocity with which S moves with respect to the frame of reference of the ether. If it were to be possible to compute v, now the speed of light measured with respect to S would now be $c + v$ (or $c - v$, depending on the direction), as expected.

The ether hypothesis was short-lived, however, since it became clear that no such substance exists. Even before experimental evidence proved the hypothesis wrong, it could hardly be possible for this ether to exist, since it would be something that is present in the entire Universe and yet being barely noticeable.

The most famous experiment that put the existence of the ether into question was the **Michelson and Morley experiment**.

A different approach was the one followed by Einstein, in its **special theory of relativity**: the name "special" comes from the fact that the theory concerns itself only with inertial frames of reference, hence "special" as in "special case". The theory is based on only two postulates:

- **Principle of relativity**. All physical laws are the same with respect to any inertial frame of reference. This is an extension of the Galilean Principle of Relativity, which states that (just) the laws of mechanics are the same with respect to any inertial frame of reference;
- **Constant speed of light**. As predicted by the Maxwell Equations, the speed of light is a universal constant, that holds the same value in any inertial frame of reference.

With these two assumptions, it is possible to construct a theory that is consistent and solves (almost) all contradictions between mechanics and electromagnetism. It introduces, however, many seemingly paradoxical consequences.

First of all, it is necessary to let go of the notion of having a "global" time that flows at the same pace for any reference frame. This is the case in Galilean transformations, where the time variable t in a given (inertial) reference frame does not change when considering the same even in a different (inertial) reference frame.

Consider an experimental setup where two sources of light, one facing the other, are placed vertically at a given distance L. The two sources are kept fixed in place, so that they either both stand still or they both have to move at the same speed. Whenever one source captures a photon emitted by the other, it sends a photon back.

If the apparatus is standing still, both from the frame of reference of an external observer and from the frame of reference of the light source, the photon has to travel along a straight line to start from one source, reach the other source and go back to the original source.

If, on the other hand, the apparatus is moving at constant velocity (for simplicity, only along the left-right axis) with respect to an external observer, the two frames of reference paint a different picture. For the frame of reference of the source, the path traced by the photon is still a vertical line, since it's moving along at constant speed. On the other hand, from the frame of reference of an external

observer the path traced by the photon is two incident lines, since both the photon *and* the apparatus are moving.
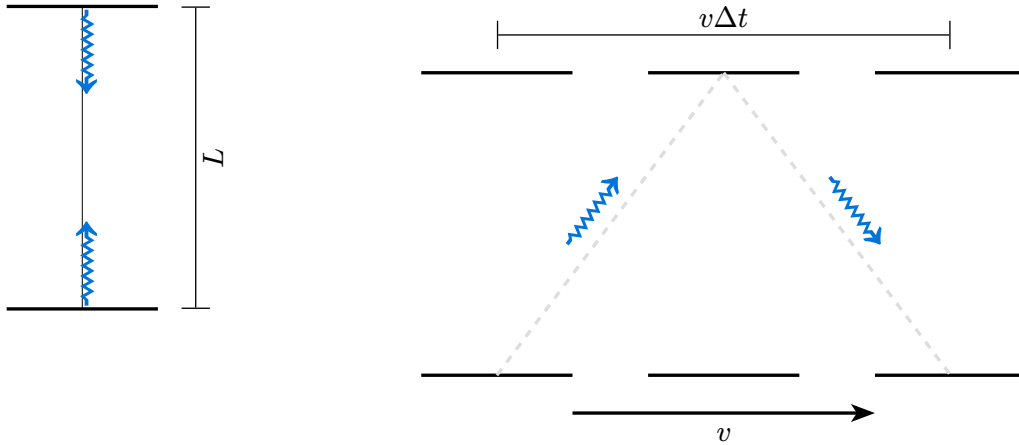

Figure 2: The apparatus consists in two facing sources of light exchanging photons. When the apparatus is still (left), the distance that a photon has to travel to go from its source to the other source and back to its original source is $2L$. When the apparatus is moving (right) the light has to travel a a distance longer than $2L$ to go back and forth, when observed from an external frame of reference.

Let v be the velocity with which the apparatus is moving with respect to an external observer. Let t_0 be the time frame in which the photon starts moving from its original source and let t_f be the time frame in which the photon has returned to its source. Let then $\Delta t = t_f - t_0$ be the time interval in which the photon moves.

In a Newtonian setting, the velocity of the photon when the apparatus is still would be c (as predicted by Maxwell's Equations), both in the frame of reference of the source and in the frame of reference of an external observer. On the other hand, when the apparatus is moving the velocity would be c in in the frame of reference of the source and $c + v$ in the frame of reference of an external observer. Also, the time interval Δt would be the same in both frames of reference and in both situations.

However, if Einstein's First Postulate is to be taken into account, the velocity for both frames of reference when the apparatus is moving has to be c. The distance travelled by the moving apparatus (with respect to an external observer) is $v\Delta t$, the velocity of the apparatus multiplied by the time the photon has taken to go back and forth once. The distance between the two light sources is still L, so the distance d travelled by the photon with respect to an external observer can be computed applying Pythagoras' Theorem:

$$\frac{d}{2} = \sqrt{\left(\frac{v\Delta t}{2}\right)^2 + L^2} \Rightarrow \frac{d^2}{4} = \frac{v^2(\Delta t)^2}{4} + L^2 \Rightarrow d^2 = v^2(\Delta t)^2 + 4L^2$$

With respect to the frame of reference of the source, the photon still travels along a straight line of length L at (constant) velocity c. However, since the velocity of the photon must be c for any observer and any inertial frame of reference, the length L must be equal to $c\Delta t'$, where $\Delta t'$ is a time interval having different size with respect to Δt:

$$d^2 = v^2(\Delta t)^2 + 4L^2 = v^2(\Delta t)^2 + c^2(\Delta t')^2$$

Where the factor of 4 was included directly into the new time $\Delta t'$, being a dummy variable.

Since the distance d must also be equal to $c\Delta t$:

$$c^2(\Delta t)^2 = v^2(\Delta t)^2 + c^2(\Delta t')^2 \Rightarrow c^2(\Delta t')^2 = (c^2 - v^2)(\Delta t)^2$$

Solving for Δt:

$$(\Delta t)^2 = \left(\frac{c^2}{c^2 - v^2}\right)(\Delta t')^2 \Rightarrow \Delta t = \Delta t'\sqrt{\frac{c^2}{c^2 - v^2}} \Rightarrow \Delta t = \Delta t'\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \Rightarrow \Delta t = \gamma\Delta t'$$

Where $\gamma = 1/\sqrt{1 - \frac{v^2}{c^2}}$ is also referred to as **Lorentz's factor**. This result does not hold just for this experimental setup, but for any pair of inertial frames of reference where one is moving at a constant speed with respect to the other.

The equation implicitly states that no velocity can be greater than the speed of light, in any inertial frame of reference. If this wasn't the case, the fraction v^2/c^2 could become greater than 1, potentially resulting in a value of γ (and of $\Delta t'$) that is a complex number. Since time intervals are necessarily real numbers, any velocity must be smaller than c, which also implies that γ is always greater or equal than 1.

This means that time effectively runs faster or slower depending on the frame of reference in which events are observed. In particular, when two inertial frames of referenece are at rest with respect to each other ($v = 0$), then $\gamma = 1$, meaning that time flows at the same pace in both reference frames. On the other hand, as v approaches c, then γ approaches $+\infty$, meaning that as velocity increases as time slows down with respect to a frame of reference of an external observer.

Since the time observed in a frame of reference at rest with respect to the events is the smallest time interval than can be observed, being $\gamma = 1$ the smallest possible value of γ, this time interval is also referred to as **proper time**. It should be noted that all measured times in all inertial frames of reference moving at any constant speed are equally valid observations of the same phenomena. The name "proper time" refers to the fact that, since the observer is itself not moving, there is no "distorsion" induced by velocity when measuring time.

> **Exercise 1.1.1**: What would be an analogy?

Solution: The Moon, when observed from the Earth, appears to be no wider than a couple of centimetres. However, the *real* width of the Moon is roughly $3.5 \times 10^6 m$, which is orders of magnitude larger. This happens because the distance from which a phenomena is observed modifies its apparent proportions: the *real* width of the Moon is the one observed when being close to it, because it's the least "distorting" point of view. This does not mean that the width of the Moon observed by the Earth is *wrong*, just that its point of view is more biased. In the same fashion, proper time describes a phenomena with the least bias, and hence it is understood to be the *real* temporal description. □

The phenomenon of time moving at slower pace in external frames of reference is called **time dilation**. Time dilation is an observable property of time and space, and it has real physical consequences. Any real-world process that can be observed, from supernovas to aging, is subjected to time dilation: a person ages faster or slower based on which frame of reference they are observed, for example.

Exercise 1.1.2: Suppose that an observer standing still is watching an alien spaceship passing by left-to-right at 85% the speed of light. If the time interval from when the spaceship is first sighted by the observer to when it is right in front of them is, from its reference frame, 2 minutes, how much time is passed in the same time frame in the frame of reference of the spaceship?

Solution: If $v = 85\%c$, then v^2/c^2 is $(0.85)^2 \approx 0.73$, which gives a Lorentz's Factor of $1/\sqrt{1-0.73} \approx 2$. If the time interval in the reference frame of the observer is $2min$, the time interval in the reference frame of the spaceship (the proper time) is $2min/2 = 1min$. $\qquad\square$

It should be noted, however, that γ scales incredibly slowly with velocity, and only velocities that are very close to the speed of light can yield a value of γ large enough to measure significative time offsets.
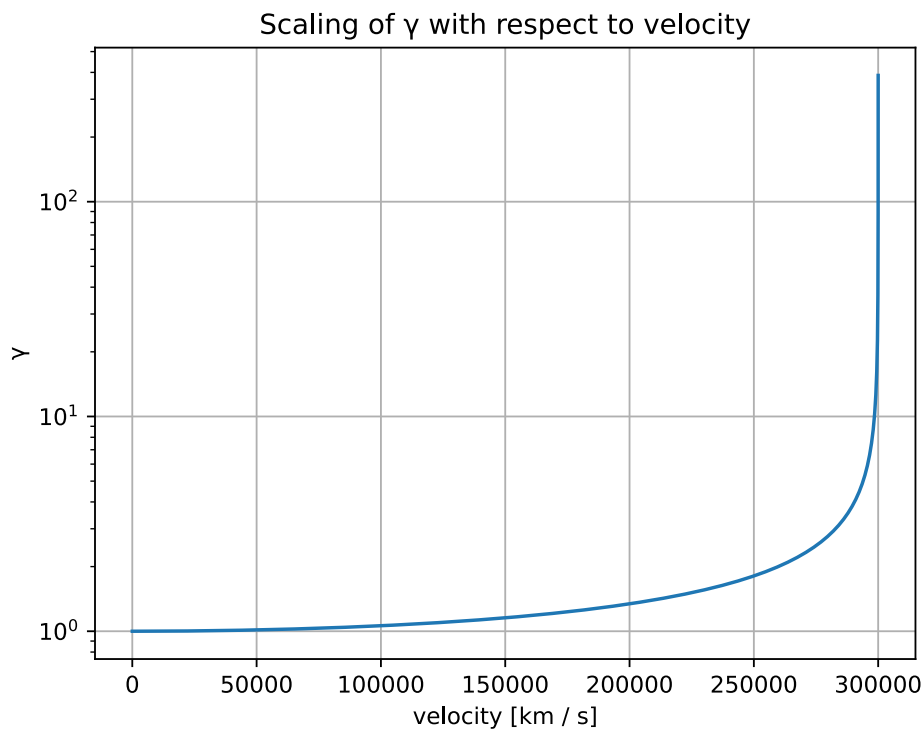

Figure 3: Value of γ with respect to different values of velocity. As can be appreciated, γ becomes noticeable only at incredibly high values of c (γ is displayed in logarithmic scale).

Also, since most real-world phenomena have a velocity that is no where near the speed of light, taking into account the γ factor is not necessary, because $\gamma \approx 1$ and hence $\Delta t' \approx \Delta t$. This is why Newtonian mechanics is still a reasonable model for interpreting reality when working at slow velocities.

Exercise 1.1.3: Fastest human-made space probes have a velocity of roughly $200km/s$. If one such probe were to be observed from an inertial frame of reference, what would be the Lorentz's Factor?

Solution: The speed of the probe is roughly 0.0006 times the speed of light. Hence:

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{1}{\sqrt{1 - \left(\frac{6 \times 10^{-4} \cancel{c}}{\cancel{c}}\right)^2}} = \frac{1}{\sqrt{1 - 3.6 \times 10^{-8}}} \approx 1.00000018$$

$\square$

The fact that the flow of time varies with different frames of reference also implies that lenghts vary with different frames of reference. That is, the same object can be observed to be longer or shorter depending on which frame of reference it is observed.

Consider, again, two inertial frames of reference, where one is moving with respect to the other in the left-right direction at constant speed v. Suppose that the first measures a certain time interval Δt when moving from point A to point B, travelling a certain distance L. Being a constant uniform motion, $L = v\Delta t'$.

The second frame of reference will instead measure a different time interval Δt when moving from B to A. If the travelled distance were to be the same as the distance observed in the other frame of reference, one would have $L = v\Delta t$, since the velocity is the same (in modulo) for both inertial frames of reference. This would mean $v\Delta t = v\Delta t'$, which entails $\Delta t = \Delta t'$. This is not possible, however, since by construction the two frames of reference are moving with respect to each other.

It is therefore necessary to assume that the second frame of reference actually observed a distance of $L' = v\Delta t$, with L' being different from L. Recalling that $\Delta t' = \gamma \Delta t$:

$$L' = v\Delta t = v\frac{\Delta t'}{\gamma} = \frac{L}{\gamma}$$

This is also referred to as **length contraction**, or **Lorentz's contraction**. Analogously to how proper time was defined, **proper length** is a length measured from the frame of reference of an external observer, which is also the longest possible.

Note that, this way, special theory of relativity presents a remarkable symmetric: a frame of reference at rest with a moving body will measure proper time but not proper distance, whereas a frame of reference external to the moving body will measure proper distance but not proper time. In particular, in the first frame of reference both time and space are shorter, in the second both time and space al longer.

It should be noted that, even in Special Relativity, some physical quantities are invariant among different frames of reference. First, note that length contraction happens only along the direction of motion, not along all direction. Consider two apparatuses such as the two mirror example, moving with respect to each other and with respect to a third, external frame of reference at velocities v_1 and v_2.

Assume that both apparatuses start in the same condition and in the same spacial coordinates. An observer in the external frame of reference would observe the photon in the first apparatus travel $\frac{1}{2}\Delta x_1 m$ in the x direction and $\frac{1}{2}c\Delta t_1 m$ in the y direction, while observing the photon in the second apparatus travel $\frac{1}{2}\Delta x_2 m$ in the x direction and $\frac{1}{2}c\Delta t_2 m$ in the y direction. In both cases, the height L is the same; applying Pythagoras' Theorem:

$$L^2 = \left(\frac{1}{2}c\Delta t_1\right)^2 - \left(\frac{1}{2}\Delta x_1\right)^2 = \left(\frac{1}{2}c\Delta t_2\right)^2 - \left(\frac{1}{2}\Delta x_2\right)^2$$

Cancelling the $\frac{1}{2}$ factor:

$$c^2\Delta t_1^2 - \Delta x_1^2 = c^2\Delta t_2^2 - \Delta x_2^2$$

The quantity $c^2\Delta t^2 - \Delta x^2$, also called the **spacetime interval**, is the same in any frame of reference.

Special relativity introduces another seemingly counterintuitive result: two events can or cannot appear to be simultaneous depending on the frame of reference in which they are observed. This is because the act of *observing* is itself dependent on the speed of light: be it a measuring instrument or the human eye, *observing* a phenomena entails capturing a light beam carrying the information associated to said phenomena, moving from the phenomena to the observer. Since this movement happens at the speed of light (that is, is not istantaneous), there is necessarily a delay. Also, since the speed of light is always the same in any inertial frame of reference, changing frame of reference also potentially changes the distance that the light has to travel and therefore the perceived time frame at which an event is observed.

Consider two inertial reference frames, one moving with respect to the other at velocity v. Suppose that one of the two frames is centered in the middle of two events, that happen simultaneously. This means that the light wave carrying the information associated to the two events reaches the observer in said frame of reference (moving at speed c) reach the observer in the same time frame. The other reference frame, having to take into account the movement of the frame of reference, will not perceive both events happening simultaneously: one will be perceived before the other.

Figure 4: In both situations, both events (a camera flash) are synchronized so that they happen simultaneously. In the first situation, an external observer lies exactly in the middle of the two, and therefore will always observe the two events simultaneously. In the second situation, the observer is closer to the of the two, and therefore will observe one (the closest) before the other.

Exercise 1.1.4: Muons are subatomic particles created in the upper atmosphere when cosmic rays collide with air molecules. Muons are unstable particles that have an half-life of $1.5\mu s$, meaning that, given a sample of muons, each $1.5\mu s$ roughly half the sample size decays into other smaller particles. Their speed is very close to the speed of light, and their presence on Earth's surface can be easily detected with ad-hoc particle detectors.

The upper atmosphere is about $60km$ above sea level, and the speed of a muon is about $0.9997c$. The time that a muon takes to travel from the atmosphere to the Earth's surface moving at constant speed straight in a downward direction (without decaying) is:

$$\Delta t \approx \frac{h}{c} = \frac{6 \times 10^4 m}{3 \times 10^8 m/s} = 2 \times 10^{-4} s = 200\mu s$$

Since $200\mu s/1.5\mu s$ is 133, it means that a muon that were to take such path would have to "dodge" decaying 133 times in a row. Since the probability of decaying (and of not-decaying) is always 0.5, the probability for a muon to reach the Earth without having decayed would be $(0.5)^{133} \approx 10^{-40}$.

This would mean that, every 10^{40} muons, only a single muon would reach the Earth's surface. However, experimental data contradicts such observation, since around 10% of the muons created in the upper atmosphere are detected on Earth. How can this be?

Solution: This apparent paradox can be solved by taking Special Relativity into account. From the muon's frame of reference, it is the Earth that is moving towards it at speed $0.9997c$: if an observer in the frame of reference of the Earth's surface measures a (dilated) time of $200\mu s$, the muon observes a (proper time) of:

$$\Delta t = \frac{\Delta t'}{\gamma} = 200\mu s\sqrt{1 - \frac{v^2}{c^2}} = 200\mu s\sqrt{1 - (0.9997)^2} = (0.0245 \cdot 200)\mu s = 4.9\mu s$$

Which gives an half-life of $4.9\mu s/1.5\mu s \approx 3.3$. This means that the "true" probability of a muon to reach the Earth's surface is $(0.5)^{3.3} \approx 0.1$, as confirmed by empirical evidence. $\square$

Galilean transformations allow one to convert a set of coordinates and velocities that are relative to a given frame of reference S to a set of coordinates relative to a different frame of reference S'. There is clearly interest in extending Galilean transformations to account for the phenomena of Special Relativity: such an extension should abide to three constraints:

1. Fall back to Galilean transformations at low speeds;
2. Take into account both transformations in space and transformations in time;
3. Ensure that the speed of light remains the same in all reference frames.

An "educated guess" would be an expression of the form:

$$x' = A(x - vt) \qquad\qquad x = A(x' + vt')$$

Where A is a to-be-determined constant that goes to 1 as the speed v goes to 0. To determine it, assume that both reference frames S and S' start in the same position $x = x' = 0$ and start in the same time frame $t = t' = 0$.

An event happens at coordinates (x, t) with respect to S and at coordinates (x', t') with respect to S'. The information carried by such events travels at the speed of light for a distance of $x = ct$ in S and $x' = ct'$ in S'. Substituting in the previous expressions:

$$ct' = A(ct - vt) = At(c - v) \qquad\qquad ct = A(ct' + vt') = At'(c + v)$$

The first expression gives $t' = At\left(1 - \frac{v}{c}\right)$. Substituting into the second:

$$ct = At'(c + v) = A\left(At\left(1 - \frac{v}{c}\right)\right)(c + v) = A^2t\left(c - \cancel{v} + \cancel{v} - \frac{v^2}{c}\right) = A^2t\left(c - \frac{v^2}{c}\right)$$

Solving for A:

$$c\cancel{t} = A^2\cancel{t}\left(c - \frac{v^2}{c}\right) \Rightarrow c^2 = A^2(c^2 - v^2) \Rightarrow A = \sqrt{\frac{c^2}{c^2 - v^2}} = \sqrt{\frac{1}{1 - \frac{v^2}{c^2}}}$$

Which is exactly the value y. An equation for time can be obtained in a similar fashion. In total, one obtains two systems of four equations each, called **Lorentz transformations**, that act as the relativistic counterpart to the Galilean transformations:

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

When the velocity v is much smaller than c, γ is close to 1 and the ratio vx/c^2 is close to 0, hence "falling back" to the Galilean transformations.

The last step in deriving the Lorentz transformations is an expression for velocities. The definition of velocity in the Galilean laws of motion is the derivative of space over the derivative of time. Carrying it out:

$$u' = \frac{dx'}{dt'} = \frac{d(\gamma(x - vt))}{d(\gamma(t - vx/c^2))} = \frac{\cancel{\gamma}d(x - vt)}{\cancel{\gamma}d(t - vx/c^2)} = \frac{dx - vdt}{dt - (v/c^2)dx} = \frac{\frac{dx}{dt} - v}{1 - (v/c^2)\frac{dx}{dt}}$$

But dx/dt is just u, giving:

$$u' = \frac{u - v}{1 - \frac{v}{c^2}u} \qquad\qquad u = \frac{u' + v}{1 + \frac{v}{c^2}u'}$$

When v is noticeably smaller than c, the ratio v/c^2 is infinitesimal, hence giving the same results as the Galilean transformations $u' = u - v$ and $u = u' + v$. Also, when $v = c$:

$$u' = \frac{u - c}{1 - \frac{\cancel{c}}{c^{\cancel{2}}}u} = \frac{u - c}{1 - \frac{u}{c}} = c\frac{\cancel{u - c}}{\cancel{c - u}} = -c \qquad\qquad u = \frac{u' + c}{1 + \frac{\cancel{c}}{c^{\cancel{2}}}u'} = \frac{u' + c}{1 + \frac{u'}{c}} = c\frac{\cancel{u' + c}}{\cancel{c + u'}} = c$$

Hence preserving the constancy of the speed of light in all reference frames.

Having introduced the relativistic γ factor also requires one to properly adapt not just Newtonian laws of motion, but also the definition of other physical quantities that depend on velocity, such as kinetic energy and momentum.

The classical definition of momentum is of mass times velocity; momentum is an important quantity in classical physics because it is always conserved. However, it can be shown that blindly applying such definition to particles where the Lorentz transformations are applied does not work, meaning that different frames of reference give different value of the momentum. So either momentum conservation should be abandoned, or the definition of momentum ought to be revisited.

Out of the two, the second is much more plausible. First, the definition of momentum of a particle is, as stated, mass times velocity, or mass times space over time. However, different frames of reference are associated to different time intervals and different lengths: the frame of reference of choice should be the one centered in the particle, since it's the one associated to proper time.

Let $u = \Delta x/\Delta t$ be the velocity of a particle in a frame of reference external to said particle. Let $\Delta \tau$ be the time measured in the frame of reference of the particle to describe the interlude of the same events described by Δt. Define the momentum of a particle with respect to its own frame of reference to be $p = mu$. Given that $\Delta\tau = \Delta t/\gamma$:

$$p = mu = m\frac{\Delta x}{\Delta\tau} = m\frac{\Delta x}{\Delta t/\gamma} = \gamma m\frac{\Delta x}{\Delta t} = \gamma mu$$

This is the expression for the relativistic momentum of a particle when observed to have velocity u from a given frame of reference. It can be proven that, as long as Lorentz transformations are used to transition from a frame of reference to another, the law of conservation of momentum still holds.

It is however important to note that the γ factor that appears in the expression for the relativistic momentum is somewhat different from the gamma factor that appears in the Lorentz transformations. The latter influences the velocity of a frame of reference with respect to another, the former influences the velocity of the particle *itself* with respect to the frame of reference in which it is observed. For this reason, to avoid distinction, the formula is sometimes written as $p = \gamma_p mu$, where the pedix p stands for "particle".

Velocity also appears in the expression for Newtonian kinetic energy, $K = \frac{1}{2}mu^2$. Since $p = mu$, the expression can also be written as $K = p^2/2m$. Having extended momentum to the relativistic case and its law of conservation, it is necessary to do the same with energy.

Consider a particle travelling a distance Δx in a time Δt, as observed from a given frame of reference S. As stated earlier, the spacetime interval $c^2\Delta^2 t - \Delta^2 x$ is invariant with respect to the frames of reference. Let $\Delta\tau$ be the time interval describing the same event when measured from the frame of reference of the particle itself.

Multiplying the expression for the spacetime interval by $(m/\Delta\tau)^2$ gives:

$$(c^2\Delta^2 t - \Delta^2 x)\left(\frac{m}{\Delta\tau}\right)^2 = \frac{m^2}{\Delta^2\tau}c^2\Delta^2 t - \frac{m^2}{\Delta^2\tau}\Delta^2 x = m^2 c^2 \frac{\Delta^2 t}{\Delta^2\tau} - m^2\frac{\Delta^2 x}{\Delta^2\tau} = (mc)^2\left(\frac{\Delta t}{\Delta\tau}\right)^2 - p^2$$

Now the expression refers to the relativistic momentum of the same particle when observed from two different reference frames. This quantity is also invariant, since both m and $\Delta\tau$ are constants.

Substituting $\Delta t = \gamma_p \Delta\tau$ and multiplying by c^2 gives:

$$c^2\left((mc)^2\left(\frac{\Delta t}{\Delta\tau}\right)^2 - p^2\right) = c^2(mc)^2\left(\frac{\gamma_p \Delta\tau}{\Delta\tau}\right)^2 - c^2 p^2 = c^2\left(\gamma_p mc\right)^2 - c^2 p^2 = \left(\gamma_p mc^2\right)^2 - (pc)^2$$

Since this quantity is invariant with respect to any frame of reference, computing it with respect to one frame of reference gives the value for any frame of reference. The most comfortable frame of reference to choose is the one of the particle itself, where $\gamma_p = 1$ and $p = 0$:

$$\left(1 \cdot mc^2\right)^2 - (0 \cdot c)^2 = \left(mc^2\right)^2 - 0^2 = \left(mc^2\right)^2$$

Giving:

$$\left(\gamma_p mc^2\right)^2 - (pc)^2 = \left(mc^2\right)^2$$

The expression involves three terms: $\gamma_p mc^2$, pc and mc^2 (all three squared). The second term is the (non relativistic) momentum of a particle multiplied by the speed of light, whereas the third term is its mass multiplied by the speed of light. Both of these are constants, since mass cannot change and momentum p is the one associated to a specific frame of reference, the one of the particle itself.

The first term is much more interesting, however. Since γ_p has no dimension, it has the same unit of measurement as mc^2: mass times velocity squared is energy. It is also not constant, since γ_p depends on the relative velocity of the particle. Applying binomial expansion:

$$\gamma_p mc^2 = \frac{mc^2}{\sqrt{1 - \frac{u^2}{c^2}}} \approx \left(1 + \frac{1}{2}\frac{u^2}{c^2}\right)mc^2 = mc^2 + \frac{1}{2}mu^2$$

The second term is clearly the (non relativistic) kinetic energy of the particle, whereas the first is a form of energy that does not depend on any property of the particle (its velocity, its position, ecc...) except for its mass. That is, simply by "existing" (almost all particles have a mass) a body possesses a form of energy.

This means that $\gamma_p mc^2$ represents the total energy of the particle. The total energy is given by a relativistic kinetic energy and a **rest energy**:

$$E = mc^2 + \left(\gamma_p - 1\right)mc^2 = E_0 + K$$

When p or u are 0, meaning that the particle is at rest, one has $E = mc^2$, which is a known result: the **mass-energy equivalence**.

2. Quantum mechanics

2.1. Introduction

Quantum mechanics was a new paradigm developed in stages to answer questions that classical physics was unable to answer. In particular, it is a framework that is necessary to model reality at very small scales (atoms and molecules).

The first staple point of quantum mechanics is the idea that energy is not a *continuous* quantity, but is instead a *discrete* quantity, that is, an integer multiple of a fixed elementary value.

The fundamental physical constant that regulates the size of this fundamental energy bit is called the **Planck constant**, denoted as h:

$$h = 6.626 \times 10^{-34} J \cdot s$$

Since this value is very small, on large scales energy appears continuous, because the "steps" between different energy values are infinitesimal and become relevant only on small scales.

The second staple point of quantum mechanics is that the difference between particles and waves becomes blurred. In this sense, it is possible for matter particles to exhibit wave-like properties and it is possible for waves to exhibit particle-like properties.

2.1.1. Black body radiation

The first problem that quantum mechanics aided in solving is the description of the emitted radiation of a **black body**. A black body is an idealized physical body that is capable of absorbing any electromagnetic radiation, regardless of its frequency or angle of incidence, and that therefore emits back energy only and exclusively because of this absorbtion.

The spectrum of all the frequencies of electromagnetic radiations emitted from a black body, also referred to as **emission spectrum**, is given by:

$$I(f) = \frac{d^4 E}{d\theta dA \cos(\theta) dt df} = \frac{d^4 \Phi}{d\theta dA \cos(\theta) df}$$

Where E is the energy, A is the surface area, θ is the angle of emission, t is the timem f is the frequency and Φ is the flux. Since A and θ are effectively chosen by the experimenter, and are therefore always known, the only variable at play is the frequency (or the wavelength, which is just its reciprocal).

Experimentally, the emission spectrum forms a curved shape with a peak at a certain wavelenght $\lambda_{\max}$. It is possible to explicitly define the relationship between these two quantities as follows:

$$I(f, T) = \frac{2\pi h f^3}{c^2} \frac{1}{\exp\left(\frac{hf}{k_B T}\right) - 1}$$

Where T is the temperature, h is the Planck constant and k_B is the **Boltzmann constant**:

$$k_B = 1.38 \times 10^{-23} \frac{J}{K}$$

The total energy emitted by the black body is given by:

$$E_{\text{tot}} = \sigma T^4$$

Where σ is the **Stefan-Boltzmann constant**:

$$\sigma 5.67 \times 10^{-8} \frac{W}{m^2 \mathrm{K}^4}$$

The formula suggests an intuitive result, mainly that the amount of energy emitted increases as the temperature increases, as vice versa.

It can be shown that the value of λ_{max} is inversely proportional to the temperature.

It is to be expected that the aforementioned equation for $I(f, T)$ were to be derived from Maxwell equations. Interestingly, there is no way to do so. To derive the expression for $I(f, T)$ it is necessary to assume that each particle acts as an harmonic oscillator that emits energy is chunks, not as a continuous stream.

Given a certain frequency, a single chunk of energy δE is given by:

$$\delta E = hf$$

2.1.2. Photoelectric effect

The second phenomena that quantum mechanics aided in explaining was the **photoelectric effect**. This phenomena is the expulsion of electrons, called *photoelectrons*, from a piece of metal when hit by light. This happens because the energy given to the plate by the light is sufficient to break the bond that links electrons to their nucleus, and are thus ejected.

It is possible to experimentally determine the number of electrons that are ejected from the plate and their energy with respect to the light frequency and intensity, and then derive from Maxwell equations the corresponding equations and see if the results match.

In the framework of classical mechanics, the energy transfer from light to electrons is no different than heating an object. This process requires time, so it is expected that when the plate is illuminated there should be some delay before the plate starts emitting electrons. Whan happens instead is that the electrons are emitted immediately, as soon as the plate is hit.

It should also be reasonable that an higher light intensity would correspond to an higher energy of the photoelectrons. Instead, the intensity of the light has no influence on the energy of the photoelectrons, which is instead proportional to the frequency of the light.

The third puzzling observation is that the photoelectric effect only happens when the incoming light has a frequency equal or above a certain threshold, specific for each metal. In classical mechanics this has no explanation, since the photoelectric effect should happen, albeit with different degree of intensity, when employing light of any frequency.

The quantum explanation is instead to assume that light is composed of elementary massless particles called **photons**, whose energy is given by the black body harmonic oscillator model. If this is the case, all three issues are solved, because:

1. If each electron is hit by photons one by one, there's no need to wait for the body to absorb energy, since energy absorbtion is "one-shot". Therefore, the electron explusion is instantaneous;
2. Since $E = hf$, energy is indeed dependent on the frequency;
3. If the absorbed energy is insufficient, the electron is immediately recaptured by the charge of the nucleus. Being the energy dependent on the frequency, this explains the existence of a frequency threshold.

2.1.3. Gasses and radiation

When a gas is traversed by light, it is expected that the overall resulting frequency of the light is lowered, but preserved. In a similar fashion, when light is induced to emit light (by a sparkle), it is

expected that every frequency is emitted. What is observed instead is that, for each gas, only specific frequencies are emitted/preserved.

The nuclear model of the atom was tested in an experiment by Rutherford. Of course, it is not possible to study atoms simply with a microscope, because the scale is too small even for the greatest magnifier. An alternative approach is to employ electric charges, since both electrons and protons are charged.

An extremely thin plate of gold is used as a probe. Gold is used because it is both very dens and very soft, and is therefore possible to craft extremely thin plates. A radioactive source emits alpha particles (which are just helium nuclei). A detector entirely surrounds the plate, so that it is possible to observe where (and if) alpha particles are deflected. What happens is that some particles, albeit in small number, are indeed deflected, sometimes with great angle, whereas in a non-nuclei model of the atom all particles would have stroke through.

The proper model of the atom was worked out by Rutherford, who imagined the atom as an incredibly dense nucleus of positive charges with electrons orbiting around it, so that the electromagnetic force between the two acts as a centripetal force. The deflection of the alpha particles are then caused by the electromagnetic repulsion of the nuclei when alpha particles get too close to them.

The problem with the model is that the electron orbiting around the nucleus, in order to be able to maintain its orbit, would emit radiation, which in turn means it would gradually loose angular momentum until it would spiral onto the nucleus. Even assuming this to be true, the time for this to happen would be too narrow for matter to exist.

A solution was to assume that the electron does indeed orbit around the nucleus, but can only find itself at very specific distances from it. For this to be possible, it is necessary to assume that the angular momentum of the electron is quantized, according to the formula:

$$|p| = mvr = n\frac{h}{2\pi} \text{ with } n \in \mathbb{N}$$

In particular, the energy of the electron can be found by employing Coulomb's Law and the expression for the electric potential:

$$E = K + U = -\frac{1}{4\pi\varepsilon_0}\frac{e^2}{2r}$$

Where e is the electric charge of the electron and r is its distance from the nucleus.

Being the energy quantized, the only non constant member of the equation, r, must also be quantized. This means that r cannot be lower than a certain threshold, which is determined as $0.53 \times 10^{-10}m$.

Each possible value of energy that an electron can possess is called an **energy level**. Each time an electron exchanges energy with the environment, it hops from one energy level to another. In particular, by releasing energy it goes down one level, by absorbing energy it goes up one level.

Let E_0 be the lowest possible energy level (the one associated to ($r = 0.53 \times 10^{-10}m$). It is possible to relate the energy of a generic level n with respect to E_0:

$$E_n = \frac{ke^2}{2n^2 E_0}$$

Since $E = hf$, this explains why gasses can absorb/release only certain frequencies, because they are the ones that match the (fixed) energy amounts needed for electrons to move.