

SPECIAL RELATIVITY

A Quick Guide

Huan Bui

Colby College
Department of Physics and Astronomy
Class of 2021

November 12, 2018

Contents

1	Galilean and Newtonian Relativity	4
1.1	Two central problems of relativity	4
1.2	Terms defined	4
1.2.1	Events	4
1.2.2	Observers	4
1.2.3	Inertial reference frame	4
1.3	Principle of Relativity	4
1.4	Coordinate Transformations	5
1.4.1	Spatial Rotation	5
1.4.2	Galilean Transformation	6
1.4.3	Newton's Laws and Galilean Relativity	9
2	Electromagnetism and Galilean Relativity	10
2.1	Maxwell's equations and light	10
2.2	The "luminiferous ether"	10
2.3	The Michelson-Morley experiment	10
2.3.1	The "horizontal"	10
2.3.2	The "vertical"	11
2.3.3	Michelson's first brilliant observation	12
2.3.4	Michelson's second brilliant observation	12
2.4	A different experiment: Bradley's Stellar Aberration (1725) . . .	14
3	Einstein's Postulates of Relativity	15
3.1	Einstein's realization about electricity and magnetism	15
3.2	Postulates of Special Relativity	16
3.2.1	Einstein's postulates	16
4	Fundamental consequences of the postulates	16
4.1	Relativity of Simultaneity	16
4.1.1	Simultaneity and clock readings	20
4.1.2	Simultaneity and Magnetism: A practical example	21
4.1.3	Simultaneity and relativity: A cautionary tale	22
4.2	Relativity and how to understand space and time	24
4.3	Transverse length measurements	25
4.4	Time dilation	26
4.5	Length contraction	28
4.6	Reflection on fundamental effects	29
4.6.1	Time dilation with regular clocks	30
4.6.2	Length contraction revisited	32
4.6.3	Velocity "addition"	32
5	Lorentz-Einstein Transformation Equations	33
5.1	Linear transformation of coordinates	34
5.2	Derivations	34
5.2.1	Transverse dimension	34

5.2.2	$x - t$ transformation	35
5.2.3	Dependence of x', t' on x, t	35
5.3	The Lorentz-Einstein Equations	37
5.4	Some properties of the Lorentz-Einstein Transformations	37
5.4.1	Spatial separation and time interval between events	37
5.4.2	Non-relativistic limit	38
5.4.3	The limiting speed	38
5.4.4	The limiting speed - REALLY!	38
5.5	The fundamental effects and consequence of the Lorentz-Einstein transformation	39
5.5.1	Time dilation	39
5.5.2	Length contraction	40
5.5.3	Relativity of simultaneity	41
6	Relativistic Kinematics: Motion of Particles	41
6.1	Longitudinal velocity transformation	42
6.2	Transverse velocity transformation	42
6.3	Relativistic speed transformation	42
6.4	Relativistic acceleration transformation	43
6.5	The Relativistic Doppler effect	43
6.5.1	The Doppler effect for sound	44
6.5.2	Light propagation	44
6.5.3	Doppler effect for light	45
7	Spacetime	46
7.1	Minkowski spacetime diagrams and the Lorentz-Einstein transformations	46
7.1.1	Spacetime axes, Events, and World-lines	46
7.1.2	Light-cones	48
7.1.3	The S' axes	49
7.1.4	Calibrating Scales	50
7.1.5	Causality and Minkowski diagrams	51
7.2	How does Relativity explain Stellar Aberration?	53
7.3	The Space-Time interval: An <i>Invariant</i>	53
7.4	The Twin Paradox	54
7.4.1	Einstein's clock paradox	54
7.4.2	Elapsed proper time	54
8	Four-vectors	54
8.1	Three-vectors	54
8.1.1	The power of vector notations	54
8.1.2	The prototype vector	54
8.1.3	What is a 3-vector?	54
8.1.4	What is a scalar?	54
8.1.5	Combination of vectors and scalar?	54
8.2	Four-vectors	54

8.2.1	The power of 4-vector notation	54
8.2.2	The prototype 4-vector	54
8.2.3	What is a 4-vector?	54
8.2.4	Combination of 4-vectors and 4-scalars	54
8.3	Four-velocity	54
8.3.1	Four-velocity transformation	54
8.3.2	The meaning of \underline{u}	54
8.4	Four-acceleration	54
9	Relativistic Dynamics	54
9.1	Classical mechanics	54
9.2	Four-momentum	55
9.3	Interpretation of four-momentum	55
9.3.1	Non-relativistic reduction	55
9.3.2	Interpretation of P_0 and \vec{P}	55
9.3.3	The equivalence of mass and energy	55
9.3.4	Relativistic kinetic energy	55
9.3.5	The energy-momentum invariant	55
9.3.6	The energy-momentum transformation	55
9.4	Testing four-momentum conservation	55
9.4.1	Relativistic billiard problem	55
9.4.2	Collisions and Particle Creation	56
9.5	Forces and relativistic dynamics	56
9.5.1	Acceleration, velocity, mass	56
9.5.2	The four-force	56
10	The General Theory of Relativity - A brief introduction	56
10.1	Why Einstein included gravity in theory of General Relativity . .	56
10.1.1	Newton's gravitation	56
10.1.2	Problems with Newton's gravitation	56
10.2	The Equivalence Principle	56
10.2.1	Freely falling light	56
10.2.2	Gravitational time dilation: Gravitational Redshift	56

1 Galilean and Newtonian Relativity

1.1 Two central problems of relativity

Two observers in different inertial reference frames S and S' make detailed observations as physical phenomena occur. Each observer records the Time and Location of a set of events. How do their records compare?

1.2 Terms defined

1.2.1 Events

Something that occurs independent of our description of it (time and location).

1.2.2 Observers

Someone who compiles measurements of where and when Events occur. Note that measurements of when and where events occur are NOT when and where the observer “sees” the events.

1.2.3 Inertial reference frame

Coordinate systems in which the law of inertia is true.

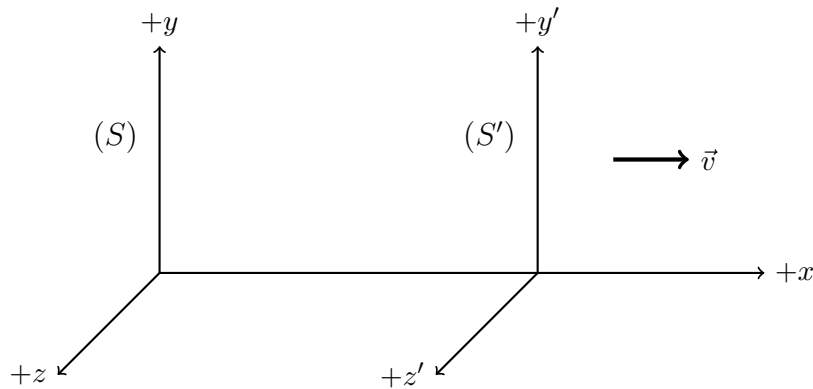


Figure 1: Frame (S') moves with speed \vec{v} relative to frame (S)

1.3 Principle of Relativity

The laws of physics are the same in all inertial reference frames.

1.4 Coordinate Transformations

If an “event” is observed in space-time ordinates (t, x, y, z) in S , how do we determine space-time coordinate (t', x', y', z') in S' ?

1.4.1 Spatial Rotation

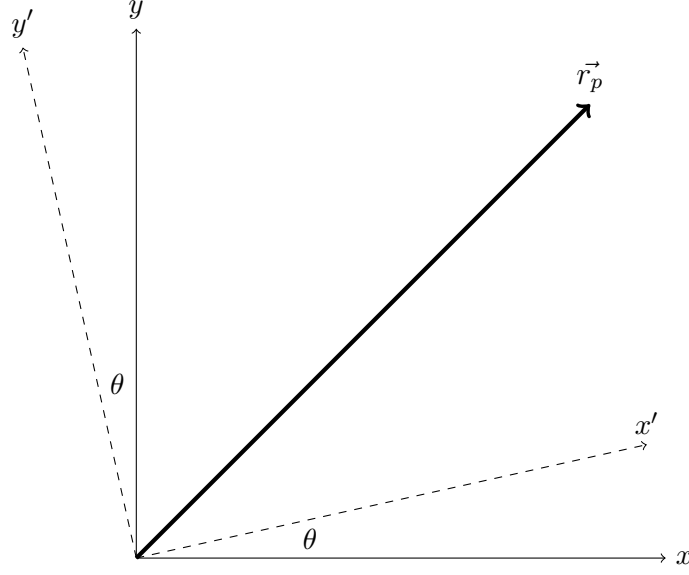


Figure 2: Spatial rotation: (S) being transformed into (S')

$$\begin{cases} (S) : \vec{r}_p = \langle x_p, y_p \rangle = x_p \hat{i} + y_p \hat{j} \\ (S') : \vec{r}_p = \langle x'_p, y'_p \rangle = x'_p \hat{i}' + y'_p \hat{j}' \end{cases} \quad (1.1)$$

For vector \vec{r}_p in our case:

$$\vec{r}_p = x_p \hat{i} + y_p \hat{j} = x'_p \hat{i}' + y'_p \hat{j}' \quad (1.2)$$

So, for any vector \vec{a} :

$$\vec{a} = x_a \hat{i} + y_a \hat{j} = x'_a \hat{i}' + y'_a \hat{j}' \quad (1.3)$$

Unit vectors:

$$\begin{cases} \hat{i}' = \hat{i} \cos \theta + \hat{j} \sin \theta \\ \hat{j}' = -\hat{i} \sin \theta + \hat{j} \cos \theta \end{cases} \quad (1.4)$$

Reversing the transformation:

$$\begin{cases} \hat{i} = \hat{i}' \cos(-\theta) + \hat{j}' \sin(-\theta) \\ \hat{j} = -\hat{i}' \sin(-\theta) + \hat{j}' \cos(-\theta) \end{cases} \quad (1.5)$$

We can thus find the coefficients x_a and y_a by combining (1.3), (1.4), and (1.5):

$$\begin{aligned}
\vec{r}_a &= x_a \hat{i} + y_a \hat{j} \\
&= x_a(\hat{i}' \cos(-\theta) + \hat{j}' \sin(-\theta)) + y_a(-\hat{i}' \sin(-\theta) + \hat{j}' \cos(-\theta)) \\
&= x_a(\hat{i}' \cos \theta - \hat{j}' \sin \theta) + y_a(\hat{i}' \sin \theta + \hat{j}' \cos \theta) \\
&= (x_a \cos \theta + y_a \sin \theta) \hat{i}' + (-x_a \sin \theta + y_a \cos \theta) \hat{j}' \\
&= x'_a \hat{i}' + y'_a \hat{j}'
\end{aligned} \tag{1.6}$$

Hence, we get:

$$\begin{cases} x'_a = x_a \cos \theta + y_a \sin \theta \\ y'_a = -x_a \sin \theta + y_a \cos \theta \end{cases} \tag{1.7}$$

We can now calculate the magnitude of vector \vec{a} in both frames (S) and (S'):

$$\begin{aligned}
(S) : ||\vec{a}|| &= x_a^2 + y_a^2 \\
(S') : ||\vec{a}|| &= x_a'^2 + y_a'^2 \\
&= (x_a \cos \theta + y_a \sin \theta)^2 + (-x_a \sin \theta + y_a \cos \theta)^2 \\
&= \dots \\
&= ((\cos \theta)^2 + (\sin \theta)^2)(x_a^2 + y_a^2) \\
&= x_a^2 + y_a^2
\end{aligned} \tag{1.8}$$

So, different observers would agree on the magnitude of vector \vec{a} . We say the magnitude of vector \vec{a} is *invariant* under orthogonal transformations.

1.4.2 Galilean Transformation

Say, we have two frames (S) and (S'). Frame (S') is moving with velocity \vec{v} relative to frame (S). At time $t = t' = 0$, $x = x' = 0$, $y = y' = 0$, and $z = z' = 0$.

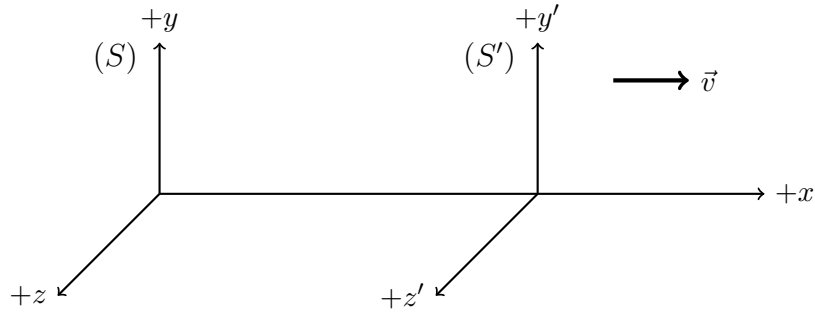


Figure 3: Frame (S') moves with speed \vec{v} relative to frame (S)

i) Galilean Transformation Equations

$$\begin{cases} x' = x - vt \\ y' = y \\ z' = z \\ t' = t \end{cases} \quad (1.9)$$

Again, we should also see what it looks like if we reversed the transformation:

$$\begin{cases} x = x' + vt' \\ y = y' \\ z = z' \\ t = t' \end{cases} \quad (1.10)$$

Another way to look at this system of transformation equations is to visualize the transformation in a space-time coordinate system. This will turn out to be a very useful strategy when we deal with more complex transformations in relativity.



Figure 4: Our first space-time diagram!

In (S) : Events A , B , and C .

$$\begin{cases} (t_A, x_A) = (2s, 3m) \\ (t_B, x_B) = (6s, 3m) \\ (t_C, x_C) = (2s, 1m) \end{cases} \quad (1.11)$$

What about movement in (S') ? Recall the transformation equations:

$$\begin{cases} t' = t \\ x' = x - vt \end{cases} \quad (1.12)$$

At $x' = 0$, $x = vt$. In other words, the line $x = vt$ is the path which the origin of (S') would draw in (S) if it were to be stationary in (S') .

ii) Galilean Velocity Transformation

We can use the “differential approach” to derive the Galilean velocity transformation. First, we write the velocities \vec{u} and \vec{u}' in (t, x, y, z) and (t', x', y', z') :

$$\vec{u} = \left\langle \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right\rangle \quad (1.13)$$

$$\vec{u}' = \left\langle \frac{dx'}{dt'}, \frac{dy'}{dt'}, \frac{dz'}{dt'} \right\rangle \quad (1.14)$$

Then, we examine the relationship between (t, x, y, z) and (t', x', y', z') :

$$u'_x = \frac{dx'}{dt'} = \frac{dx'}{dt} \frac{dt}{dt'} = \frac{d}{dt}(x - vt) = \frac{dx}{dt} - v \frac{dt}{dt} = u_x - v \quad (1.15)$$

Using a similar approach, we can also evaluate u'_y and u'_z . We can also evaluate u'_y and u'_z :

$$\begin{cases} u'_y = \frac{dy'}{dt'} = \frac{dy'}{dt} \frac{dt}{dt'} = \frac{dy}{dt} = \frac{dy}{dt} = u_y \\ u'_z = \frac{dz'}{dt'} = \frac{dz'}{dt} \frac{dt}{dt'} = \frac{dz}{dt} = \frac{dz}{dt} = u_z \end{cases} \quad (1.16)$$

Combining (1.16) and (1.17), we obtain the Galilean velocity transformation:

$$\begin{cases} u'_x = u_x - v \\ u'_y = u_y \\ u'_z = u_z \\ t' = t \end{cases} \quad (1.17)$$

Again, we should also look at the inverse transformation from (S') to (S) :

$$\begin{cases} u_x = u'_x + v \\ u_y = u'_y \\ u_z = u'_z \\ t = t' \end{cases} \quad (1.18)$$

iii) Direction

Imagine an object moving in the xy coordinates in frame (S) , at an angle θ to the x -axis. How does the object move in frame (S') ? Specifically, what angle does the object make with the x' -axis in (S') ?

First, we can easily work out the relationship between the angle θ and the vector components of \vec{u} .

$$\begin{cases} u_x = ||\vec{u}|| \cos \theta = u \cos \theta \\ u_y = ||\vec{u}|| \sin \theta = u \sin \theta \end{cases} \quad (1.19)$$

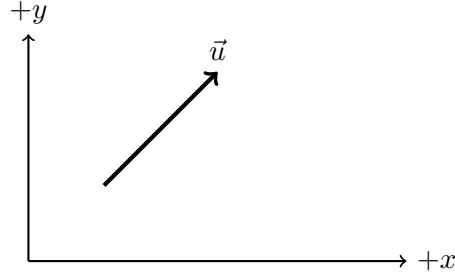


Figure 5: Velocity vector \vec{u} in Cartesian coordinates

$$\tan \theta = \frac{\sin \theta}{\cos \theta} = \frac{u_y}{u_x} \quad (1.20)$$

Similarly, we can write an expression for $\tan \theta'$ in (S') . Combining this newly derived expression with the velocity transformation equations, we obtain the relationship between θ' and θ :

$$\tan \theta' = \frac{\sin \theta'}{\cos \theta'} = \frac{u'_y}{u'_x} = \frac{u_y}{u_x - v} = \frac{u \sin \theta}{u \cos \theta - v} \quad (1.21)$$

This is the relationship between θ' and θ . If you might wonder what you can do with (1.21), think about the problem of calculating the angle a raindrop makes on your car window as you move with some speed v , assuming that raindrops fall vertically relative to ground.

iv) Acceleration Transformation

Surprise surprise! We use the good 'ol “differential approach” to derive the Galilean acceleration transformation.

$$\begin{cases} a'_x = \frac{du'_x}{dt'} = \frac{du'_x}{dt} \frac{dt}{dt'} = \frac{d}{dt}(u_x - v) = a_x \\ a'_y = \frac{du'_y}{dt'} = \frac{du'_y}{dt} \frac{dt}{dt'} = \frac{du'_y}{dt} = a_y \\ a'_z = \frac{du'_z}{dt'} = \frac{du'_z}{dt} \frac{dt}{dt'} = \frac{du'_z}{dt} = a_z \end{cases} \quad (1.22)$$

Note that these transformation equations are only true if v is constant. Otherwise, we would be dealing with noninertial reference frames, which is beyond our consideration of special relativity.

1.4.3 Newton's Laws and Galilean Relativity

$$\vec{F} = m\vec{a} \quad (1.23)$$

We have found that:

$$\begin{cases} a'_x = a_x \\ a'_y = a_y \\ a'_z = a_z \end{cases} \quad (1.24)$$

Hence, Newton's laws are *invariant* under a Galilean transformation.

2 Electromagnetism and Galilean Relativity

2.1 Maxwell's equations and light

Maxwell's equations summarize the rules for \vec{E} and \vec{B} . They predict that electromagnetic waves travel at speed:

$$c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}} = 3 \times 10^8 \text{ m/s} \quad (2.1)$$

where μ_0 is the permeability constant or the magnetic constant, and ε_0 is the permittivity of free space.

2.2 The “luminiferous ether”

Proposes that there is a medium in which $c = 3 \times 10^8 \text{ m/s}$. Are Maxwell's equations wrong? Or are Galilean's transformation wrong?

2.3 The Michelson-Morley experiment

A test of the “luminiferous ether” hypothesis. Experiment for measuring c .

2.3.1 The “horizontal”

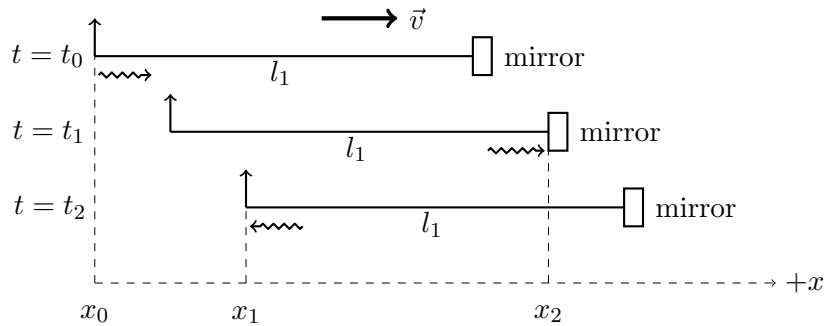


Figure 6: Horizontal arm, length l_1 , moving with speed v relative to the “ether”

Our goal is to work out Δt_h , the time it takes for light to travel from x_0 to x_2 , and x_2 back to x_1 . Now, we can let the speed of light be c . We should also let $t_a = t_1 - t_0$ and $t_b = t_2 - t_1$.

First, we can easily obtain the time it takes for light to travel each path:

$$\begin{cases} ct_a = l_1 + vt_a \\ ct_b = l_1 - vt_b \end{cases} \quad (2.2)$$

Following some algebraic rearrangements, we get t_a and t_b .

$$\begin{cases} t_a = \frac{l_1}{c-v} \\ t_b = \frac{l_1}{c+v} \end{cases} \quad (2.3)$$

Since the total time, Δt_h , is the sum of t_a and t_b :

$$\Delta t_h = t_a + t_b = \frac{l_1}{c-v} + \frac{l_1}{c+v} = \frac{2l_1c}{c^2 - v^2} = \frac{2l_1}{c} \cdot \frac{1}{1 - \frac{v^2}{c^2}} \quad (2.4)$$

2.3.2 The “vertical”

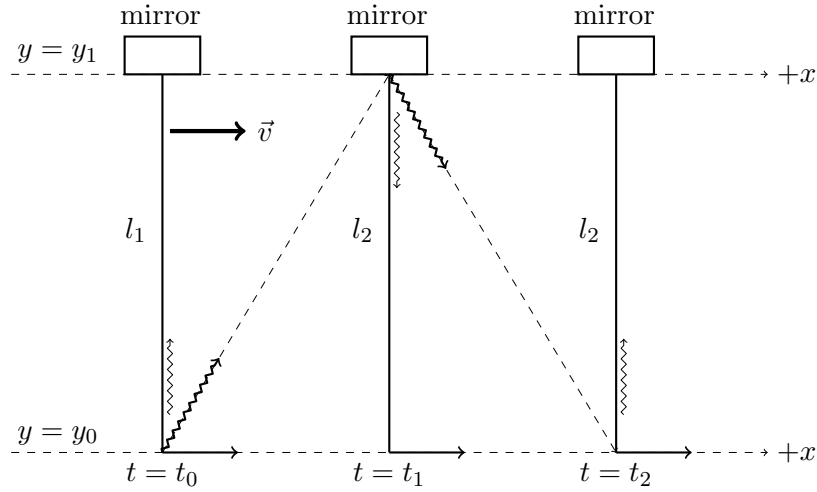


Figure 7: Vertical arm, length l_2 , moving with speed v relative to the “ether”

Our goal, similar that in the previous section, is to work out Δt_v , the time it takes for light to travel from y_0 to y_1 , and y_1 back to y_0 . Let $t_c = t_1 - t_0$ and $t_d = t_2 - t_1$.

Again, we obtain the time it takes for light to travel each path. However, note that the light paths in this case are different from those in the “horizontal” scenario. From the ground frame, light travels diagonal paths, as illustrated in

Figure 7 as dashed lines, rather than along the vertical arm of the interferometer. Hence, we use the Pythagorean theorem to find the path length.

$$\begin{cases} ct_c = \sqrt{l_2^2 - (vt_c)^2} \\ ct_d = \sqrt{l_2^2 - (vt_d)^2} \end{cases} \quad (2.5)$$

Following some algebraic rearrangements, we obtain t_c and t_d . Not surprisingly, these two quantities are equal.

$$t_c = t_d = \frac{l_2}{c} \cdot \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2.6)$$

Again, we obtain Δt_v by summing t_c and t_d .

$$\Delta t_v = t_c + t_d = \frac{2l_2}{c} \cdot \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2.7)$$

2.3.3 Michelson's first brilliant observation

Suppose that the speed v at which the interferometer travels relative to the "ether" is approximately the speed at which Earth orbits the Sun, the speed of light c is so great compared to v that $\Delta t_h \approx \Delta t_v$, and thus any difference between Δt_h and Δt_v would be humanly undetectable.

However, Michelson's use of interferometry allows for detections of slightest time differences between the vertical and horizontal arms. Interferometry, as its name suggests, works based on the ability of light to interfere. Differences between Δt_h and Δt_v , if they existed, would result in constructive and destructive interference, creating detectable "fringes." Hence, if "fringes" were observed, Michelson would be able to confirm that existence of the "luminiferous ether."

2.3.4 Michelson's second brilliant observation

Michelson's second brilliant idea is to rotate the interferometer, alternating the roles of the vertical and horizontal arms - vertical becoming horizontal, horizontal becoming vertical.

The purpose of alternating the roles of the vertical and horizontal arms of the interferometer is to create visible "fringe shifts." Mathematically speaking, rotating the apparatus allows for the removal of the speed of light c from part of the calculations and for the fractional time differences to be dependent only on the sum of the arm lengths rather than individual arm lengths, thus simplifying the machining and detection process.

While there are infinitely many angles the apparatus can be rotated relative to its original position, we can fully show the effect of rotating the interferometer by considering special cases, when $\theta = 0^\circ$ and $\theta = 90^\circ$.

First, we consider $\theta = 0^\circ$. We are interested in the difference between the time light takes in the vertical (Δt_{v_0}) versus horizontal arms (Δt_{h_0}). We shall call this quantity Δt_{0° .

$$\Delta t_{0^\circ} = \Delta t_{v_0} - \Delta t_{h_0} = \frac{2l_2}{c} \cdot \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} - \frac{2l_1}{c} \cdot \frac{1}{1 - \frac{v^2}{c^2}} \quad (2.8)$$

Now, let's consider Δt_{90° :

$$\Delta t_{90^\circ} = \Delta t_{v_{90}} - \Delta t_{h_{90}} = \frac{2l_2}{c} \cdot \frac{1}{1 - \frac{v^2}{c^2}} - \frac{2l_1}{c} \cdot \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2.9)$$

What we are interested now the difference between the time differences at Δt_{0° and Δt_{90° . Let this quantity be Δt .

$$\Delta t = \Delta t_{90^\circ} - \Delta t_{0^\circ} = \frac{2}{c}(l_1 + l_2) \left(\frac{1}{1 - \frac{v^2}{c^2}} - \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \right) \quad (2.10)$$

By now, we must have found the term $\frac{v^2}{c^2}$ quite ubiquitous. So, we can (in fact we really should) assign to it a more elegant symbol β . (2.10) now becomes:

$$\Delta t = \Delta t_{90^\circ} - \Delta t_{0^\circ} = \frac{2}{c}(l_1 + l_2) \left(\frac{1}{1 - \beta^2} - \frac{1}{\sqrt{1 - \beta^2}} \right) \quad (2.11)$$

Since $v \ll c$, or in other words $\beta \ll 1$, we can simplify (2.11) by approximating the terms containing β . Specifically, we are using the following approximation,

$$(1 + x)^p \approx 1 + px \quad (2.12)$$

for $x \ll 1$, to obtain a simpler version of (2.11):

$$\begin{aligned} \Delta t &= \frac{2}{c}(l_1 + l_2) \left(\frac{1}{1 - \beta^2} - \frac{1}{\sqrt{1 - \beta^2}} \right) \\ &\approx \frac{2}{c}(l_1 + l_2) \left[(1 + \beta^2) - \left(1 + \frac{1}{2}\beta^2 \right) \right] \\ &= \left(\frac{l_1 + l_2}{c} \right) \beta^2 \end{aligned} \quad (2.13)$$

Now, in order to achieve maximum "fringe shifts," we (or Michelson more exactly) would want to have the time difference Δt to be half the "period of light"

- meaning that light on one path travels a half wavelength more than light on the other path. Hence, we can let:

$$\begin{aligned}\frac{\Delta t}{T} &= \left(\frac{l_1 + l_2}{c} \right) \beta^2 \left(\frac{c}{\lambda} \right) \\ &= \left(\frac{l_1 + l_2}{\lambda} \right) \beta^2 = 0.5\end{aligned}\tag{2.14}$$

where $\lambda = 590 \text{ nm}$ is the wavelength of the sodium source which Michelson used in his experiment and $T = \frac{1}{\nu} = \frac{\lambda}{c}$ is the "period" of light.

Earlier, we made an assumption that v is the speed of the Earth around the Sun. Hence, $\beta \approx 10^{-4}$. So, to take $\frac{\Delta t}{T} \approx 0.5$, the sum of the arm lengths should be around $20 - 30 \text{ m}$, which was totally achievable in a laboratory setting. In fact, Michelson chose to make the arm lengths 11 m each, which, if the "ether" hypothesis were true, should give $\frac{\Delta t}{T} \approx 0.4$.

As it turned out, however, Michelson did not detect any changes in the interference patterns. In fact, he found that $\frac{\Delta t}{T} < 0.005$.

2.4 A different experiment: Bradley's Stellar Aberration (1725)

A different experiment, on the other hand, gave results in agreement with the "ether drift" theory.

Astronomers in the 18th century realized they had to tilt their telescopes by some angle α away from the direction in which light was coming towards their location θ , in order to be able to see the stars. We immediately obtain the relationship:

$$\alpha = \theta - \theta' \tag{2.15}$$

While not absolutely necessary, we should take a peek at how the angle θ' can be calculated. We will not go into detail how the following expression is obtained. However, its derivation is rather simple solution to a geometry problem.

$$\frac{u_x'}{u} = \cos(\theta') = \frac{u \cos(\theta) - v}{\sqrt{u^2 + v^2 - 2uv \cos(\theta)}} \tag{2.16}$$

where u is the speed of light c , and u_x' is the speed of light in the x -direction, in the moving frame. Now, we shall consider the case in which star light is coming straight down, making a 90° angle with the horizontal:

$$\cos(\theta') = \left(\frac{-v}{c} \right) \frac{1}{\sqrt{1 + \frac{v^2}{c^2}}} \approx -\frac{v}{c} \tag{2.17}$$

which makes a lot of geometric sense. (This is similar to the raindrop problem).

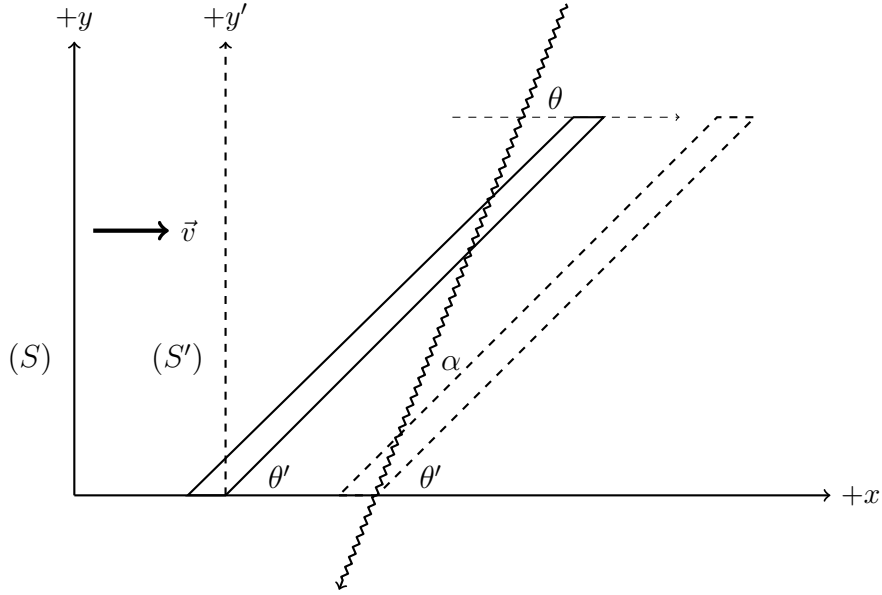


Figure 8: Bradley's stellar aberration set-up (1725) (exaggerated)

3 Einstein's Postulates of Relativity

Can we reconcile the Michelson-Morley experiment and Bradley's Stellar Aberration? Are the laws of Electricity and Magnetism independent of any inertial frame?

3.1 Einstein's realization about electricity and magnetism

Einstein's theory of special relativity came about as a result of his studying Maxwell's equations of electricity and magnetism. In fact, in his first paper on special relativity - *On the Electrodynamics of Moving Bodies*, published in his golden years (*Annus Mirabilis*), Einstein wrote:

"It is known that Maxwell's electrodynamics—as usually understood at the present time—when applied to moving bodies, leads to asymmetries which do not appear to be inherent in the phenomena. Take, for example, the reciprocal electrodynamic action of a magnet and a conductor. The observable phenomenon here depends only on the relative motion of the conductor and the magnet, whereas the customary view draws a sharp distinction between the two cases in which either the one or the other of these bodies is in motion. For if the magnet is in motion and the conductor at rest, there arises in the neighbourhood of the magnet an electric field with a certain definite energy, producing a current at the places where parts of the conductor are situated. But if the magnet is stationary and the conductor in motion, no electric field arises in the neighbourhood of the magnet. In the conductor, however, we find an electromotive force, to which in itself there is no corresponding energy, but which gives rise—assuming equality

of relative motion in the two cases discussed—to electric currents of the same path and intensity as those produced by the electric forces in the former case.”

Essentially, what Einstein realizes from studying Maxwell’s equations is that Electricity and Magnetism are two sides of the same coin.

3.2 Postulates of Special Relativity

3.2.1 Einsteins postulates

(a) All laws of physics are the **same** in all inertial reference frames. In other words, all laws of physics must retain their mathematical forms regardless of any reference frames in which they are considered.

(b) The speed of light is the **same** in all inertial reference frames. This postulate is also called the *The Principle of the Constancy of the Speed of Light*. This means the speed of light is always measured to be c , independent of reference frames. For example, both you (standing on Earth) and the I.S.S. (which is orbiting Earth at dazzling speeds) agree that $c = 299\,792\,458\,\text{m s}^{-1}$.

4 Fundamental consequences of the postulates

This is the prelude to the “Lorentz-Einstein Transformation,” a very useful mathematical tool for solving special relativity problems. However, before we can into more rigorous and “mathy” meat of special relativity, we will first discuss the consequences of Einstein’s postulates—in order to gain a better understanding of the rules of special relativity and hopefully develop some intuition and prepare our minds for the less intuitive stuff ahead.

4.1 Relativity of Simultaneity

In this section, we will attempt to answer whether simultaneous events in some frame (S) remain simultaneous in some other frame (S') that is moving with speed v relative to (S). In order to do so, let’s meet Alice and Bob, our lab assistants:



It’s okay if you are socially awkward. You’ll see them enough times that greetings will eventually become unnecessary. Anyway, consider the following situation: Alice is in a train moving at speed v relative to the ground. Bob stands on the ground. A light source in the middle of Alice’s car fires two pulses in opposite directions, parallel to the train’s direction of motion.

For Alice, the two pulses arrive at the ends of the train simultaneously, illustrated below:

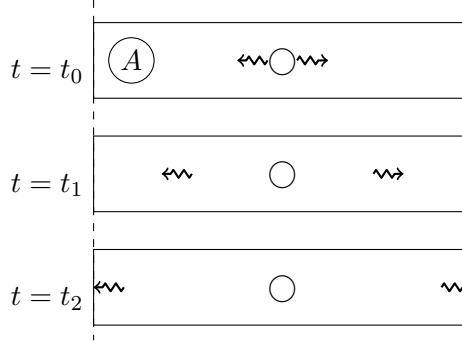


Figure 9: Alice's train in Alice's frame, over time

For Bob, however, things are little different. According to Einstein's second postulate of relativity, the speed of light is constant. Therefore, we should expect that Bob, in the ground frame, sees the left pulse arrive at the end of Alice's train **before** the right pulse arrive at the other end. As we visualize the problem, it becomes fairly obvious that simultaneous events in Alice's frame are not simultaneous in Bob's frame.

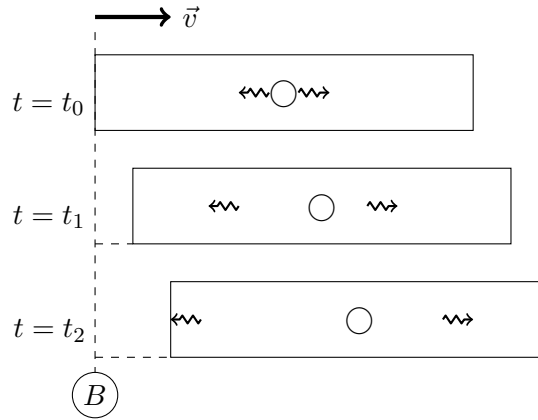


Figure 10: Alice's train in Bob's frame, over time

This visualization certainly useful - because it approximates very well what happens in space at certain points in time. However, the sketching process can become too tedious when we deal with more complex relativity problems that involves more relatively moving objects. An alternative and more effective route to visualizing relativity problem, as we have touched on earlier, is drawing space-time diagrams: one for Alice, one for Bob. For Alice's diagram, we can see that the pulses arrive at the rear and front of the train at the same time.

As we have shown earlier, this (simultaneity) is not the case in Bob's frame. In

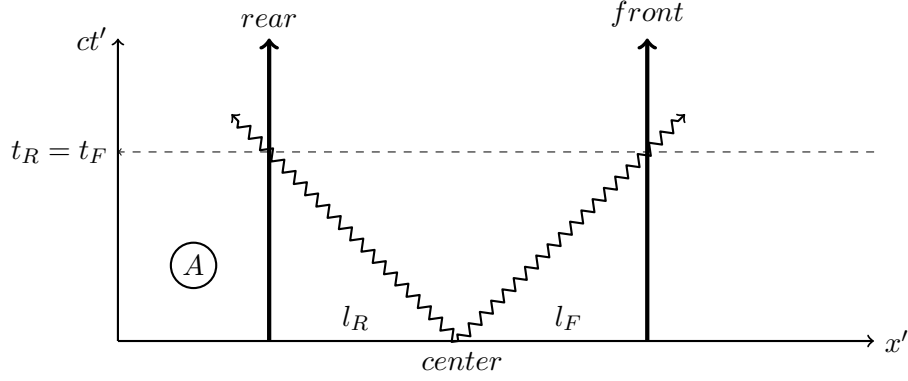


Figure 11: Alice's spacetime diagram

fact, the left pulse hits the **rear** of the Alice's train **before** the right pulse hits the **front** of the train.

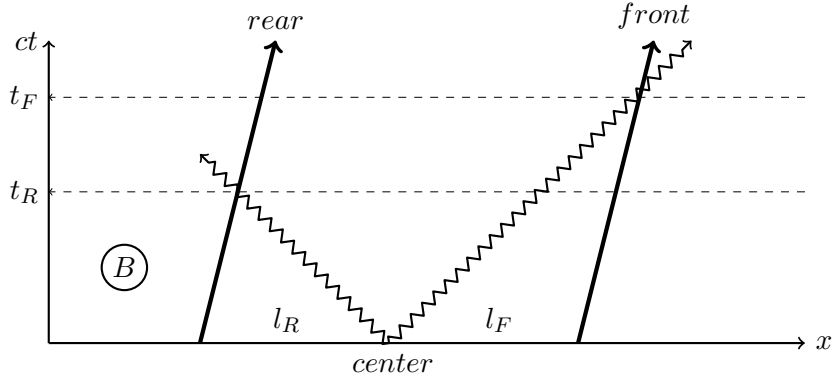


Figure 12: Bob's spacetime diagram

Now that we know simultaneity is not independent of reference frames, we shall attempt to find the difference between the pulses' arrival times. We shall stick with our previous setup, with a little change. For later convenience, we will change the location of the light source in the train so that the pulses arrive at the front and rear of Alice's train **simultaneously** in Bob's frame. Hence, we should re-draw Alice's and Bob's spacetime diagrams.

Since the light pulses arrive the ends of Alice's train simultaneously in Bob's frame, we can write l_r and l_f in terms of the length of the train, in Bob's frame, L_B . We have the following system of equations to start:

$$\begin{cases} l_R + l_F = L_B \\ t_R = t_F \\ ct_R = l_R - vt_R \\ ct_L = l_L + vt_L \end{cases} \quad (4.1)$$

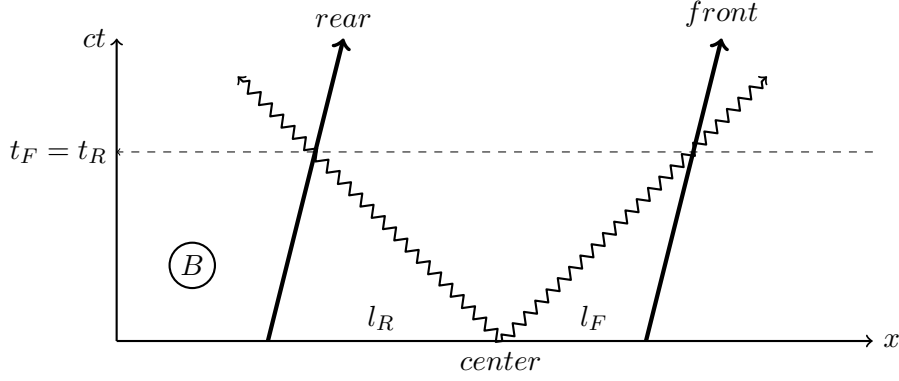


Figure 13: Bob's new spacetime diagram

With a little bit of algebraic rearrangements, we can figure out how L_B should be divided so that the light pulses reach the ends of Alice's train simultaneously in Bob's frame:

$$\begin{cases} l_F = \frac{L_B}{2} \left(1 - \frac{v}{c}\right) \\ l_R = \frac{L_B}{2} \left(1 + \frac{v}{c}\right) \end{cases} \quad (4.2)$$

This is also how L_A is divided in Alice's frame.

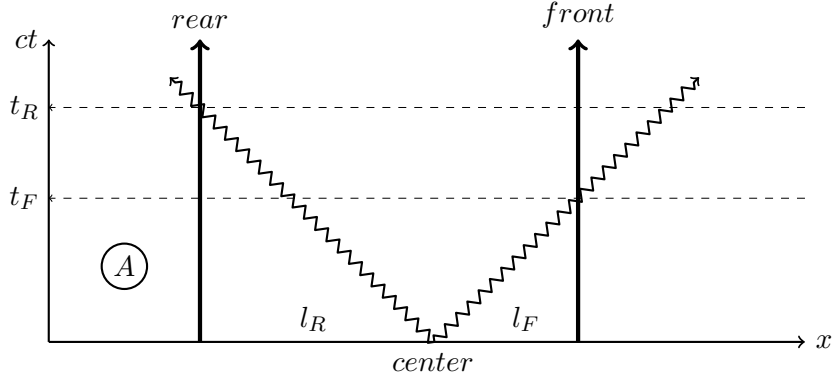


Figure 14: Alice's new spacetime diagram

We can express l_R and l_F in Alice's frame in terms of L_A , v , and c .

$$\begin{cases} l_F = ct_F' = \frac{L_A}{2} \left(1 - \frac{v}{c}\right) \\ l_R = ct_R' = \frac{L_A}{2} \left(1 + \frac{v}{c}\right) \end{cases} \quad (4.3)$$

Now we can write down the expression for the time interval between the pulses' arrivals in Alice's frame:

$$\Delta t' = t_R' - t_F' = \frac{L_A v}{c^2} \quad (4.4)$$

Keep in mind that all of the calculations we have been doing are in Bob's frame. More specifically, we based all of our derivations on events that are simultaneous in Bob's frame. What we found, according to (4.4), is that Bob sees not only the **loss of simultaneity** in Alice's frame, but also Alice's rear clock giving larger reading than Alice's front clock. This is another fundamental effect of Einstein's postulates: **rear clock ahead**, which we will visualize and simplify in the following section.

4.1.1 Simultaneity and clock readings

The diagram below sums up what happens to Alice's clock readings **in Bob's frame**.

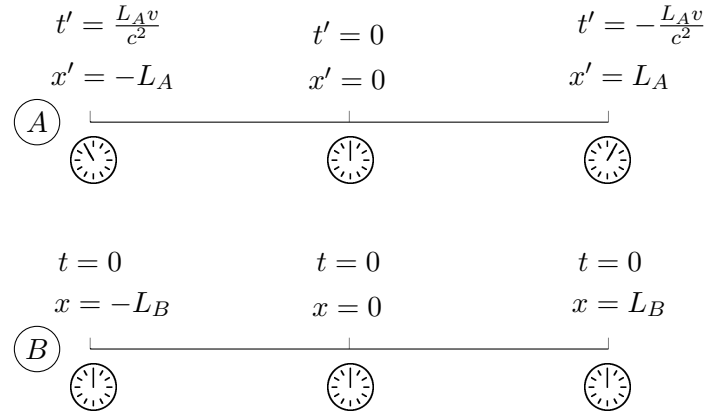


Figure 15: Alice's clock readings, in Bob's frame

In relativity, it is very important to specify in which frame we are working, because as we demonstrated earlier Alice's and Bob's measurements of the same events do not necessarily agree at all times. As we dig deeper into other fundamental effects of relativity, we have to be more and more careful with our assumptions, especially those about time intervals and distances!

It is appropriate now to point out that all of our sketches so far (from spacetime diagrams to trains) have been based on an assumption that Bob and Alice both agree on the length of Alice's train. However, while we are still unsure whether our assumption will hold true, which unfortunately (or fortunately) it will not, we shall not panic and there is a reason not to be: we did not make such assumption in our mathematics! In fact, keen readers will have noticed the use of different symbols L_A and L_B to express the length of the same train, in different frames. So, while our sketches maybe a little too sloppy, we can be confident in the results we have derived so far.

In the next section, we will take a short break from sketches of train, and rather focus on simultaneity and electromagnetism. Specifically, we will play the role

of Einstein and look into a puzzling connection between relativity and electromagnetism.

4.1.2 Simultaneity and Magnetism: A practical example

In this practical example, we will be (safe) in the stationary frame of our laboratory. In the moving frame, on the other hand, are electric charges in a long, thin conductor. In this exercise, we will attempt to derive the expression for the Lorentz force acting on the charged particle from the stationary lab frame in which the particle moves with speed \vec{v} and in the particle's frame in which it is at rest.

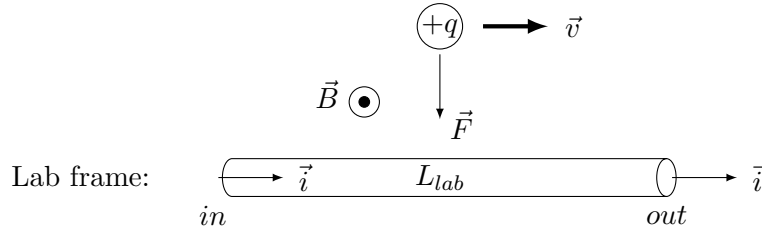


Figure 16: Lab frame: EM and Relativity "experimental" setup

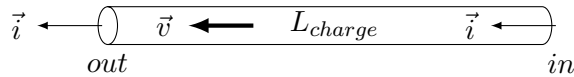
Recall the expression for the magnitude of the magnetic field as a function of distance from a long, thin conductor with current \vec{i} :

$$||\vec{B}|| = \frac{\mu_0 i}{2\pi r} \quad (4.5)$$

Note that the only force acting the charged particle is the Lorentz force because the conductor, while contains a current, is neutral: $q_{in} = q_{out}$. Hence, from (4.6), working the lab frame, we can write down the expression for the magnitude of the Lorentz force acting on the particle:

$$||\vec{F}_{Lorentz}|| = q||\vec{v} \times \vec{B}|| = \frac{qv\mu_0 i}{2\pi r} \quad (4.6)$$

The next step is to teleport ourselves into the particle's frame. As a result, the particle is now at rest while the conductor and the current inside the conductor move in the opposite direction.



Particle frame:



Figure 17: Particle frame: EM and Relativity "experimental" setup

What's happening here is quite different from the previous scenario. The conductor in this case is no longer neutral because of the **rear clock ahead** effect. Precisely speaking, in the particle's frame, there is a net charge in the conductor at any given length L_{charge} because the particle always sees charges coming into the conductor **before** it sees charges leave. Therefore, $q_{in} > q_{out}$. There exists a net charge, and we should be able to work out how much net charge there is. First, based on the **rear clock ahead** effect, working in the frame of the charged particle, we know the time difference $\Delta t'$ between when charges come in and when charges come out of the conductor:

$$\Delta t' = \frac{v L_{charge}}{c^2} = \frac{L' v}{c^2} \quad (4.7)$$

where $L_{charge} = L'$ is the length of the conductor in the $+q$ particle's frame. So the net charge q' on the conductor, in the particle's frame is

$$q' = i \Delta t' = i \frac{L' v}{c^2} \quad (4.8)$$

Our goal is to find the electric force acting on the charged particle from a charged, long, thin conductor. Recall the expression for the electric field from a line of charge:

$$E = \frac{1}{2\pi\epsilon_0} \frac{\lambda}{r} \quad (4.9)$$

where r is the distance from the particle to the conductor, and $\lambda = \frac{q'}{L'}$ is the conductor's linear charge density.

Combining this expression with (4.7) and (4.8),

$$E = \frac{1}{2\pi\epsilon_0} \frac{\lambda}{r} = \frac{1}{2\pi\epsilon_0} \frac{1}{r} \frac{q'}{L'} = \frac{1}{2\pi\epsilon_0} \frac{iv}{rc^2} \quad (4.10)$$

Hence, the electric force acting on the charged particle is

$$F_q = qE = \frac{q}{2\pi\epsilon_0} \frac{iv}{rc^2} \quad (4.11)$$

Now, recall the relation between the speed of light c and μ_0 and ϵ_0 .

$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}} \quad (4.12)$$

So (4.11) is essentially (4.6)

$$F_q = qE = \frac{q}{2\pi\epsilon_0} \frac{iv}{rc^2} = \frac{qv\mu_0 i}{2\pi r} \quad (4.13)$$

4.1.3 Simultaneity and relativity: A cautionary tale

So far, we have mostly worked in Bob's frame, deriving expressions for time differences in her frame. We have come to a conclusion that simultaneous events

in Bob's frames are not necessarily simultaneous in Alice's frame. We have also found that while Alice and Bob both have synchronized clocks in their respective frames, Bob perceives Alice's rear clocks to have higher readings than her front clocks do. The next question we will be trying to answer, working in Bob's frame is: What readings does Alice see on Bob's clock if events are simultaneous in Alice's frame? To answer this question, we will briefly take a look at an old example involving some spacetime diagrams. Our goal is to write the expression for the time difference in Bob's frame (the ground frame), if events are simultaneous in Alice's frame. Since Alice is moving to the "right" of Bob, in Alice's frame, Bob is moving to the "left" of Alice. We obtain the two following spacetime diagrams.

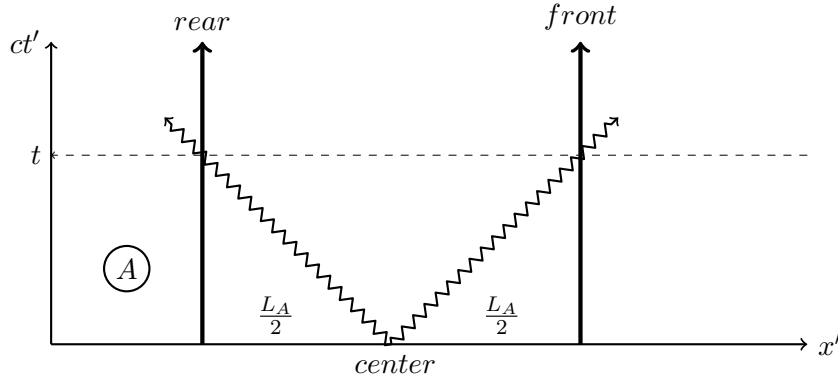


Figure 18: Alice's spacetime diagram, simultaneous events

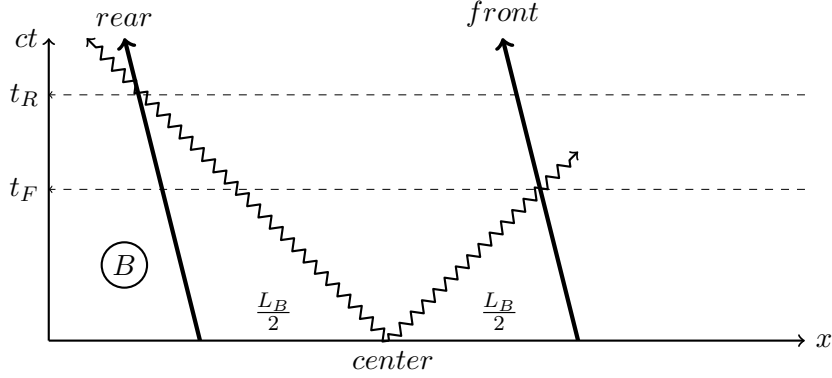


Figure 19: Bob's spacetime diagram, simultaneity in Alice's frame

Using our approach to the last example, we can quickly write down the expression for the time difference:

$$\Delta t_B = t_R - t_F = \frac{L_B}{2} \left(\frac{1}{c-v} - \frac{1}{c+v} \right) = \frac{L_B v}{c^2 - v^2} = \frac{L_B v}{c^2} \left(\frac{1}{1 - \frac{v^2}{c^2}} \right) \quad (4.14)$$

At this point we must have seen this term $\frac{1}{1-v^2/c^2}$, or more precisely, its square root $\frac{1}{\sqrt{1-v^2/c^2}}$ enough times that we should assign to it a name. We shall be conventional and call this factor γ .

$$\Delta t_B = t_R - t_F = \frac{L_B v}{c^2} \left(\frac{1}{1 - \frac{v^2}{c^2}} \right) = \frac{v L_B \gamma^2}{c^2} \quad (4.15)$$

If we compare the form of (4.15) to that of (4.4), we can see they differ by the square of factor γ . In the following sections, we will explore in more detail how the square of γ comes to appear in (4.15), as well as its ubiquity and vital role in relativity.

4.2 Relativity and how to understand space and time

So, just a recap, we shall review a few yet important rules to keep in mind when dealing with problems in special relativity:

First, all observers agree on events. Recall our example of Alice and Bob. While they may disagree on the simultaneity of events, depending on their relative motions, they must and always agree that the light pulses arrive at the ends of the train. In same rule also applies to our previous example involving relativity and magnetism. Both the moving charged particle $+q$ and us agree that it is attracted to the conductor with current i . We should also extend beyond the scope of our examples. The order of events is also crucial in relativity. If, for instance, Alice calls and Bob responds, then regardless of which frame we are in (which results in certain time differences between the two events), we should never see the events in reverse orders, that is Bob responds **before** Alice calls. This is also known as the principle of causality.

Second, the principle of relativity and constancy of the speed of light. Einstein's postulates are the foundation for all of our derivations. Indeed, without the assumption that all observers agree on the universal speed of light regardless of their reference frame, we would not be able to derive **rear clock ahead**, and different approaches to find the force on the charged particle will give different answer in different reference frames, which violate the first rule.

Third, all situations need to be measured with real tools: Sticks and Clocks. Note that we abstain from using the phrase: "Bob sees the light pulse arriving..." or "Alice sees Bob's clock readings..." This is intentional in order for an observer to "see" an event, light (or more precisely, information) from that event must reach the observer's eye, and then be processed by the brain... The point is that in special relativity, we often talk about measurements made in one frame or another, and these measurements are typically of "time," which is given by clock readings, and "location," which is given by the number of meter sticks. Our goal is to associate measurements with physical objects that belong to certain frames because as we will find out, one meter or one second in one frame are not necessarily the same in other frames. By knowing from which

frames certain measurements are made and how these frames move relative to other frames, we can then “convert” these measurements between frames.

So, a quick recap on 3 important rules to keep in mind:

- All observers agree on events.
- The principle of relativity and constancy of the speed of light.
- All situations need to be measured with real tools.

In the following sections, we will be investigating other effects of special relativity. Specifically, we will explore how relative motions between different frames affect measurements of lengths and time intervals in another frame. In a physical sense, we are trying to answer questions like: “How many meters is a meter in a moving frame?” or “How many seconds is a second in a moving frame?”

4.3 Transverse length measurements

First, we are interested in whether different observers agree on transverse lengths. Consider a thought experiment in which a cylinder of radius r enters an annular cylinder with inner radius r with velocity \vec{v} .

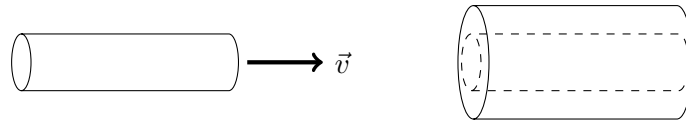


Figure 20: Thought experiment: Cylinders and transverse lengths

We as observers are in the ground frame along with the annular cylinder. In our frame, the smaller cylinder is moving to the right with speed v . Three things can happen to transverse lengths (specifically radius r) of the smaller cylinder in our frame:

- One, transverse lengths contract.
- Two, transverse lengths expand.
- Three, transverse lengths neither contract nor expand.

Now, all observers, regardless of their reference frames, must agree that the cylinder has to fit the annular cylinder (because radius of the smaller cylinder is the same as the inner radius of the annular cylinder). Therefore, in our frame, the transverse lengths of the smaller cylinder must not expand. Thus, we rule out the second hypothesis.

The first hypothesis is a no more simple to rule out. Assume that transverse lengths in moving frames contract. Then if we put ourselves in the smaller cylinder’s frame, we would measure the inner radius of the annular cylinder to be smaller than r , so the smaller cylinder does not fit inside the annular cylinder.

This contradicts a point we made earlier about all observers must agree that the smaller cylinder can fit inside the annular cylinder. Hence, we rule out the first hypothesis.

This leaves us with the only (and correct) hypothesis that transverse lengths neither expand nor contract. Based on the “invariance” of events, we have successfully argued that moving objects “retain” their stationary transverse dimensions.

4.4 Time dilation

The next fundamental effect of relativity is **time dilation**, or “moving clocks run slow.” In order to demonstrate this effect and mathematically describe its consequences, we use a classic (thought) experiment which involves a light-pulse clock, Alice, Bob, and their respective trains.

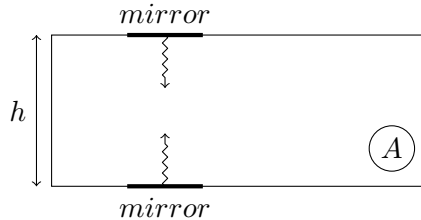


Figure 21: Light-clock experiment: Alice’s frame

A light-pulse clock in Alice’s frame is shown in Figure 21. The clock works by sending a pulse of light into a pair of mirrors, one placed on the floor of the train, the other on the ceiling, both aiming to keep light bouncing back-and-forth between themselves. Unlike regular clocks which ticks every second, the light-pulse clock ticks every time the pulse of light hits the bottom mirror. Given c is the speed of light and h is the height of the train, we obtain the tick-period of this clock, in Alice’s frame

$$\tau_A = \frac{2h}{c} \quad (4.16)$$

In Bob’s frame, Alice (and her clock) is moving to the right with speed v . As we might expect, Bob sees things a little bit differently.

Our goal is to find τ_B , which in this case is the time light takes to travel the diagonal path twice (up and down in Alice’s train). We can approach this problem the same way we did in an earlier example with the Michelson-Morley interferometer. First, we can express the distance light travels in τ_B . Note here that the train’s height h is retained.

$$c\tau_B = 2\sqrt{h^2 + \left(\frac{v\tau_B}{2}\right)^2} \quad (4.17)$$

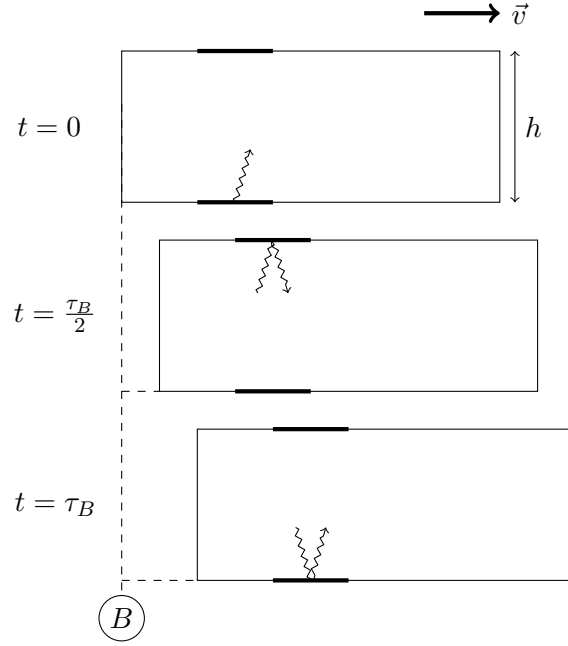


Figure 22: Light-clock experiment: Bob's frame

With some rearrangements, we obtain an expression for τ_B :

$$\begin{aligned}
 \tau_B &= \sqrt{\frac{4h^2}{c^2 - v^2}} \\
 &= \left(\frac{2h}{c}\right) \frac{1}{\sqrt{1 - v^2/c^2}} \\
 &= \frac{2h\gamma}{c}
 \end{aligned} \tag{4.18}$$

From (4.16) and (4.18), we obtain a very important relation:

$$\tau_B = \gamma\tau_A \tag{4.19}$$

This relation tells us that Bob would measure the tick-period of Alice's clock to be γ times larger than the tick-period of the clock in Alice's frame. In other words, **in Bob's perspective, Alice's clock runs slow.**

It is also worth mentioning that the same logic can be applied to a different situation where Alice measures Bob's tick-period. Using the same derivation and setup, we should arrive at $\tau_A = \gamma\tau_B$. It can be tempting to think of this solution as paradoxical and contradictory to what we obtained earlier. However, if we think very carefully about how measurements are made, then there is no such paradox! Consider the diagram below, which represents a visual resolution to this apparent "paradox."

Figure 23 aims to show an **asymmetry** in the comparison we are trying to make that leads to an apparent "paradox." While Bob uses **his clocks** to measure

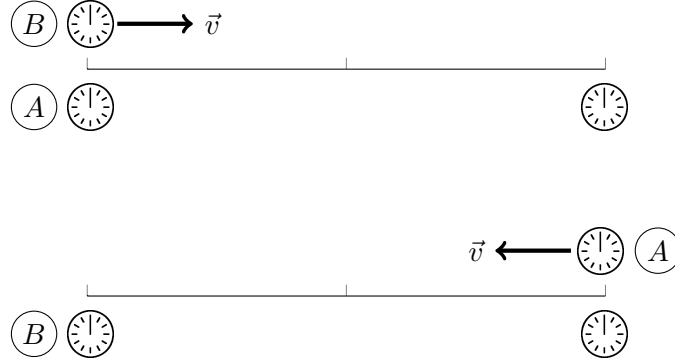


Figure 23: Resolution to time dilation “paradox”

Alice’s time, Alice uses **her clocks** to measure Bob’s time. And as we have found earlier, Alice and Bob do not always agree on their clock readings. Therefore, the statements “Alice’s clocks run slow” or “Bob’s clock run slow” are valid only if we also indicate the reference frames on which these statements are based.

4.5 Length contraction

In this section we are going to explore the fourth and last fundamental effect of special relativity: **length contraction**. Using a similar setup as that in the previous section, except for the orientation of the light clock, we will find the relation between L_A and L_B . Instead of sending vertical pulse, the light clock

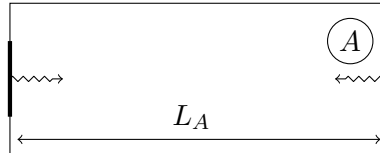


Figure 24: Light-clock experiment: Alice’s frame

in this setup sends a pulse of light in the direction of motion. In Alice’s frame, the tick-period is

$$\tau_{A'} = \frac{2L_A}{c} \quad (4.20)$$

where the prime is there just to help us distinguish the the longitudinal tick-period (which is what we are trying to find) from the vertical tick-period that we have found earlier.

In Bob’s frame, the light pulses have to travel extra/less distances depending on their direction of motion. This is due to the rightward motion of Alice’s frame with respect to the ground frame.

Similar to what we did in the previous section, we can set up an equation to

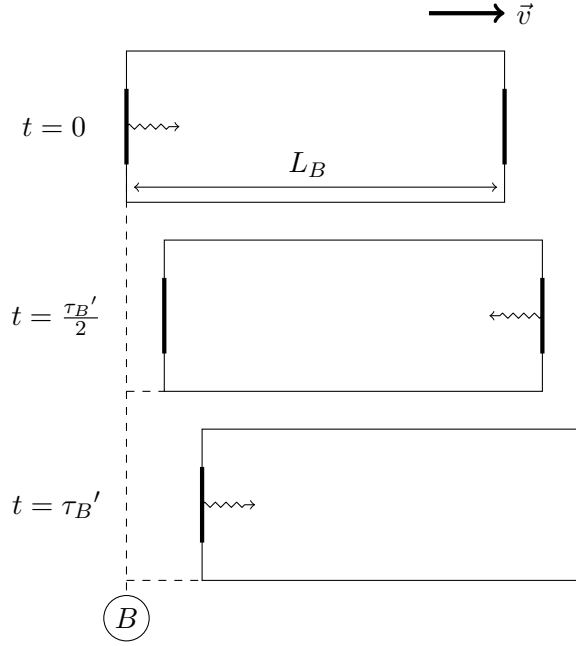


Figure 25: Light-clock experiment: Bob's frame

solve for τ_B'

$$\tau_B' = \frac{L_B}{c+v} + \frac{L_B}{c-v} = \frac{2L_B c}{c^2 - v^2} = \left(\frac{2L_B}{c} \right) \frac{1}{1 - v^2/c^2} = \frac{2L_B}{c} \gamma^2 \quad (4.21)$$

From (4.19), we know that $\tau_B' = \gamma \tau_A'$. So from (4.20) and (4.21), we get the relation between length measurements between relatively moving frames:

$$L_B = \frac{L_A}{\gamma} \quad (4.22)$$

This relation tells us that Alice's (longitudinal) lengths appear to be γ times smaller in Bob's frame. Hence we call this effect **length contraction**. Note that Alice would also measure longitudinal lengths in Bob's frame to be γ times smaller than their lengths in Bob's frame. While these (true) statements seem paradoxical, recall the argument we made about **asymmetrical** comparisons. Both Alice and Bob measure the other's longitudinal lengths as contracted. But this is totally fine (no physics is violated here) because both are in different point of view. Special relativity is all about perspectives—frames of reference.

4.6 Reflection on fundamental effects

So, just a quick recap on the basic idea—the one single cause of all fundamental effects of special relativity. We have found that clocks are out-of-sync between relatively moving frames. We have found that clocks in moving frames run slow.

We have also found that objects shrink in their direction of motion. We should ask ourselves: “What is the one cause of all these phenomena?” The answer is **perspective**.

In the following two subsections, we shall both test and apply the fundamental effects of special relativity to derive and make predictions of other consequences of relativity. In the first subsection, we will first look at a different approach to derive time dilation, using a regular clock instead of a light clock. Then, we will check for consistency among time dilation, length contraction, and rear clock ahead by deriving length contraction from the two other fundamental effects. In the second subsection, we will use the known fundamental effects to explore how we can convert measurements of velocities from one frame to another.

4.6.1 Time dilation with regular clocks

We shall attempt to derive time dilation in a more practical manner (by which I mean clock that does not require a single pulse of light to bounce indefinitely many times between two mirrors): by using a clock that sends out radio pulses every τ seconds in its rest frame. Now, we assume that this clock ticks with period $g_v\tau$ and is moving **away** from us, the observers, at speed v . Let $f_v\tau$ denotes the period which we **detect** the pulses. Since following every tick, the clock moves an extra distance $c\Delta t = vg_v\tau$ to get to us, we get an expression for $f_v\tau$:

$$f_v\tau = g_v\tau + \frac{v}{c}g_v\tau \quad (4.23)$$

Hence,

$$f_v = g_v \left(1 + \frac{v}{c}\right) \quad (4.24)$$

If the clock moves **towards** us, then we get an expression for $f_{-v}\tau$

$$f_{-v}\tau = g_v\tau - \frac{v}{c}g_v\tau \quad (4.25)$$

which gives

$$f_{-v} = g_v \left(1 - \frac{v}{c}\right) \quad (4.26)$$

In the trivial case in which the clock is at rest, we would detect the pulses coming with period τ . Our goal now is to relate f_v , f_{-v} , and g_v in a way such that we can express g_v in terms of v and show how τ depends on g_v . In order to do so, we shall consider the following scenario:

The light clock with period τ is in the ground frame, but at some distance away from us. Now, a spaceship is traveling from the light clock towards us at speed v . Its job is to receive the light pulses from the clock and **instantaneously** transmit the pulse towards us (so the pulse is not interrupted). Since the spaceship is moving with speed v away from the clock, it **sees** the pulses with period $f_v\tau$. Therefore, it also transmit at period $f_v\tau$. As a result, we (the ground frame) receives the pulses from the ship (which is moving towards us) at period $f_{-v}f_v\tau$.

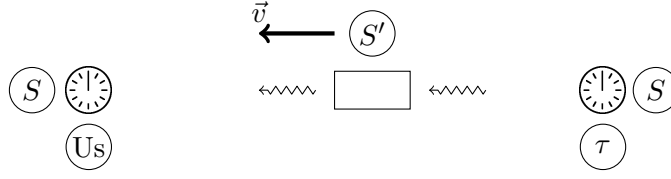


Figure 26: Time dilation from a regular clock

However, at the same time, we must see the light coming from the stationary clock period τ because the pulses are not delayed at the spaceship. Hence, we obtain an equation:

$$\tau = f_{-v} f_v \tau \quad (4.27)$$

Therefore,

$$f_{-v} f_v = g_v^2 \left(1 - \frac{v}{c}\right) \left(1 + \frac{v}{c}\right) = 1 \quad (4.28)$$

Following some rearrangements, we obtain this familiar expression for g_v :

$$g_v = \sqrt{\frac{1}{1 - v^2/c^2}} \quad (4.29)$$

Indeed, $g_v = \gamma$. So we can conclude clocks run slow in moving frames:

$$\tau_v = \gamma \tau_0 \quad (4.30)$$

At this point, we might be tempted to ask ourselves whether time dilation has ever been observed before—whether time dilation is **real**. And the answer is yes! In fact, the only reason we can detect large numbers (or any number at all) of muon particles μ^- from outer space at very low altitudes on Earth is due solely to time dilation.

Muons are subatomic particles that are very much similar to the electrons except for their greater mass and incredibly short half-life—approximately $150\mu s$. As these particles enter Earth's atmosphere, they enter at relativistic speeds close to the speed of light. Hence, despite their short lifetime, some muons indeed survive their entry to Earth and trigger ground detectors. What is so fascinating about the detection of muon particles, however, is the fact that they are detected in large numbers. At $2000m$, there are approximately $570\mu^-/hr$. Given their velocities and decay rate (which is found experimentally using particle accelerator), it is expected that at $50m$ above ground there should be less than $50\mu^-/hr$. In reality, however, scientists detected around $400\mu^-/hr$, significantly greater than the value predicted by our non-relativistic hypothesis.

We are certainly not so much puzzled by this fact because we have been equipped with the knowledge of time dilation. Muons are moving towards us with great (and by great I mean relativistic) speeds, thus their “clocks,” as seen from our frame, run slow. As a result, in our frame, these muons take longer than $150\mu s$ to decay and therefore are more likely to travel further down towards Earth. In their rest frame, however, a muon's half-life is still $150\mu s$.

4.6.2 Length contraction revisited

Before moving on to the next subsection, we shall take a look back at the fundamental effects of special relativity—especially to check their consistency with each other. In this section, we will derive (yet again) length contraction, not using a light clock, but from rear clock ahead and time dilation. If our derivation leads us to the same conclusion we made about length contraction, then we can be confident that special relativity works (at least so far).

4.6.3 Velocity “addition”

In this subsection we explore how velocities change as we move from one frame to another. At the moment, we have known that Alice and Bob disagree on their measurements of distances and time intervals, so it is reasonable to be careful about their agreement on velocities—the change of distance over time.

Specifically, we are interested in how Alice’s measurements of velocities change as we move from her frame to Bob’s frame. Our goal is to setup an experiment in which a particle moves with speed u' in Alice’s frame and apply our knowledge of the fundamental effects of special relativity (time dilation, length contraction, read clock ahead) to (hopefully) derive an expression for the speed u of the same particle in Bob’s frame, in relation to u' .

Consider two events: a ball (1) leaving the rear of a train, and (2) arriving at the front of the train. In Alice’s frame, she measures $\Delta x' = L_A$ as the distance the ball traverses and $\Delta t'$ as the time between these two events. In Bob’s frame, he measures $\Delta x = L_B$ and Δt . Let us define $u' = \Delta x' / \Delta t'$ as the ball’s velocity in Alice’s frame, and $u = \Delta x / \Delta t$ as the ball’s velocity in Bob’s. The train is stationary in Bob’s frame and is moving at speed v in Alice’s frame. Hence:

$$\Delta t' = \frac{L_A}{u'}, \quad (4.31)$$

and

$$\Delta t = \frac{L_B}{u - v} \quad (4.32)$$

Working in Bob’s frame, from length contraction, we know that:

$$L_B = \frac{L_A}{\gamma} = \frac{L_A}{\sqrt{1 - v^2/c^2}}. \quad (4.33)$$

And from time dilation and loss of simultaneity:

$$\Delta t = (t_2 - t_1) = \gamma \Delta t' = \gamma(t'_2 - t'_1) = \gamma \left(\frac{L_A}{u'} + \frac{L_A v}{c^2} \right). \quad (4.34)$$

From 4.31 to 4.34, we get:

$$\begin{aligned}
\gamma \left(\frac{L_A}{u'} + \frac{L_A v}{c^2} \right) &= \frac{L_A}{\gamma(u-v)} \\
\therefore \gamma \left(\frac{1}{u'} + \frac{v}{c^2} \right) &= \frac{1}{\gamma(u-v)} \\
\therefore \frac{c^2 - u'v}{u'c^2} &= \frac{1}{\gamma^2} \cdot \frac{1}{u-v} = \left(1 - \frac{v^2}{c^2} \right) \cdot \frac{1}{u-v} \\
\therefore (c^2 - u'v)(u-v) &= (u'c^2) \left(1 - \frac{v^2}{c^2} \right) \\
\therefore \dots \\
\therefore u' &= \frac{u-v}{1 - \frac{uv}{c^2}},
\end{aligned} \tag{4.35}$$

which is the expression for the ball's velocity in Alice's frame, working in Bob's frame.

Similarly, working in Alice's frame:

$$u = \frac{u' + v}{1 + \frac{u'v}{c^2}}. \tag{4.36}$$

So far we have focused on deriving the expressions for velocity addition; however, the larger and more important point to note here is that velocities can never exceed c the speed of light regardless of where the observation takes place.

For example, let Bob be in a stationary frame. Alice is moving relative to Bob at speed $0.9c$, and a ball is also moving at speed $0.9c$ in Alice's frame. Without any knowledge of special relativity, one might say the speed of the ball in Bob's frame is $u = 0.9c + 0.9c = 1.8c$. However, with velocity addition:

$$u = \frac{u' + v}{1 + \frac{u'v}{c^2}} = \frac{1.80c}{1 + \frac{0.81c^2}{c^2}} \approx 0.9945c < c. \tag{4.37}$$

If $u' = c$, then:

$$u = \frac{c + u}{1 + \frac{cv}{c^2}} = \frac{c(1 + v/c)}{1 + v/c} = c. \tag{4.38}$$

What if $u' = 0$?

$$u = \frac{v}{1 + 0} = v. \tag{4.39}$$

Both of these two extreme cases make sense.

5 Lorentz-Einstein Transformation Equations

In previous sections, we have derived expressions for time and spatial quantities of the same events in different frames. However, just in case you haven't noticed by now, going between frames can get quite cumbersome.

The goal of this section is to formulate a way to transform spatial and time coordinates from one frame to another. The Lorentz-Einstein transformation equations will be our tool.

5.1 Linear transformation of coordinates

Let us define a set of equations which transform time and spatial coordinates in Bob's frame $(t, x, y, z)_S$ into Alice's frame $(t', x', y', z')_{S'}$:

$$\begin{aligned}x' &= a_{11}x + a_{12}y + a_{13}z + a_{14}t \\y' &= a_{21}x + a_{22}y + a_{23}z + a_{24}t \\z' &= a_{31}x + a_{32}y + a_{33}z + a_{34}t \\t' &= a_{41}x + a_{42}y + a_{43}z + a_{44}t\end{aligned}\tag{5.1}$$

We can write these equations in a form of a matrix equation:

$$\begin{bmatrix} x' \\ y' \\ z' \\ t' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix},\tag{5.2}$$

just to emphasize a point that the Lorentz-Einstein equations make a “linear transformation.” Our goal is to figure out what the matrix elements are.

5.2 Derivations

5.2.1 Transverse dimension

Let us revisit the thought experiment in 4.3 where a solid cylinder of radius r and length l travels at speed v and fits into a (stationary) annular cylinder of inner radius r and length l . (Note: all lengths are lengths) measured in the rest frame).

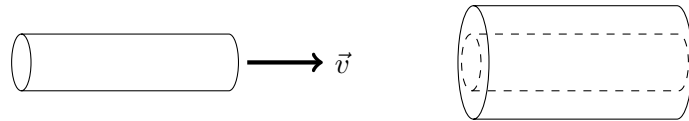


Figure 27: Thought experiment: Cylinders and transverse lengths

We have argued in 4.3 that since observers have to agree on events, observers agree on the transverse lengths y and z . So Eq. 5.2 becomes:

$$\begin{aligned}x' &= a_{11}x + a_{12}y + a_{13}z + a_{14}t \\y' &= y \\z' &= z \\t' &= a_{41}x + a_{42}y + a_{43}z + a_{44}t\end{aligned}\tag{5.3}$$

And the matrix equation becomes:

$$\begin{bmatrix} x' \\ y' \\ z' \\ t' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix}, \quad (5.4)$$

5.2.2 $x - t$ transformation

Since x' and t' cannot depend on y and z :

$$a_{12} = a_{13} = a_{42} = a_{43} = 0,$$

our matrix equation reduces to:

$$\begin{bmatrix} x' \\ y' \\ z' \\ t' \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 & a_{14} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ a_{41} & 0 & 0 & a_{44} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix}. \quad (5.5)$$

5.2.3 Dependence of x', t' on x, t

Now, we will break down the $x - t$ dependence on $x' - t'$, in three steps:

1/ Consider a pulse of light traveling in the +x-axis

Let λ and μ be arbitrary constants such that:

$$\begin{aligned} \lambda(x - ct) &= (x' - ct') \\ \mu(x + ct) &= (x' + ct') \end{aligned}$$

It follows that:

$$\begin{aligned} 2ct' &= (\mu - \lambda)x + c(\mu + \lambda)t \\ 2x' &= (\mu + \lambda)x + c(\mu - \lambda)t. \end{aligned}$$

Let $a = (\mu + \lambda)/2$ and $b = (\lambda - \mu)/2$, we get:

$$ct' = -bx + act \quad (5.6)$$

$$x' = ax - bct. \quad (5.7)$$

2/ Consider the motion of the origins of S and S'

$x' = 0$, so $avt - bct = 0$. Eq. 5.6 and Eq. 5.7 become:

$$ct' = a[ct - (v/c)x] = a(ct - \beta x) \quad (5.8)$$

$$x' = a[x - (v/c)ct] = a(x - \beta ct). \quad (5.9)$$

So in matrix form:

$$\begin{bmatrix} x' \\ y' \\ z' \\ t' \end{bmatrix} = \begin{bmatrix} a & 0 & 0 & -a\beta c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -a\beta/c & 0 & 0 & a \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix}. \quad (5.10)$$

3/ Apply the principle of relativity

The next reasonable step is to find what a is. Let us return to our classic thought experiment: Alice moving relative to Bob at speed v , both making measurements of a train of proper length l_0 .

Let everyone be initially at rest. Since Alice and Bob both have to agree on the length of the train:

- Working in (S) (Bob's frame), at $t = 0$: $x'_L = 0$ and $x'_R = l_0$, where x'_L is the position of the tail of the train, and x'_R is the position of the head of the train in Alice's frame.

- Working in (S') (Alice's frame), at $t' = 0$: $x_L = 0$ and $x_R = l_0$, where x_L is the position of the tail of the train, and x_R is the position of the head of the train in Bob's frame.

The question, then, is: "In (S') , at $t' = 0$, if $x'_L = 0$ then what is x'_R ?"

To answer this question we use Eq. 5.8 and Eq. 5.9. Substituting the expression for ct from Eq. 5.8 into Eq. 5.9, we get:

$$x'_R = a(x_R - \beta ct) = a \left[x_R - \frac{v}{c} \left(\frac{ct'}{a} + \frac{v}{c} x_R \right) \right]. \quad (5.11)$$

At $t' = 0$, with $x'_R = l'$ and $x_R = 0$, Eq. 5.11 simplifies to:

$$l' = al_0 \left(1 - \frac{v^2}{c^2} \right). \quad (5.12)$$

But we also know that at $t' = 0$ (in Alice's frame, of course):

$$x'_R = ax_R \equiv l_0 = al'. \quad (5.13)$$

Combining Eq. 5.12 and 5.13:

$$\frac{l_0}{a} = al_0 \left(1 - \frac{v^2}{c^2} \right). \quad (5.14)$$

Therefore:

$$a = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \gamma. \quad (5.15)$$

Our matrix equation becomes:

$$\begin{bmatrix} x' \\ y' \\ z' \\ t' \end{bmatrix} = \begin{bmatrix} \gamma & 0 & 0 & -\gamma\beta c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma\beta/c & 0 & 0 & \gamma \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix}. \quad (5.16)$$

We can tweak a little so that Eq. 5.16 looks more “symmetric” and conventional:

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix}. \quad (5.17)$$

5.3 The Lorentz-Einstein Equations

We can extract from Eq. 5.17 a set of equations called Lorentz-Einstein Equations.

$$\begin{cases} x' = \gamma(x - \beta ct) \\ y' = y \\ z' = z \\ ct' = \gamma(ct - \beta x) \end{cases} \quad (5.18)$$

Of course Eq. 5.17 and Eq. 5.18 are just different ways to express the same idea. Some may find the matrix equation more elegant and symmetric. Others may prefer the “equations notation.” At our level there is little advantage difference between the two expressions.

5.4 Some properties of the Lorentz-Einstein Transformations

5.4.1 Spatial separation and time interval between events

In (S) , let Event 1 be $(ct_1, x_1, y_1, z_1)_S$, and Event 2 be $(ct_2, x_2, y_2, z_2)_S$. (Recall that an “Event” can be represented by a set of spatial and time coordinates).

The spatial separation of Event 2 and Event 1 in (S) is:

$$\Delta x = x_2 - x_1. \quad (5.19)$$

The spatial separation of Event 2 and Event 1 in another frame (S') is:

$$\begin{aligned} \Delta x' &= x'_2 - x'_1 \\ &= \gamma(x_2 - \beta\gamma t_2) - \gamma(x_1 - \beta\gamma t_1) \\ &= \gamma(x_2 - x_1) - \gamma\beta c(t_2 - t_1) \\ &= \gamma(\Delta x - \beta c\Delta t). \end{aligned} \quad (5.20)$$

Similarly, we can work out the time interval in (S') :

$$\begin{aligned} c\Delta t' &= ct'_2 - ct'_1 \\ &= \gamma(ct_2 - \beta x_2) - \gamma(ct_1 - \beta x_1) \\ &= \gamma(c\Delta t - \beta\Delta x). \end{aligned} \quad (5.21)$$

So spatial separations and time intervals transform like positions and time.

5.4.2 Non-relativistic limit

As the ratio v/c approaches 0, or analogously $c \rightarrow \infty$, γ approaches 1:

$$\lim_{c \rightarrow \infty} \gamma = 1. \quad (5.22)$$

At these limits, called “non-relativistic“ limits, the Lorentz-Einstein Equations become:

$$\begin{cases} x' = x - vt \\ y' = y \\ z' = z \\ t' = t \end{cases} \quad (5.23)$$

which are nothing but the Galilean transformation equations that we are more familiar with.

5.4.3 The limiting speed

The ubiquitous γ factor plays a crucial role in coordinate transformations.

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

Intuitively, γ represents how much time and space is stretched or contracted as we move between frames. However, the mathematical form of γ places a limiting value on the speed v in terms of the speed of light c .

Specifically we observe that if $v = c$, then $\gamma \rightarrow \infty$, and if $v > c$ then γ becomes imaginary. Therefore the speed of light c is the limiting speed. Consequently no two frames can travel at c relative to each other. In fact, no physical object (with non-zero mass) can travel at the speed of light or beyond.

5.4.4 The limiting speed - REALLY!

Physics describes sequences of events that are causally related. No matter in which reference frame observations are made, the law of physics (of our world at least) must retain the chronological order of events. The speed of light as the universal speed limit, which is the limiting speed of any signal of any type known or unknown, is placed under this assumption.

Let us imagine another thought experiment in which Event 1 $(ct_1, x_1)_S$ causes Event 2 $(ct_2, x_2)_S$ by sending a signal with speed u_{Signal} through space. The spatial separation between the two Events in (S) is:

$$\Delta x = x_2 - x_1 = u_{Signal} \Delta t. \quad (5.24)$$

In another reference frame (S'):

$$c\Delta t' = \gamma \left(c\Delta t - \frac{v}{c} \Delta x \right) = \gamma c\Delta t \left(1 - \frac{v}{c^2} u_{Signal} \right). \quad (5.25)$$

Since $\Delta t = t_2 - t_1 > 0$ and $\gamma \geq 1$ (and $c > 0$ of course), if $u_{Signal} > c$ then $\Delta t' < 0$. Therefore, if $u_{Signal} > c$ then there exists a frame (S') with $v/c < 1$ in which the order of events 1 and 2 are reversed, violating causality. So if causality should never be violated, then u_{Signal} must not exceed the speed of light c .

5.5 The fundamental effects and consequence of the Lorentz-Einstein transformation

Let us revisit the fundamental effects of special relativity (time dilation, length contraction, and loss of simultaneity) in the context of the Lorentz-Einstein Equations, which can be written in the matrix form as:

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix}.$$

In this section we will re-derive/verify these effects, using the Lorentz-Einstein transformation equations.

5.5.1 Time dilation

Consider a thought experiment in which a clock moves at speed v relative to Bob (rest frame). At time t_1 , the clock is at point x_1 in Bob's frame. We call this Event 1. At time t_2 , the clock is at x_2 . We call this Event 2. Alice also moves at the same speed v relative to Bob. The clock is at rest in Alice's frame.

Let Event 1 (ct_1, x_1) and Event 2 (ct_2, x_2) be given, then in (S), let the spatial separation of these events be:

$$\Delta x = x_2 - x_1 = v\Delta t = v(t_2 - t_1). \quad (5.26)$$

In Alice's frame (S'), the spatial separation $\Delta x'$, according to the Lorentz-Einstein transformation equations, is:

$$\begin{aligned} \Delta x' &= \gamma(\Delta x - \beta c\Delta t) \\ &= \gamma \left(\Delta x - \frac{v}{c} \frac{\Delta x}{v} c \right) \\ &= 0, \end{aligned} \quad (5.27)$$

which makes sense because (again) the clock is at rest in Alice's frame. But what about the time interval between Event 2 and Event 1?

In Alice's frame (S'):

$$\begin{aligned} c\Delta t' &= \gamma(c\Delta t - \beta\Delta x) \\ &= \gamma\left(c\Delta t - \frac{v}{c}v\Delta t\right). \end{aligned} \quad (5.28)$$

Therefore,

$$\begin{aligned} \Delta t' &= \gamma\left(\Delta t - \frac{v^2}{c^2}\Delta t\right) \\ &= \gamma\Delta t\left(1 - \frac{v^2}{c^2}\right) \\ &= \gamma\Delta t\frac{1}{\gamma^2} \\ &= \frac{\Delta t}{\gamma}, \end{aligned} \quad (5.29)$$

which is what “time dilation” means.

5.5.2 Length contraction

Imagine a train at rest in Alice's frame, which is moving at speed v relative to Bob. In Alice's frame, the train has length $l_0 = x'_R - x'_L = \Delta x'$. In Bob's frame, the length of the train is $l' = x_R - x_L = \Delta x$.

According to the Lorentz-Einstein transformations:

$$\Delta x' = \gamma(\Delta x - \beta c\Delta t). \quad (5.30)$$

Since measurements are made simultaneously in Bob's frame, $\Delta t = 0$. Therefore Eq. 5.30 becomes:

$$\Delta x' = \gamma\Delta x. \quad (5.31)$$

Or, equivalently:

$$l_0 = \gamma l', \quad (5.32)$$

which is what “length contraction” means.

We can, of course, derive the same result differently. Imagine if Bob measures the length of the train by measuring the time interval Δt between the front and the back of the train passing through a point x_0 in Bob's frame.

According to Lorentz-Einstein equations, the length of the train in Alice's frame is:

$$\Delta x' = \gamma(\Delta x - \beta c\Delta t) = -l_0 = x'_L - x'_R, \quad (5.33)$$

note the minus sign represents how Bob's approach in making measurements is different than our previous approach.

Since Bob makes measurements at the same location, $\Delta x = 0$. From Eq. 5.33:

$$l_0 = \gamma\beta c\Delta t = \gamma c\left(\frac{v}{c}\right)\left(\frac{l'}{v}\right) = \gamma l', \quad (5.34)$$

which is the same result we obtained earlier.

5.5.3 Relativity of simultaneity

We return to the question: “If two events occur simultaneously in (S) then what is the time interval between (S') ?”

According to the Lorentz-Einstein equations:

$$\begin{aligned}\Delta x' &= \gamma(\Delta x - \beta c \Delta t) \\ c \Delta t' &= \gamma(c \Delta t - \beta \Delta x).\end{aligned}\tag{5.35}$$

$\Delta t = 0$ since the two events occur simultaneously in (S) . From Eq. 5.35 we get the expression for the time interval in (S') :

$$\Delta t' = \frac{-\gamma \beta \Delta x}{c} = \frac{-v \Delta x'}{c^2},\tag{5.36}$$

which is nothing but the “rear-clock ahead” effect.

6 Relativistic Kinematics: Motion of Particles

In this section we study how to describe the motions of particles in special relativity. Specifically, we study how to represent force, velocity, acceleration in space-time coordinates, as well as new kinetic effects that emerge as a result of relativity.

In previous sections, we have been studying how to transform spatial and time coordinates between frames: given $(x(t), y(t), z(t))_S$, we can find $(x'(t'), y'(t'), z'(t'))_{S'}$. Now, the question simply becomes: given $(u_x(t), u_y(t), u_z(t))_S$, can we find an expression for $(u'_{x'}(t'), u'_{y'}(t'), u'_{z'}(t'))_{S'}$?

While there are multiple ways to tackle this problem, one approach we will use here is the “differential approach” in which we convert finite intervals into differentials. What is means is to bring the “finite interval form” of the Lorentz-Einstein equations that we are familiar with:

$$\begin{aligned}c \Delta t' &= \gamma(c \Delta t - \beta \Delta x) \\ \Delta x' &= \gamma(\Delta x - \beta c \Delta t) \\ \Delta y' &= \Delta y \\ \Delta z' &= \Delta z\end{aligned}\tag{6.1}$$

into “differentials form:”

$$\begin{aligned}cdt' &= \gamma(cdt - \beta dx) \\ dx' &= \gamma(dx - \beta cdt) \\ dy' &= dy \\ dz' &= dz.\end{aligned}\tag{6.2}$$

Replacing βc with v , which is the relative velocity between two reference frames, and dx/dt with u_x :

$$\begin{aligned} cdt' &= \gamma (c - vu_x/c) dt \\ dx' &= \gamma(u_x - v)dt \\ dy' &= dy \\ dz' &= dz. \end{aligned} \tag{6.3}$$

6.1 Longitudinal velocity transformation

From Eq. 6.3, we derive the expression for velocities in the longitudinal direction:

$$u'_x = \frac{dx'}{dt'} = \frac{u_x - v}{1 - \frac{vu_x}{c^2}}. \tag{6.4}$$

Inversely,

$$u_x = \frac{dx}{dt} = \frac{u'_x + v}{1 + \frac{vu'_x}{c^2}}. \tag{6.5}$$

6.2 Transverse velocity transformation

From Eq. 6.3, we derive the expression for velocities in the transverse direction:

$$u'_y = \frac{dy'}{dt'} = \frac{1}{\gamma (1 - \frac{vu_x}{c^2})} \frac{dy}{dt} = \frac{u_y/\gamma}{1 - \frac{vu_x}{c^2}}. \tag{6.6}$$

Inversely,

$$u_y = \frac{u'_y/\gamma}{1 + \frac{vu'_x}{c^2}}. \tag{6.7}$$

what's worth noticing here is that the transverse velocity depends on longitudinal velocity - which makes sense because how fast different reference frames move relative to each other determines how much time is stretched/contracted between these frames.

6.3 Relativistic speed transformation

The next question is that given $(u_x(t), u_y(t), u_z(t))_S$ and $(u'_{x'}(t'), u'_{y'}(t'), u'_{z'}(t'))_{S'}$, can we work out a relationship between the speeds $u = \sqrt{u_x^2 + u_y^2 + u_z^2}$ and $u' = \sqrt{u'^2_{x'} + u'^2_{y'} + u'^2_{z'}}$ in different reference frames?

After some careful algebra, we can get to the following expression:

$$\left(1 - \frac{v^2}{c^2}\right) \left(1 - \frac{u^2}{c^2}\right) = \left(1 - \frac{u'^2}{c^2}\right) \left(1 - \frac{u_x v}{c^2}\right)^2. \tag{6.8}$$

Or equivalently,

$$\gamma(u') = \gamma(u)\gamma(v) \left(1 - \frac{u_x v}{c^2}\right). \quad (6.9)$$

And inversely, of course:

$$\gamma(u) = \gamma(u')\gamma(v) \left(1 + \frac{u'_x v}{c^2}\right). \quad (6.10)$$

6.4 Relativistic acceleration transformation

We can use the differential approach to derive the expressions that relate accelerations between frames:

$$a_x = \frac{du_x}{dt},$$

and so on.

After some algebra, we obtain:

$$a_x = \frac{a_x'}{\gamma^3 \left(1 + \frac{u_x' v}{c^2}\right)^3} \quad (6.11)$$

and

$$\begin{aligned} a_y &= \frac{1}{\gamma^2 \left(1 + \frac{u_x' v}{c^2}\right)^2} \left\{ a_y' - a_x' \frac{u_y' v / c^2}{1 + \frac{u_x' v}{c^2}} \right\} \\ a_z &= \frac{1}{\gamma^2 \left(1 + \frac{u_x' v}{c^2}\right)^2} \left\{ a_z' - a_x' \frac{u_z' v / c^2}{1 + \frac{u_x' v}{c^2}} \right\}. \end{aligned} \quad (6.12)$$

Inversely,

$$\begin{aligned} a_x' &= \frac{a_x}{\gamma^3 \left(1 - \frac{u_x v}{c^2}\right)^3} \\ a_y' &= \frac{1}{\gamma^2 \left(1 - \frac{u_x v}{c^2}\right)^2} \left\{ a_y + a_x \frac{u_y v / c^2}{1 - \frac{u_x v}{c^2}} \right\} \\ a_z' &= \frac{1}{\gamma^2 \left(1 - \frac{u_x v}{c^2}\right)^2} \left\{ a_z + a_x \frac{u_z v / c^2}{1 - \frac{u_x v}{c^2}} \right\}. \end{aligned} \quad (6.13)$$

6.5 The Relativistic Doppler effect

The Doppler Effect is fundamentally rooted in relativity - it requires relative motions in order to manifest. For instance, the sound of approaching jet fighters is slightly more high-pitched than stationary jets.

In this section we study how the Doppler Effect is different under different laws of relativity. We will study the Doppler Effect for sound waves, which travel in the non-relativistic regime and can be well-described using Galilean relativity, as well as light wave, which special relativity better describes.

6.5.1 The Doppler effect for sound

Consider a detector and sound source in relative motion. The detector travels at speed v_D , and the sound source at v_S , both relative to us the observer. Let v be the speed of sound, then the frequency picked up by the detector ν and the frequency of the sound at source ν_0 are related by the following expression:

$$\nu = \nu_0 \left(\frac{v - v_D}{v + v_S} \right). \quad (6.14)$$

Already we can see a problem with the Doppler effect in Galilean relativity: the speed of sound is frame-dependent, whereas the speed of light is not (recall the first postulate of relativity). Therefore we can claim that the Galilean Doppler Effect is at best an approximation, and it is no longer a good model in the relativistic limit. Our goal in the next sections is to derive the relativistic version of Eq. 6.14, and hopefully to verify Eq. 6.14 for non-relativistic cases.

6.5.2 Light propagation

i) 1-D

Consider a beam of light traveling along the x -axis. Assume that this propagating wave can be described as simply as a sinusoidal oscillation with angular frequency ω and wavelength λ . Let $E(x, t)$ represents the amplitude of this wave at any given point in spacetime x and t :

$$E = E_0 \cos(kx - \omega t + \phi_0), \quad (6.15)$$

where ϕ_0 is the phase.

ii) 2-D

Let us assume that the wave carries the same mathematical form in 2-dimension, with the only exception that the spatial coordinate is now parameterized:

$$x_R = x \cos \theta + y \sin \theta. \quad (6.16)$$

Let $k = 2\pi/\lambda$ and $\omega = 2\pi c/\lambda$, Eq. 6.15 becomes:

$$\begin{aligned} E(x_R, t) &= E_0 \cos(kx_R - \omega t + \phi_0) \\ &= E_0 \cos \left(\frac{2\pi}{\lambda} (x \cos \theta + y \sin \theta - ct) + \phi_0 \right). \end{aligned} \quad (6.17)$$

Our goal is to answer the following question: given a plane wave $E(x, y, t)$ in frame (S) , how do we express the same plane wave in another reference frame (S') , $E'(x', y', t')$?

6.5.3 Doppler effect for light

We can start by applying the Lorentz-Einstein transformation equations to the x , y , and t coordinates in Eq. 6.17:

$$E(x', y', t') = E_0 \cos \left(\frac{2\pi}{\lambda} [\gamma \cos \theta (x' + \beta ct') + y' \sin \theta - \gamma(ct' + \beta x')] + \phi_0 \right). \quad (6.18)$$

With some tidying-up, we get the expression for a simple plane wave in (S) :

$$E(x', y', t') = E_0 \cos \left(\frac{2\pi}{\lambda} [\gamma(\cos \theta - \beta)x' + y' \sin \theta - \gamma(1 - \beta \cos \theta)ct'] + \phi_0 \right). \quad (6.19)$$

Now, in (S') , the same wave can be written as:

$$E'(x', y', t') = E_0 \cos \left(\frac{2\pi}{\lambda'} [x' \cos \theta' + y' \sin \theta' - ct'] + \phi_0 \right) \quad (6.20)$$

Since Eq. 6.19 and 6.20 describe the same wave, we can equate them variable-by-variable can see what properties we can obtain from doing so.

Equating the “ ct' ” terms, we get:

$$\lambda' = \frac{\lambda}{\gamma(1 - \beta \cos \theta)} = \frac{\lambda \sqrt{1 - \beta^2}}{1 - \beta \cos \theta}. \quad (6.21)$$

Inversely,

$$\lambda = \frac{\lambda' \sqrt{1 - \beta^2}}{1 + \beta \cos \theta'}. \quad (6.22)$$

Writing Eq. 6.21 and 6.22 in terms of wavelengths gives us the “Doppler Equations.”

$$\nu' = \frac{\nu(1 - \beta \cos \theta)}{\sqrt{1 - \beta^2}} \leftrightarrow \nu = \frac{\nu'(1 + \beta \cos \theta')}{\sqrt{1 - \beta^2}} \quad (6.23)$$

Equating the “ y' ” terms, we get:

$$\frac{\sin \theta'}{\sin \theta} = \frac{\lambda'}{\lambda} = \frac{\sqrt{1 - \beta^2}}{1 - \beta \cos \theta}. \quad (6.24)$$

And equating the “ x' ” terms:

$$\cos \theta' = \frac{\cos \theta - \beta}{1 - \beta \cos \theta}. \quad (6.25)$$

Surprisingly, not only do we see changes in the wavelengths (and frequency) of the waves, but we also see a change in θ , which is the angle the propagation makes with the x -axis.

Now, let us consider a special case in which the light travels only along the x -axis ($\theta = 0$) and frame (S') also travels along the x -axis at speed v relative to (S) . From Eq. 6.25:

$$\cos \theta' = \frac{1 - \beta}{1 - \beta} = 1. \quad (6.26)$$

Therefore $\theta' = 0$, which means the angle of the propagation does not change. However, if we look at the wavelength in (S') :

$$\frac{\lambda'}{\lambda} = \sqrt{\frac{1+\beta}{1-\beta}}. \quad (6.27)$$

For approaching light, Eq. 6.27 becomes:

$$\lambda' = \lambda \sqrt{\frac{1-\beta}{1+\beta}}. \quad (6.28)$$

And for receding light:

$$\lambda' = \lambda \sqrt{\frac{1+\beta}{1-\beta}}. \quad (6.29)$$

Note that for the approaching case, the ratio $\sqrt{(1-\beta)/(1+\beta)} \leq 1$. So the light is “blue shifted,” whereas for the receding case the light is “red shifted.”

Let us consider another special case in which light travels only in the y -axis while (S') still travels in the x -axis relative to (S) at speed v . From Eq. 6.25 we easily get:

$$\lambda' = \frac{\lambda}{\gamma} \leftrightarrow \nu' = \gamma\nu \quad (6.30)$$

7 Spacetime

7.1 Minkowski spacetime diagrams and the Lorentz-Einstein transformations

Spacetime diagrams are a useful graphical representation of the relationship between measurements in different frames. While we have touched in spacetime diagrams in Section 4, in this section we will study spacetime diagrams in better detail. Specifically we will study how a moving frame appears on a spacetime diagram, and how we can use spacetime diagrams to solve problems in special relativity.

We will also look at how the properties of the Lorentz-Einstein transformation, for instance the physics of causality, manifest in the geometry of spacetime diagrams.

7.1.1 Spacetime axes, Events, and World-lines

Let us return to Bob’s spacetime diagram, which was introduced in Section 4. Just a quick reminder of how we can “read” a spacetime diagram. Unlike in most kinematic diagrams where position is often plotted against time, in spacetime diagrams the vertical axis represents “time” and the horizontal axis represents

“space.” Trajectories of particles are represented by arrows. For instance, a vertical arrow pointing upwards represents a particle that is at rest in space. In Figure 28, the rear of the train is moving in the $+x$ direction.

In spacetime diagrams, the “time” axis has units of “space.” The slopes of lines are consequently unit-less, and they represent the speed of particles. More vertical paths correspond to slower particles.

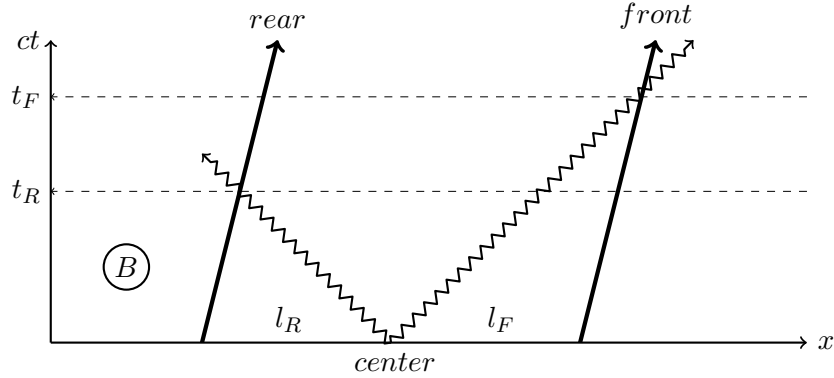


Figure 28: Bob’s spacetime diagram, revisited

We often define the scaling so that light paths (wavy lines) make a 45-degree angle with the horizontal axis. Because no massive physical object can exceed the speed of light, particle paths are always more vertical than light paths.

Mathematically, the speed of the particle relates to the slope of its path by:

$$\beta = \frac{u}{c} = \frac{\Delta x}{c\Delta t} = \frac{\Delta x}{\Delta(ct)} = \frac{1}{\text{slope}}. \quad (7.1)$$

Therefore the slope magnitude of the object’s path in spacetime must always be greater than 1 at any given point.

Minkowski (spacetime) diagram conventionally has a light-line intersecting the origin, as shown in Figure 29:

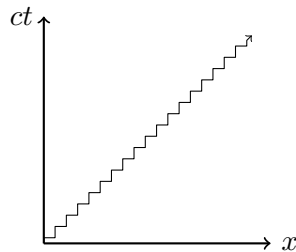


Figure 29: (S') , traveling at speed v relative to (S) , represented in (S) .

7.1.2 Light-cones

Figure 30 shows a 2-dimensional projection of a light-cone onto a spacetime diagram. It is quite straightforward how we can construct a light-cone in a spacetime diagram: crossing two light-lines, one representing a beam in the $+x$ direction and the other in the $-x$, together roughly at their mid-points. However, it is crucial that we understand what light-cones do and their significance in special relativity.

Light-cones are a good graphical representation of how events evolve in space-time. What do we mean by this? Well, if we define the intersection of the light-lines to be the “present,” then the region between the light-lines above the “present” point represents the “future,” and the region below represents the “past.” Why?

As discussed in 7.1.1, because any physical object must never exceed the speed of light, its path in spacetime always has slope with magnitude greater than 1. Consequently, if we let any point on its path be the “present” point, then the rest of the object’s path is confined within two regions defined by a light-cone centered at the chosen “present” point. It makes sense for us to call the upper region “future” and the lower region “past” simply because time flows upwards in a spacetime diagram.

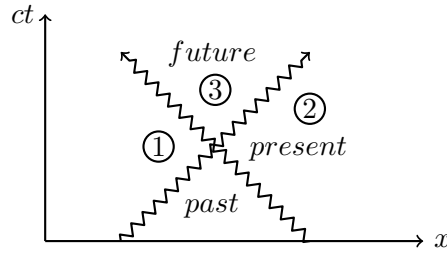


Figure 30: An exemplary light-cone

The division of a spacetime diagram into different “time domains” allows us to categorize events into those that can be causally related and those that cannot. Let us define the first group as “**time-like**” events and the latter as “**space-like**.”

Let Event 1 be the “present” point on our spacetime diagram. Consider Events 1 and 2:

$$\begin{aligned}\Delta x_{12} &= x_2 - x_1 \\ c\Delta t_{12} &= c(t_2 - t_1)\end{aligned}\tag{7.2}$$

We observe that $\Delta x_{12} > c\Delta t_{12}$, which means Events 1 and 2 are too far apart that one event cannot trigger another without breaking the speed of light limit. In other words, Events 1 and 2 cannot be causally related, or they are “space-like separated.” Note that Event 2 is not in the light-cone.

Consider Events 1 and 3:

$$\begin{aligned}\Delta x_{13} &= x_3 - x_1 \\ c\Delta t_{13} &= c(t_3 - t_1)\end{aligned}\tag{7.3}$$

Because $\Delta x_{13} < c\Delta t_{13}$, a subluminal signal can trigger one event after the other. In other words, Events 1 and 3 can be causally related, or they are “time-like separated.” Note that Event 3 is in the light-cone.

7.1.3 The S' axes

Now we will study how to represent another moving frame (S') with axes ct' and x' on a spacetime diagram of a rest frame (S).

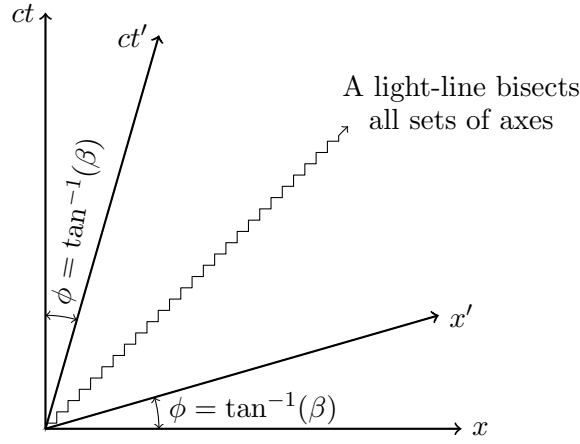


Figure 31: (S'), traveling at speed v relative to (S), represented in (S).

Note that the transformation of coordinates between (S) and (S') aren't orthogonal because lengths (and time intervals) are not *invariant* quantities in special relativity. To state more simply, the spacetime diagram representing (S') is “squeezed” between the axes of (S) and the light-line instead of rotated or reflected about some axis, effectively stretching (and/or compressing) along the light-line. So what (physically) explains this “squeezing” action? Let us answer this question using (1) a physical, more intuitive, reasoning and (2) the Lorentz-Einstein transformation equations. We will see how the latter is a more advantageous approach.

(For readers with background in linear algebra: The “squeezing action” is nothing but a geometrical representation of the Lorentz-Einstein transformation matrix in one spatial dimension.)

i. A physical explanation

Consider an object at $x' = 0$ in (S'), which is traveling at speed v relative to (S). Assuming that the object is at rest in (S'), the line $x' = 0$ (or equivalently ct')

is the object's path in the spacetime of (S') . In our rest frame (S) , the object's path in the spacetime of (S) is tilted an angle ϕ from the vertical (ct) because the object is effectively traveling at speed v in (S) . As discussed in 7.1.1, ϕ characterizes how fast an object travels by:

$$\phi = \tan^{-1}(\beta) = \tan^{-1}\left(\frac{v}{c}\right) \quad (7.4)$$

What about the “up-tilt” of the x' axis from x ? This is where our “physical explanation” gets less intuitive, because in order to explain how an object at $ct' = 0$ (in (S') , of course) can be represented in (S) , we have to take into account two fundamental effects of relativity, namely length contraction and loss of simultaneity. While a careful inspection can lead us to $\phi = \tan^{-1}(\beta)$, the Lorentz-Einstein equations are obviously a better set of tools for tackling this problem.

ii. The Lorentz-Einstein explanation

For the $x' = 0$ case:

$$\begin{aligned} x &= \beta ct \\ \therefore \frac{ct}{x} &= \frac{1}{\beta} = \frac{1}{\text{slope}} \\ \therefore \phi &= \tan^{-1}(\beta). \end{aligned} \quad (7.5)$$

For the $ct' = 0$ case:

$$\begin{aligned} ct' &= 0 \\ \therefore ct &= \beta x \\ \therefore \frac{x}{ct} &= \frac{1}{\beta} = \frac{1}{\text{slope}} \\ \therefore \phi &= \tan^{-1}(\beta). \end{aligned} \quad (7.6)$$

7.1.4 Calibrating Scales

Now we will demonstrate how “lengths” in the spacetime diagram of (S') relate to “lengths” in the spacetime diagram of (S) . Consider the point $(x_0, 0)$ in (S') . In (S') , the distance between this point and the origin is:

$$r' = \sqrt{x_0^2 + 0^2} = x_0 \quad (7.7)$$

The question is “In (S') , what is the distance between this point and the origin?” To answer this we first find the x and ct components of this segment in (S) . The “ x ” component of the line segment:

$$\begin{aligned} x &= \gamma(x' + \beta ct') \\ &= \gamma x' \\ &= \gamma x_0 \end{aligned} \quad (7.8)$$

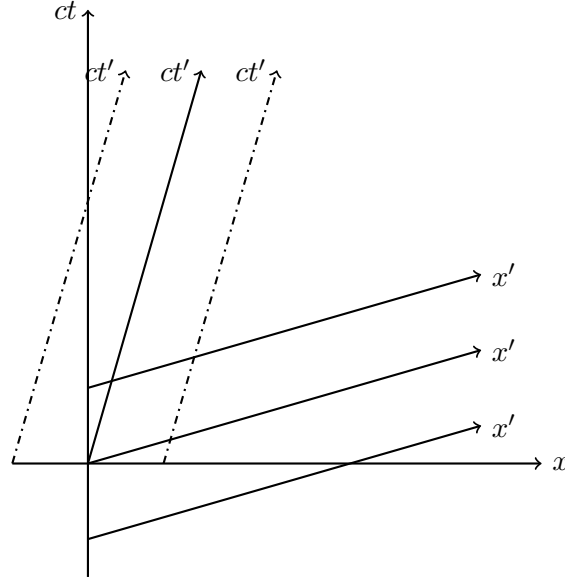


Figure 32: (S') , traveling at speed v relative to (S) , represented in (S) .

The “ ct ” component of the line segment:

$$\begin{aligned} ct &= \gamma(ct' + \beta x') \\ &= \gamma\beta x_0 \end{aligned} \tag{7.9}$$

Hence the length of the line segment in (S) , by Pythagorean theorem, is:

$$\begin{aligned} r &= x_0 \sqrt{\gamma^2 + (\gamma\beta)^2} \\ &= x_0 \sqrt{\gamma^2(1 + \beta^2)} \\ &= x_0 \sqrt{\frac{1 + \beta^2}{1 - \beta^2}} \\ &= r' \sqrt{\frac{1 + \beta^2}{1 - \beta^2}} > 1 \end{aligned} \tag{7.10}$$

We conclude that intervals in (S') spacetime diagrams is stretched by $(1 + \beta^2)/(1 - \beta^2)$ as we move to (S) .

7.1.5 Causality and Minkowski diagrams

In 7.1.2 we briefly discussed “time-like” and “space-like” separated events. In this section we revisit these concepts in a more graphical sense through Minkowski spacetime diagrams.

Consider two pairs of events: $(1,2)$ and $(3,4)$ in (S) .

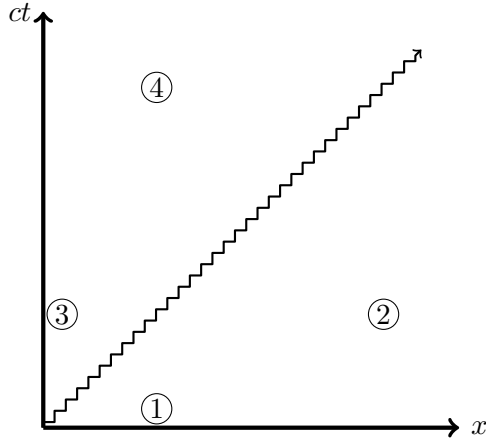


Figure 33: (S') , traveling at speed v relative to (S) , represented in (S) .

Consider the first pair, (1,2). There exists a frame, say (S') in which 1 and 2 are simultaneous. There also exists a frame, say (S'') , in which the chronological order of these two events is reversed:

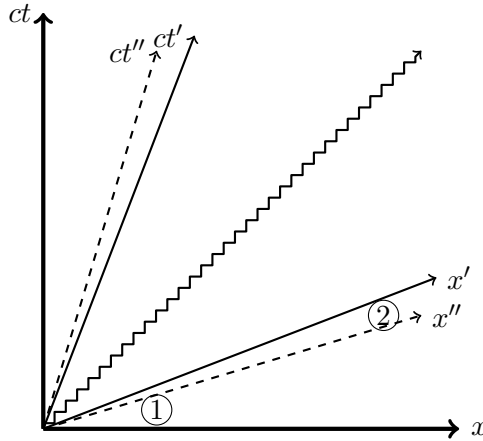


Figure 34: (S') , traveling at speed v relative to (S) , represented in (S) .

Events 1 and 2 cannot be causally related (since their chronological order can be reversed). We say events 1 and 2 are “space-like separated.”

Similarly, consider the second pair of events, (3,4). There exists a frame, say (S') in which 3 and 4 are at the same location. However, the chronological order of events 3 and 4 cannot be reversed.

Events 3 and 4 can be causally related (since their chronological order cannot be reversed). We say events 3 and 4 are “time-like separated.”

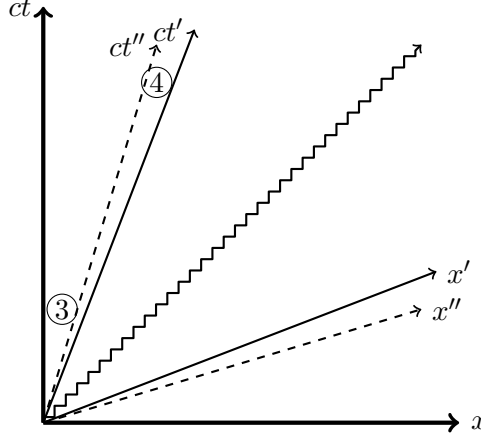


Figure 35: (S') , traveling at speed v relative to (S) , represented in (S) .

7.2 How does Relativity explain Stellar Aberration?

7.3 The Space-Time interval: An *Invariant*

Is there a quantity that uniquely and, in a frame-independent way, identifies the “kind” of separation between events? The answer is yes, and this quantity is called “spacetime interval,” defined mathematically as:

$$(\Delta S)^2 = (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2. \quad (7.11)$$

Note that this quantity is a scalar. We can in fact show, quite quickly, that this quantity is frame-independent. Consider a narrow case with two frames (S) and (S') , where (S') is traveling at speed v in the x -direction relative to (S) . We will show $(\Delta S)^2 = (\Delta S')^2$, using the Lorentz-Einstein transformation equations:

$$\begin{aligned} (\Delta S')^2 &= (c\Delta t')^2 - (\Delta x')^2 - (\Delta y')^2 - (\Delta z')^2 \\ &= \gamma^2(c\Delta t - \beta\Delta x)^2 - \gamma^2(\Delta x - \beta c\Delta t)^2 - (\Delta y')^2 - (\Delta z')^2 \\ &= \gamma^2((c\Delta t)^2 - 2c\Delta t\beta\Delta x + (\beta\Delta x)^2) - \gamma^2((\Delta x)^2 - 2\beta c\Delta t\Delta x + (\beta c\Delta t)^2) \\ &\quad - (\Delta y')^2 - (\Delta z')^2 \\ &= (\Delta t)^2(\gamma^2 c^2 - \gamma^2 v^2) + (\Delta x)^2(\gamma^2 \beta^2 - \gamma^2) - (\Delta y')^2 - (\Delta z')^2 \\ &= (c\Delta t)^2(1 - \beta^2) - (\Delta x)^2(1 - \beta^2)\gamma^2 - (\Delta y')^2 - (\Delta z')^2 \end{aligned} \quad (7.12)$$

Now, since $\gamma^2 = 1/(1 - \beta^2)$, $\Delta y = \Delta y'$, and $\Delta x' = \Delta x$, we obtain the following:

$$(\Delta S')^2 = (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2 = (\Delta S)^2 \quad (7.13)$$

Now we

7.4 The Twin Paradox

7.4.1 Einstein's clock paradox

7.4.2 Elapsed proper time

8 Four-vectors

8.1 Three-vectors

8.1.1 The power of vector notations

8.1.2 The prototype vector

8.1.3 What is a 3-vector?

8.1.4 What is a scalar?

8.1.5 Combination of vectors and scalar?

8.2 Four-vectors

8.2.1 The power of 4-vector notation

8.2.2 The prototype 4-vector

8.2.3 What is a 4-vector?

8.2.4 Combination of 4-vectors and 4-scalars

8.3 Four-velocity

8.3.1 Four-velocity transformation

8.3.2 The meaning of \underline{u}

8.4 Four-acceleration

9 Relativistic Dynamics

9.1 Classical mechanics

hi

9.2 Four-momentum

hi

9.3 Interpretation of four-momentum

hi

9.3.1 Non-relativistic reduction

hi

9.3.2 Interpretation of P_0 and \vec{P}

hi

9.3.3 The equivalence of mass and energy

hi

9.3.4 Relativistic kinetic energy

hi

9.3.5 The energy-momentum invariant

hi

9.3.6 The energy-momentum transformation

hi

9.4 Testing four-momentum conservation

9.4.1 Relativistic billiard problem

i) Newtonian mechanics results

ii) Relativistic results

9.4.2 Collisions and Particle Creation

9.5 Forces and relativistic dynamics

9.5.1 Acceleration, velocity, mass

9.5.2 The four-force

10 The General Theory of Relativity - A brief introduction

- Special Relativity considers observers in inertial frames.
- General Relativity observations in relativity accelerating frames. This is the theory of gravity!

10.1 Why Einstein included gravity in theory of General Relativity

10.1.1 Newton's gravitation

10.1.2 Problems with Newton's gravitation

10.2 The Equivalence Principle

- In a truly falling frame, you do eliminate gravity. Hence, the rules of special relativity hold.
- There is NO observable difference between a real acceleration and gravity.
- The “strong” equivalence principle: In a freely falling frame, all of the law of physics obey the rules of special relativity.

10.2.1 Freely falling light

Light's path in a gravitational field is curved.

10.2.2 Gravitational time dilation: Gravitational Redshift