Vol. 4 – MODERN PHYSICS

The Undergraduate Companion to Theoretical Physics

Andrea Kouta Dagnino[‡]

[‡] *Open University, Milton Keynes, UK.*

E-mail: k.y.dagnino@gmail.com

Contents

The enchanting charms of this sublime science reveal only to those who have the courage to go deeply into it.

— Carl Friedrich Gauss

Part I

Fluid dynamics

Part II

Chaos theory and stochastic dynamics

Part III

Relativity

Basic postulates of Special relativity

1

1.1 Reference frames

What is a frame of reference?

Consider an scaffolding of ruler sticks arranged in space in such a way as to denote every point in space with a set of coordinates (x, y, z), and endowed with a clock keeping track of time (by some physical, periodic phenomenon, such as a fixed number of radiative transitions in a caesium-133 atom).

Such an object is known as a frame of reference, with each space-time point (t, x, y, z), known as **events**, being specified. An inertial frame of reference where an object which is not acted upon by an external force moves at a constant velocity. In other words, it is a frame where Newton's first law holds (thus ruling out accelerating frames of references where fictitious forces are not considered to be external forces).

In classical physics, inertial frames of references satisfy galilean transformations. Consider two frames $\mathcal{S}$ and $\mathcal{S}'$ with coordinates (t, x, y, z) and (t', x', y', z'), with $\mathcal{S}'$ moving with velocity $\mathbf{v} = v_x\mathbf{x} + v_y\mathbf{y} + v_z\mathbf{z}$ as measured in $\mathcal{S}$. Then, the following transformation law is satisfied in galilean relativity:

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v_x & 1 & 0 & 0 \\ -v_y & 0 & 1 & 0 \\ -v_z & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} \tag{1.1.1}$$

It is paramount to note that the time parameter is not affected at all by this transformation, in classical physics all clocks are assumed to be synchronized, even if they are moving relative to each other.

Maxwell vs Newton

This however leads to several contradictions and paradoxical conclusions, especially when put to the test with Maxwell's electromagnetism. For example, consider an electromagnetic wave $\mathbf{E} = \mathbf{E}_0 \sin(\mathbf{k} \cdot \mathbf{x} - \omega t)$ travelling at c as measured in the inertial frame $\mathcal{S}$. In the frame $\mathcal{S}'$, the same wave will be of the form $\mathbf{E}' = \mathbf{E}'_0 \sin(\mathbf{k}' \cdot \mathbf{x}' - \omega' t')$. We now argue that the phase of a plane wave must be an invariant quantity under a change of frame, since everyone must agree on how many crests a wave has undergone in a certain time/distance within their own frame. Consequently, we need

$$\begin{aligned}\mathbf{k}' \cdot \mathbf{x}' - \omega' t' &= \mathbf{k}' \cdot \mathbf{x}' - \omega' t' && (1.1.2)\\ &= \mathbf{k}' \cdot \mathbf{x} - (\mathbf{k}' \cdot \mathbf{v} + \omega')t && (1.1.3)\end{aligned}$$

from which we identify $\mathbf{k}' = \mathbf{k}$ and $\omega = \mathbf{k}' \cdot \mathbf{v} + \omega' t'$. As we let $v \to c$, the observer in $\mathcal{S}'$ will observe a frozen wave with no time-dependence. This clearly isn't a plane wave solution to Maxwell's equations. So are we to believe that Maxwell's equations are only true in a specific frame of reference, the so-called aether?

The Aether

We define the aether as the frame of reference (if it even exists) in which light propagates at the conventional speed of light $c \approx 3 \times 10^8$ m/s.

Consider the following experiment. A person and a mirror are placed on the ends of a platform of length L moving at a speed $v_p \ll c$ relative to the aether. The platform is oriented so that when at rest (relative to the aether), a light beam travelling between its end has speed c. The observer sends a light beam to the mirror, which reflects back and is detected after some time. If the platform is moving along the distance between the observer and the mirror, then this time interval will be:

$$t_1 = \frac{L}{c + v_p} + \frac{L}{c - v_p} \approx \frac{2L}{c}\left(1 + \frac{v_p^2}{c^2}\right) \tag{1.1.4}$$

while if the platform is moving perpendicular to the distance L, then:

$$t_2 = \frac{2L}{\sqrt{c^2 - v_p^2}} \approx \frac{2L}{c}\left(1 + \frac{v_p^2}{2c^2}\right) \tag{1.1.5}$$

There will be a noticeable difference between these time intervals:

$$\Delta t = t_1 - t_2 \approx \frac{L v_p^2}{c^3} \tag{1.1.6}$$

which would cause a beam travelling in the parallel direction to interfere with a beam travelling in the perpendicular direction.

In the Michelson interferometer, a beam splitter is used to split a beam into two travelling in perpendicular directions, and which will interfere according to our above argument when recombining. However, no such interference effects were ever observed.

To explain this shortcoming of Galilean relativity, Lorentz and Fitzgerald argued that the aether could exert some sort of pressure on objects travelling within it, causing a contraction in its direction of motion by a factor γ:

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}} \implies L \to \gamma L \tag{1.1.7}$$

$$\implies t_2 = \frac{2L/c}{1 - v^2/c^2} = t_1 \tag{1.1.8}$$

Such an explanation, although numerically correct, fails to give the proper picture as to why such a contraction should occur. The correct explanation would ultimately arrive with Einstein.

1.2 Fundamental postulates and definitions

Postulates

The basic postulates of special relativity are the following:

(i) **Principle of relativity**: all inertial frames of reference are equivalent, and the laws of physics apply equally.

(ii) **Light speed**: the speed of light in vacuum is c irrespective of its source.

The first postulate is shared with Newtonian physics. A nice way to put it is "if you can juggle at rest, you can also juggle in an IRF", or alternatively "a blind man cannot tell if they are moving in an IRF". The second postulate, on the other hand, is shared with electromagnetism.

The problem of synchronization

We now tackle the question of synchronizing clocks. Suppose an observer sends a light beam at time t_1. It gets reflected by a mirror at an event A and reaches the observer at some time t_2. How do we synchronize the mirror's clock with the observer's clock? If we assume that light travels equally in all directions in vacuum (i.e. space is isotropic) then we can claim that the light beam reached the mirror at $\tau = \frac{1}{2}(t_1 + t_2)$ thus travelling a distance $c\tau = \frac{1}{2}c(t_1 + t_2)$.

Note however that this is just a convention. There is no way to measure the one-way speed of light and hence no way to know exactly when the light beam hit the mirror. Luckily for special relativity, it makes no difference whether or not the one way speed of light is c or some other value. Suppose that for some reason light travels at $c/2$ in the AB direction and instantaneously in the BA direction. An observer is placed at A, and another at B. Their clocks may or may not be synchronized.

At $t_0^A = 0$, the observer at A sends a message to B asking "what does your clock read". The observer at B will receive this message at $t_1^A = \frac{2l}{c}$ in A's clock, and some t_1^B in B's clock. B can respond and instantly and say "t_1^B", which will arrive at $t_2^A = t_1^A$. The observer at A then erroneously changes his clock to $t_3^A = t_1^B + \frac{l}{c}$, thinking that the message must have taken $\frac{l}{c}$ seconds to arrive since it was sent by B. He sends a message saying that his clock now reads $t_1^B + \frac{l}{c}$, arriving at $t_1^B + \frac{2l}{c}$. B then thinks that this makes sense, for A's message must have taken $\frac{l}{c}$ second to arrive.

As can be seen, even though their messages were travelling at different speeds, there were no contradictions in assuming that the one-way speed of light was c. With this convention in mind, then two people can synchronize their clocks by sending a light beam to another observer sitting exactly midway between them.

1.3 Space-time diagrams

An extremely useful tool in special relativity are space-time diagrams. It is common convention to place ct on the z-axis and x, y on the x, y-axes. A trajectory in this is known as the worldline.

We can revisit the problem of synchronization using these space-time diagrams. Consider two frames $\mathcal{S}$ and $\mathcal{S}'$ moving relative to each other at speed v. Three observers, A, B, C are in the frame $\mathcal{S}'$ separated by 1 unit each, and initially set their clocks so that $t = t' = 0$. In the $\mathcal{S}'$ frame, x_A, x_B, x_C's world-lines would satisfy $x' = 0, x' = 1, x' = 2$ respectively. To synchronize their clocks according to Einstein's convention, A and C must send a light

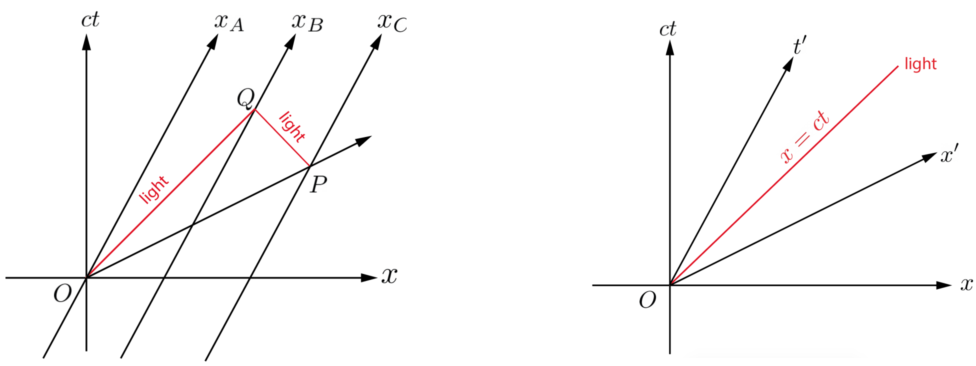

Figure 1.1. Synchronization of clocks

beam to B. If their clocks are synchronized, then B will receive the signals simultaneously, making O and P synchronous in the $\mathcal{S}'$. The point P will thus also be a $t' = 0$ point since it is synchronized with O where $t' = 0$.

To find Q, we solve:

$$ct_Q = vt_Q + 1 \implies t_Q = \frac{1}{c - v} \implies x_Q = \frac{v}{c - v} + 1 \tag{1.3.1}$$

Now QP must have the form $x = c_1 - ct$ where c_1 can be found by imposing that Q lies on the line:

$$\frac{v}{c - v} + 1 = c_1 - \frac{c}{c - v} \implies c_1 = \frac{2c}{c - v} \tag{1.3.2}$$

so that P has coordinates satisfying:

$$\frac{2c}{c - v} - ct_P = vt_P + 2 \implies t_P = \frac{2v}{c^2 - v^2} \implies x_P = \frac{2c^2}{c^2 - v^2} \tag{1.3.3}$$

Consequently, the line OP for which $t' = 0$ must satisfy:

$$ct = \frac{v}{c}x \iff x = \frac{c}{v}ct \tag{1.3.4}$$

We may therefore label the line OP as the x' axis. In the $\mathcal{S}'$ frame we therefore have two tilted axes, which are reflections of each other along $x = ct$.

1.4 Fundamental consequences

Loss of simultaneity

Consider a light bulb on a moving. Observer B is inside the train while observer A is outside, they are moving at a speed v relative to each other. Two receivers are on either side of the light bulb at a distance l, and will activate when hit by a light ray.

In B's frame, the two receivers will clearly activate simultaneously after time $t_1 = t_2 = \frac{l}{c}$. In A's frame, the light from the bulb travels at speed c, but the receivers are also moving with speed v to the right. Consequently, receiver 1 will activate first after time $t_1 = \frac{l}{c+v}$ while the second will activate after time $t_2 = \frac{l}{c-v}$. The two events are not simultaneous for A even though they are for B.

This is a clear example of simultaneity being broken for two inertial observers.

We can view this in the form of a space-time diagram:

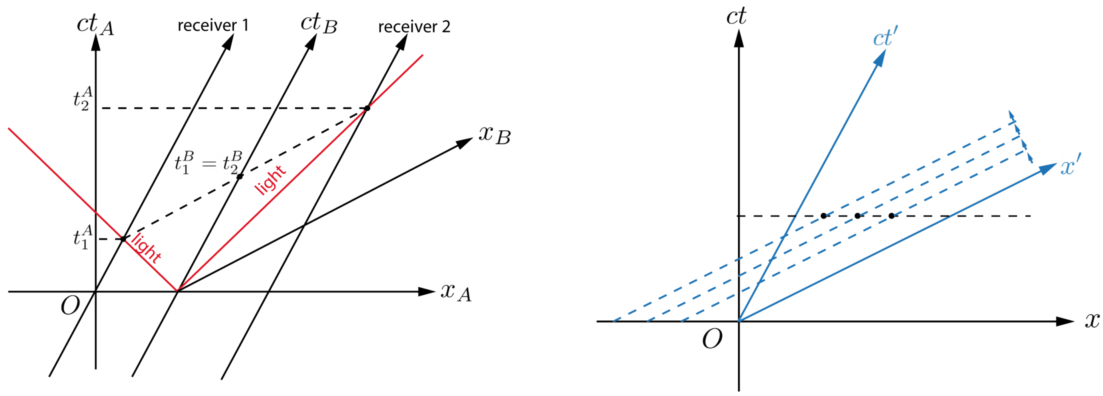

Figure 1.2. Frame dependence of simultaneity

One can also view the loss of simultaneity as a result of the "moving" observer's x'-axis being tilted. Indeed, if we envision a line parallel to the x'-axis moving along the ct'-axis, then clearly three events that are simultaneous in the stationary frame will be crossed at different times in the moving frame.

Time dilation

Consider once again a train containing an observer A moving to with speed v to the right relative to an observer B. The train has a mirror attached to its ceiling at a height h, and the observers have synchronized their clocks at time $t = 0$.

Observer A sends a light beam to the mirror at $t = 0$, in its frame it will see the reflection of the beam at time $t_A = \frac{2h}{c}$.

From observer B's point of view, the light beam has speed c along a diagonal direction, its vertical component will therefore be $\sqrt{c^2 - v^2}$. Consequently, the reflection will be observed

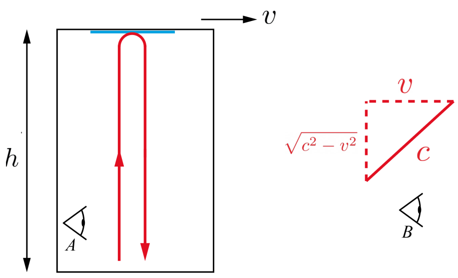

Figure 1.3. Time dilation as a result of loss of simultaneity

at time $t_B = \frac{2h}{\sqrt{c^2-v^2}}$. Hence:

$$\boxed{t_B = \frac{t_A}{\sqrt{1-v^2/c^2}}} \tag{1.4.1}$$

Interestingly, these two times are different, the "moving observer"'s clock will run slowly compared to the "stationary observer".

We can view this more intuitively by looking at the following comic by Tatsu Takeuchi https://www1.phys.vt.edu/~takeuchi/relativity/notes/section12.html:

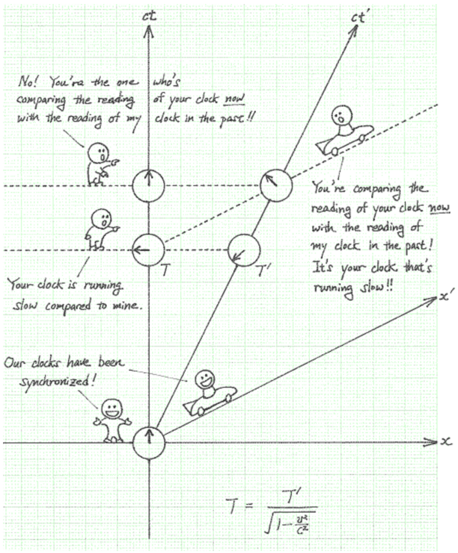

Figure 1.4. Time dilation as a result of loss of simultaneity

Due to the loss of simultaneity between two inertial observers, when they compare their clocks their definitions of simultaneity will cause them to compare their clocks with the other's clock in the past. Hence, the moving observer will always have a clock running more slowly since by the definition of simultaneity the stationary observer is looking at the moving observer's clock in the past.

Length contraction

Observer A stands on one end of a train which they have measured to have length l_A, and sends a light beam to a mirror on the other side. To them the time taken by the light beam is:

$$t_A = \frac{2l_A}{c} \tag{1.4.2}$$

For an observer B on the platform moving with speed v relative to the train, the train has length l_B, and the time taken is:

$$t_B = \frac{l_B}{c-v} + \frac{l_B}{c+v} = \frac{2l_B c}{c^2 - v^2} \tag{1.4.3}$$

since on the first trip of the light beam, the train is trying to move away from it, while on the return trip the train is moving towards it, as shown below: Consequently, using the

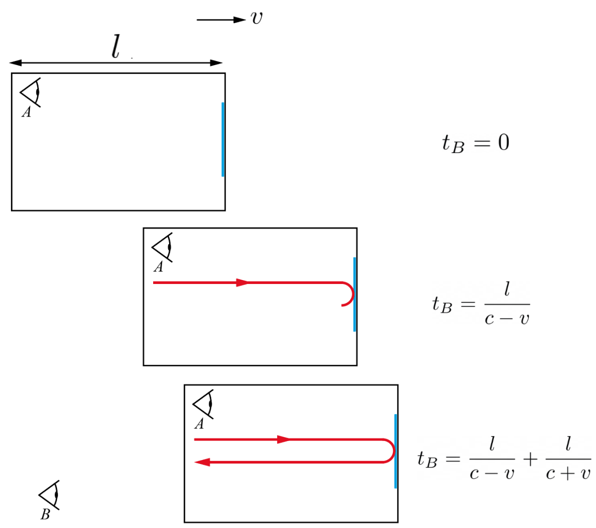

Figure 1.5. Length contraction

time dilation formula we found earlier:

$$t_B = \frac{t_A}{\sqrt{c^2 - v^2}} \implies \boxed{l_B = l_A\sqrt{1 - v^2/c^2}} \tag{1.4.4}$$

Let's consider a rod moving at speed v relative to a frame $\mathcal{S}$. We can express the position of the rod by drawing the world-lines of the front and back end of the rods, as shown below: We center the axes so that the back world-line has equation $x = vt$, while the front world-line has equation $x = vt + l$. In the still frame, the length of the rod is given by the difference in positions of the back and front world-lines at a given time t, which is $QS = l$.

In the moving frame, the length of the rod l' is given by the difference in positions of the back and front world-lines at a given time t'. From the diagram it is clear that this length

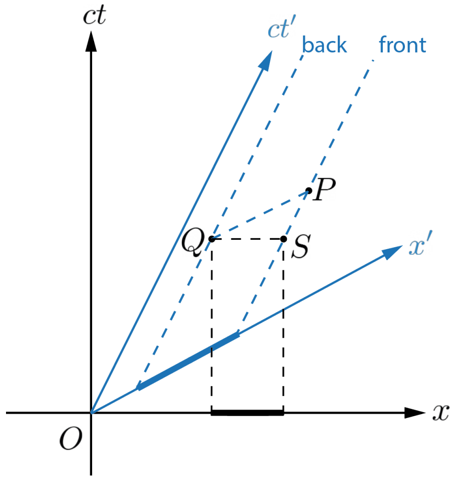

Figure 1.6. Length contraction

is shorter. Indeed:

$$x = \frac{c^2}{v}t = vt + l \implies t = \frac{v}{c^2 - v^2}l \implies x = \frac{c^2}{c^2 - v^2}l \tag{1.4.5}$$

giving a length of:

$$l' = \sqrt{c^2t^2 - x^2} = \frac{l}{\sqrt{1 - v^2/c^2}} \tag{1.4.6}$$

The physical explanation of the minus sign will come later when we encounter the Minkowski metric, but for now let us take it as a postulate.

Interestingly, these two lengths are different, the "moving observer"'s rod will be shorter compared to the "stationary observer".

Lorentz transformations

2

2.1 Derivation

We now seek to find a transformation between two inertial frames $\mathcal{S} : \mathbf{x} = (ct, x, y, z)^T$ and $\mathcal{S} : \mathbf{x}' = (ct', x', y', z')^T$, where $\mathcal{S}'$ moves with velocity $\mathbf{v} = v\hat{\mathbf{e}}_x$ relative to $\mathcal{S}$. We assume that the clocks of these two frames have been synchronized at $t = t' = 0$. Firstly, by the principle of relativity if an object moves with constant velocity in one frame it must move with constant velocity in the other as well. Consequently, the transformation must be a linear one, mapping lines to lines, and keeping the origin fixed. Hence:

$$\mathbf{x}' = \Lambda\mathbf{x},\ \Lambda = \begin{pmatrix} \alpha_1 & \alpha_2 & 0 & 0 \\ \alpha_3 & \alpha_4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.1.1}$$

where the y, z variables are left unchanged from this change of basis. Indeed, if we did have transverse effects, then this would lead to contradictions. For example, if we consider two metal pipes of equal rest diameters D_0 moving towards each other. In pipe 1's frame, pipe 2 has diameter D_2, while of course $D_1 = D_0$ is pipe 1's diameter. If $D_2 > D_0 = D_1$ (transverse length dilation), then this would mean that pipe 1 is inside pipe 2. However from pipe 2's point of view, $D_1 > D_0 = D_2$ so that pipe 2 is inside pipe 1. This is clearly a contradiction. By similar arguments, transverse length contraction is also not feasible, showing that $D_1 = D_2 = D_0$ as desired.

Now the line $x = vt$ must get mapped to $x' = 0$ so that:

$$0 = \alpha_3 ct + \alpha_4 vt \implies \alpha_3 = -\alpha_4 \frac{v}{c} \tag{2.1.2}$$

Similarly, the line $x = 0$ must get mapped to $x' = -vt'$ so that:

$$\begin{cases} -vt' = -\alpha_4 vt \\ t' = \alpha_1 t \end{cases} \implies \alpha_4 = \alpha_1 \tag{2.1.3}$$

Also, by the Light speed postulate, $x = ct$ gets mapped to $x' = ct'$ so that:

$$\begin{cases} x' = ct' = -\alpha_4 vt + \alpha_4 ct \\ ct' = \alpha_4 ct + \alpha_2 ct \end{cases} \implies \alpha_2 = -\alpha_4 \frac{v}{c} = \alpha_3 \tag{2.1.4}$$

Consequently:

$$\Lambda = \alpha_4 \begin{pmatrix} 1 & -\frac{v}{c} & 0 & 0 \\ -\frac{v}{c} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.1.5}$$

Finally, we use the principle of relativity. We know that from the perspective of $\mathcal{S}'$, it is $\mathcal{S}$ that moves with velocity $\mathbf{v} = -v\hat{\mathbf{e}_x}$. Consequently, since $\mathbf{x} = \Lambda^{-1}\mathbf{x}'$, we should have that $\Lambda(v) = \Lambda^{-1}(v)$, and thus:

$$\Lambda^{-1} = \frac{1}{\alpha_4\sqrt{1 - v^2/c^2}} \begin{pmatrix} 1 & \frac{v}{c} & 0 & 0 \\ \frac{v}{c} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \alpha_4 \begin{pmatrix} 1 & \frac{v}{c} & 0 & 0 \\ \frac{v}{c} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.1.6}$$

$$\iff \alpha_4 = \frac{1}{\sqrt{1 - v^2/c^2}} \equiv \gamma(v) \tag{2.1.7}$$

Consequently, the transformation from $\mathcal{S}$ to $\mathcal{S}'$, known as a **Lorentz transformation**, can be written as:

$$\boxed{\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma(v) & -\gamma(v)\frac{v}{c} & 0 & 0 \\ -\gamma(v)\frac{v}{c} & \gamma(v) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}} \tag{2.1.8}$$

or alternatively:

$$t' = \frac{t - \frac{v}{c^2}x}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{2.1.9}$$

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{2.1.10}$$

$$y' = y \tag{2.1.11}$$

$$z' = z \tag{2.1.12}$$

In three dimensions it is easy to see how they generalize to:

$$t' = \gamma_v\left(t - \frac{\mathbf{r}\cdot\mathbf{v}}{c^2}\right) \tag{2.1.13}$$

$$\mathbf{r}'_\parallel = \gamma_v\left(\mathbf{r}_\parallel - \mathbf{v}t\right) \tag{2.1.14}$$

$$\mathbf{r}'_\perp = \mathbf{r}_\perp \tag{2.1.15}$$

2.2 Velocity addition

We know that when velocities are measured in the same frame, they add in the typical Galilean way. However, how do we deal with velocities being measured in different frames?

Longitudinal addition

Suppose we have a frame $\mathcal{S}$ in which an observer A measures another frame $\mathcal{S}'$ moving at speed v to the right. Another observer B is inside $\mathcal{S}'$ and measures the speed of a ball moving to the right to be u. What will the speed w of the ball be in $\mathcal{S}$?

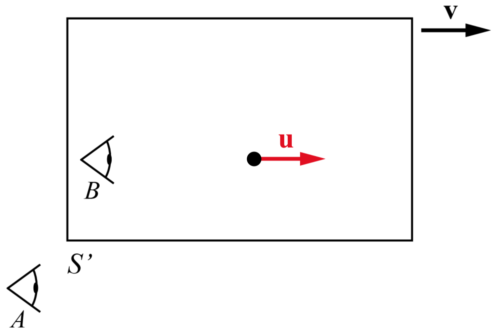

Figure 2.1. Velocity addition

We have that if the ball follows a wordline $(ct, x, 0)$ in frame $\mathcal{S}$ and $(ct', x', 0)$ in $\mathcal{S}'$, then:

$$w = \frac{x}{t} = \frac{x' + vt'}{t' + \frac{v}{c^2}x'} = \frac{u + v}{1 + \frac{uv}{c^2}} \tag{2.2.1}$$

Transverse addition

Suppose now that the ball moves in the transversally in $\mathcal{S}'$.

If the ball follows a wordline $(ct', u_x t', u_y t', u_z t')$ in $\mathcal{S}'$ then in $\mathcal{S}$ it follows a wordline (ct, x, y, z) where:

$$t = \gamma(v)(t' + \frac{u_x v}{c^2}t') \tag{2.2.2}$$
$$x = \gamma(v)(u_x t' + vt') \tag{2.2.3}$$
$$y = u_y t' \tag{2.2.4}$$
$$z = u_z t' \tag{2.2.5}$$

Consequently:

$$w_x = \frac{u_x + v}{1 + \frac{u_x v}{c^2}} \tag{2.2.6}$$
$$w_y = \frac{u_y}{\gamma(v)(1 + \frac{u_x v}{c^2})} \tag{2.2.7}$$
$$w_z = \frac{u_z}{\gamma(v)(1 + \frac{u_x v}{c^2})} \tag{2.2.8}$$

More generally, for a frame $\mathcal{S}'$ moving with velocity $\mathbf{v}$ relative to $\mathcal{S}$, if the ball moves with velocity $\mathbf{u}$ in $\mathcal{S}'$ then $\mathcal{S}$ measures:

$$\boxed{\mathbf{w}_\parallel = \frac{\mathbf{u}_\parallel + \mathbf{v}}{1 + \frac{\mathbf{u}\cdot\mathbf{v}}{c^2}}, \qquad \mathbf{w}_\perp = \frac{\mathbf{u}_\perp}{\gamma(v)(1 + \frac{\mathbf{u}\cdot\mathbf{v}}{c^2})}} \tag{2.2.9}$$

Rapidity

Another way to derive this result is using a quantity known as the rapidity ρ satisfying $\cosh\rho = \gamma$, $\sinh\rho = \gamma\frac{v}{c}$. The Lorentz transformation can now be written in a handy way:

$$\Lambda(\rho) = \begin{pmatrix} \cosh\rho & -\sinh\rho & 0 & 0 \\ -\sinh\rho & \cosh\rho & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.2.10}$$

Due to the additivity of cosh and sinh, the composition of Lorentz transformations is simplified. Suppose in a frame $\mathcal{S}$ we measure a rapidity ρ_1 for frame $\mathcal{S}'$ in which the ball has rapidity ρ_2. Then:

$$\Lambda(\rho_2)\Lambda(\rho_1) = \begin{pmatrix} \cosh\rho_2 & -\sinh\rho_2 & 0 & 0 \\ -\sinh\rho_2 & \cosh\rho_2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cosh\rho_1 & -\sinh\rho_1 & 0 & 0 \\ -\sinh\rho_1 & \cosh\rho_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.2.11}$$

$$= \begin{pmatrix} \cosh(\rho_1+\rho_2) & -\sinh(\rho_1+\rho_2) & 0 & 0 \\ -\sinh(\rho_1+\rho_2) & \cosh(\rho_1+\rho_2) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.2.12}$$

Consequently the rapidity of the ball in the frame $\mathcal{S}$ is $\rho \equiv \rho_1 + \rho_2$ implying that:

$$\tanh\rho = \tanh(\rho_1+\rho_2) = \frac{\tanh\rho_1 + \tanh\rho_2}{1 + \tanh\rho_1 \tanh\rho_2} \tag{2.2.13}$$

and recalling that $\tanh\rho = \frac{w}{c}, \tanh\rho_1 = \frac{v}{c}, \tanh\rho_2 = \frac{u}{c}$ we finally get the velocity addition rule:

$$\boxed{w = \frac{u+v}{1+\frac{uv}{c^2}}} \tag{2.2.14}$$

The ease with which we can combine Lorentz transformations is once again reminiscent of how one can compose rotations in a similar fashion. In the case of typical rotations, the rapidity ρ would be substituted by the

This makes sense, since in a space-time diagram $\tanh\rho$ corresponds to $\tan\theta$ where θ is the angle between the stationary and moving frames' axes.

The use of hyperbolic trigonometric functions allows us to sum angles the way we would conventionally do in euclidean geometry, only that angles now correspond to rapidities (see chapter on spinors for more details).

Rapidities also have a physical interpretation related to classical acceleration. Consider a rocket moving at speed v relative to frame $\mathcal{S}$ and with acceleration a. At time $t + dt$ the rocket is moving with velocity adt relative to its rest frame at time t. Using velocity addition, in the frame $\mathcal{S}$ we have that:

$$v(t+dt) = \frac{v(t) + adt}{1 + v(t)adt/c^2} \approx v(t) + adt - \frac{v(t)^2}{c^2} adt \tag{2.2.15}$$

$$\implies \frac{dv(t)}{dt} = a\left(1 - \frac{v(t)^2}{c^2}\right) \tag{2.2.16}$$

$$\implies \frac{v(t)}{c} = \tanh\left(\frac{1}{c}\int_0^t adt\right) = \tanh\rho \tag{2.2.17}$$

so that:

$$\boxed{\rho = \frac{1}{c}\int_0^t adt \iff \frac{d\rho}{dt} = \frac{a}{c}} \tag{2.2.18}$$

2.3 Lorentz invariance

The quantity $\mathbf{x} = (ct, x, y, z)^T$ is known as a 4-vector, any quantity that transforms as $\mathbf{x}$ under Lorentz boosts, that is through $\mathbf{x}' = \Lambda\mathbf{x}$ is known as a 4-vector. The coordinates of a 4-vector are denoted by a greek script, typically μ or ν running from 0 to 3.

A quantity is said to be Lorentz invariant if it is left unchanged under Lorentz transformation. In Newtonian mechanics, the length of a vector with Euclidean metric is invariant under rotations. This allows us to express the laws of mechanics in a frame-independent way. In a similar way it is useful to find quantities related to 4-vectors that are frame-independent in special relativity.

As one would guess from looking at the, the typical Euclidean length of $\mathbf{x}$ vector is not invariant. Indeed:

$$\mathsf{X}^T\mathsf{X} = (ct)^2 + x^2 + y^2 + z^2 \tag{2.3.1}$$

while:

$$\mathsf{X}'^T\mathsf{X}' = (\Lambda\mathsf{X})^T(\Lambda\mathsf{X}) = \mathsf{X}^T\Lambda^T\Lambda\mathsf{X} = \mathsf{X}^T\Lambda^2\mathsf{X} \tag{2.3.2}$$

where we used the symmetry of Λ. So clearly the notion of length in Euclidean geometry will not do.

Let us impose a metric $g = [\eta_{\mu\nu}]$ such that the norm of a 4-vector in this metric is Lorentz-invariant. In other words, we need the quadratic form:

$$X_\mu X^\mu = \mathsf{X}^T g \mathsf{X} = \eta_{\mu\nu} X^\mu X^\nu \tag{2.3.3}$$

and

$$X'_\mu X'^\mu = \mathsf{X}'^T g \mathsf{X} = \mathsf{X}^T(\Lambda^T g \Lambda)\mathsf{X} = X^a \Lambda^\mu_a \eta_{\mu\nu} \Lambda^\nu_b X^b \tag{2.3.4}$$

to be equal, giving an orthogonality condition:

$$\boxed{\Lambda^T g \Lambda = g \iff \eta_{ab} = \Lambda^\mu_a \eta_{\mu\nu} \Lambda^\nu_b} \tag{2.3.5}$$

Matrices Λ satisfying this condition form the Lorentz group, which are discussed in detail in the Mathematical methods volume. The Lorentz group has a remarkable resemblance with the rotation group $O(3)$, which satisfies a similar orthogonality condition in Euclidean space:

$$R^T \mathbb{1} R = \mathbb{1} \iff \delta_{ab} = R^i_a \delta_{ij} R^j_b \tag{2.3.6}$$

since $\mathbb{1} = [\delta_{ij}]$ is the Euclidean metric.

Going back to the postulate of light speed, we can gain insight into the form of g by imposing that two light-like separated events in one inertial frame be so in all inertial frames. In other words, if say an event with $\mathbf{x} = (ct, x, y, z)$ is light-like separated from the origin in one frame:

$$(ct)^2 - x^2 - y^2 - z^2 = 0 \tag{2.3.7}$$

then similarly:

$$(ct')^2 - x'^2 - y'^2 - z'^2 = 0 \tag{2.3.8}$$

in any other arbitrary primed frame. One should therefore choose a metric of the form:

$$\boxed{g = [\eta_{\mu\nu}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}} \tag{2.3.9}$$

known as the **Minkowski metric** with $(+ - - -)$ signature. It is easy to verify that this metric does indeed satisfy the orthogonality condition (2.3.6).

2.4 Space-time intervals

Given two events (ct_1, x_1, y_1, z_1) and (ct_2, x_2, y_2, z_2), their space-time interval is thus defined as:

$$\boxed{(\Delta s)^2 = \eta_{\mu\nu} \Delta X^\mu \Delta X^\nu = (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2} \tag{2.4.1}$$

The sign of the space-time interval between two events can give insight into their properties:

(i) if $\Delta s > 0$ then the events are **time-like** separated, that is, a physical signal could travel between the two events. It corresponds to the region contained within the light cone. Alternatively, one can find a frame where the two events occur at the same position, but there does not exist a frame where they are simultaneous.

(ii) if $\Delta s < 0$ then the events are **space-like** separated, that is, no physical signal can travel between the two events. It corresponds to the region outside the light cone. Alternatively, one can find a frame where the two events are simultaneous, but there does not exist a frame where they occur at the same position.

(iii) if $\Delta s = 0$, then the events are **light-like** separated, that is, only a light signal can travel between the two events. It corresponds to the surface of the light cone.

As can be seen from the figure below, the surfaces of constant space-time interval form hyperboloids.

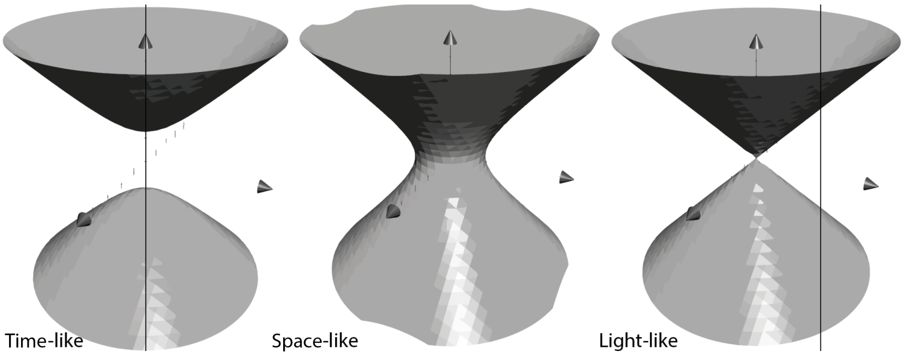

Figure 2.2. Surfaces of constant space-time interval in 2+1 space, with ct on the z-axis, and x, y in the $x - y$ plane.

Using the space-time interval, which is a Lorentz invariant quantity, we may also formally define the concepts of distance and time. For two events that are time-like separated, the distance between them is given by the proper length:

$$\Delta r = -\Delta s \tag{2.4.2}$$

Since we can find a frame $\tilde{\mathcal{S}}$ where the events are simultaneous, we see that Δr is the distance between the events measured simultaneously in $\tilde{\mathcal{S}}$.

For two events that are space-like separated, the time between them is given by the proper time:

$$\Delta \tau = \frac{\Delta s}{c} \tag{2.4.3}$$

2.5 4-vectors

4-velocity

Consider the world-line of a particle moving through space relative to an inertial frame. The differential proper time between any two $(ct, \mathbf{r})$ and $(c(t + dt), \mathbf{r} + d\mathbf{r})$ is:

$$d\tau = \frac{ds}{c} = \frac{1}{c}\sqrt{g_{\mu\nu} dX^\mu dX^\nu} \tag{2.5.1}$$

$$= \frac{1}{c}\sqrt{g_{\mu\nu} \frac{dX^\mu}{dt}\frac{dX^\nu}{dt}} dt \tag{2.5.2}$$

$$= \frac{1}{c}\sqrt{c^2 - v^2} dt \tag{2.5.3}$$

$$= \frac{dt}{\gamma(v)} \tag{2.5.4}$$

where $v = \sqrt{\delta_{ij} \frac{dX^i}{dt} \frac{dX^j}{dt}}$ is the conventional 3-velocity of the particle. This allows us to find the proper time between any two events A and B on this world-line:

$$\Delta\tau = \int_A^B \frac{dt}{\gamma(v)} = \frac{\Delta t}{\gamma(v)} \tag{2.5.5}$$

as we found earlier when discussing time-dilation.

Using proper-time, we can create a 4-velocity whose norm which will be Lorentz invariant:

$$\mathsf{U} = \frac{d\mathsf{X}}{d\tau} = \frac{d}{d\tau}(ct, \mathbf{r}) = \gamma(v) \begin{pmatrix} c \\ \mathbf{v} \end{pmatrix} \tag{2.5.6}$$

Its norm is clearly:

$$||\mathsf{U}|| \equiv \mathsf{U}^T g \mathsf{U} = \gamma(v)\sqrt{c^2 - v^2} = c \tag{2.5.7}$$

which is not only Lorentz-invariant as desired, but also constant.

4-momentum

In Newtonian mechanics, momentum is defined as $\mathbf{p} = m\mathbf{v}$, where m is a Galilean-invariant quantity. Similarly, in Special relativity we can define the 4-momentum using a Lorentz-variant mass, the rest mass m_0, which is defined as the mass of the object as measured in its frame. Hence:

$$\mathsf{P} = m_0 \mathbf{v} = m_0 \gamma(v) \begin{pmatrix} c \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} E/c \\ \mathbf{p} \end{pmatrix} \tag{2.5.8}$$

where we defined:

$$E = \gamma(v) m_0 c^2, \;\; \mathbf{p} = \gamma(v) m_0 \mathbf{v} \tag{2.5.9}$$

to be the relativistic energy and momenta respectively (we shall motivate the definition for the former later).

Its norm is found to be:

$$||\mathsf{P}|| = m_0 \gamma(v) \sqrt{c^2 - v^2} = m_0 c \tag{2.5.10}$$

which is Lorentz invariant as desired. Consequently, we find that:

$$\boxed{E^2 - p^2 c^2 = m^2 c^4} \tag{2.5.11}$$

4-gradient

Note that we can write the transformation law for 4-position as:

$$X'^{\nu} = \Lambda^{\nu}_{\mu} X^{\mu} = \frac{\partial X'^{\nu}}{\partial X^{\mu}} X^{\mu} \tag{2.5.12}$$

$$X'_{\nu} = \Lambda^{\mu}_{\nu} X_{\mu} = \frac{\partial X^{\mu}}{\partial X'^{\nu}} X_{\mu} \tag{2.5.13}$$

which gives us the typical definition of contravariant and covariant vectors. It then follows that:

$$\partial'_\nu \equiv \frac{\partial}{\partial X'^\nu} = \frac{\partial X^\mu}{\partial X'^\nu}\frac{\partial}{\partial X^\mu} = \Lambda^\mu_\nu \partial_\mu \tag{2.5.14}$$

$$\partial'^\nu \equiv \frac{\partial}{\partial X'_\nu} = \frac{\partial X'^\nu}{\partial X^\mu}\frac{\partial}{\partial X_\mu} = \Lambda^\nu_\mu \partial^\mu \tag{2.5.15}$$

Hence, we see that we may define a new 4-operator $\Box$, known as 4-gradient, with contravariant components ∂^μ by differentiating with respect to covariant position components:

$$\partial^\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\nabla\right) \tag{2.5.16}$$

and with covariant components ∂_μ by differentiating with respect to contravariant position components:

$$\partial_\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, \nabla\right) \tag{2.5.17}$$

When we operate on some Lorentz scalar ϕ with the 4-gradient, we get a 4-vector since:

$$\partial'^\nu \phi = \Lambda^\nu_\mu \partial^\mu \phi \tag{2.5.18}$$

If instead we operate on a 4-vector, then:

$$\Box' \cdot \mathsf{V}' = g_{\mu\nu}\partial'^\mu V'^\nu = (\Lambda^\mu_\alpha g_{\mu\nu} \Lambda^\mu_\beta)\partial^\alpha V^\beta = g_{\alpha\beta}\partial^\alpha V^\beta \tag{2.5.19}$$

so we get a Lorentz scalar. For example, $\Box \cdot \mathsf{X} = 4$.

It follows that $\Box = \partial^\mu \partial_\mu$ must be a scalar operator, known as the d'Alembertian operator. It is equivalent to the classical wave operator:

$$\boxed{\Box^2 \equiv \partial^\mu \partial_\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2} \tag{2.5.20}$$

4-wavevector

Let us assume that the phase $\phi = \mathbf{k}\cdot\mathbf{r} - \omega t$ of a plane wave be Lorentz-invariant (this should be case, since all observers should agree on how many cycles a wave has gone through). This is a well motivated choice as we will soon explain. Noting that $\phi = (\frac{\omega}{c}, \mathbf{k}) \cdot (ct, \mathbf{r})$, one would be inclined to define the following quantity:

$$\mathsf{K} = \begin{pmatrix} \frac{\omega}{c} \\ \mathbf{k} \end{pmatrix} \tag{2.5.21}$$

To see that our instincts are justified, consider the following thought experiment. Suppose an observer in some frame measures the number of wave fronts crossing a finite volume in some time interval. The number of crests will be proportional to the measured phase. Now another observer in a frame moving relate to the initial one will still record the same number of crests even though the finite volume and time intervals will be different. Hence the measured phase must be invariant.

Taking the 4-gradient of the phase we obtain a 4-vector known as the 4-wavevector:

$$\boxed{\mathsf{K} = \Box\phi = \begin{pmatrix} \frac{\omega}{c} \\ \mathbf{k} \end{pmatrix}} \tag{2.5.22}$$

The norm of the 4-wavevector is:

$$||\mathsf{K}|| = \frac{\omega^2}{c^2} - k^2 = \omega^2\left(\frac{1}{c^2} - \frac{1}{v_p^2}\right) \tag{2.5.23}$$

where $v_p = \frac{\omega}{k}$ is the phase-speed of a mode ω.

2.6 The Doppler effect

Suppose in frame $\mathcal{S}'$ we have a plane wave moving in the $x'y'$ plane, making an angle θ' with the x' axis, with wave-number k' and angular frequency ω'. Hence we have that:

$$\mathsf{K}' = (\frac{\omega'}{c}, k'\cos\theta', k'\sin\theta', 0) \tag{2.6.1}$$

In the stationary frame $\mathcal{S}$, we have that:

$$\begin{pmatrix} \frac{\omega}{c} \\ k\cos\theta \\ k\sin\theta \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\omega'}{c} \\ k'\cos\theta' \\ k'\sin\theta' \\ 0 \end{pmatrix} \tag{2.6.2}$$

implying that:

$$\omega = \gamma\omega'\left(1 + \frac{v}{\omega'}k'\cos\theta'\right), \qquad \tan\theta = \frac{\sin\theta'}{\gamma(\frac{v\omega'}{k'c^2} + \cos\theta')} \tag{2.6.3}$$

Defining the phase velocity in $\mathcal{S}'$ to be $v_p = \frac{\omega'}{k'}$ then these become:

$$\omega = \gamma\omega'\left(1 + \frac{v}{v_p}\cos\theta'\right) \tag{2.6.4}$$

$$\tan\theta = \frac{\sin\theta'}{\gamma(\cos\theta' + \frac{v_p v}{c^2})} \tag{2.6.5}$$

These equations define the relativistic Doppler effect. There are two special cases of the Doppler effect, the transverse effect where $\cos\theta = 0$, and the longitudinal effect where $\cos\theta' = 1$, both of which can be understood through time dilation and length contraction.

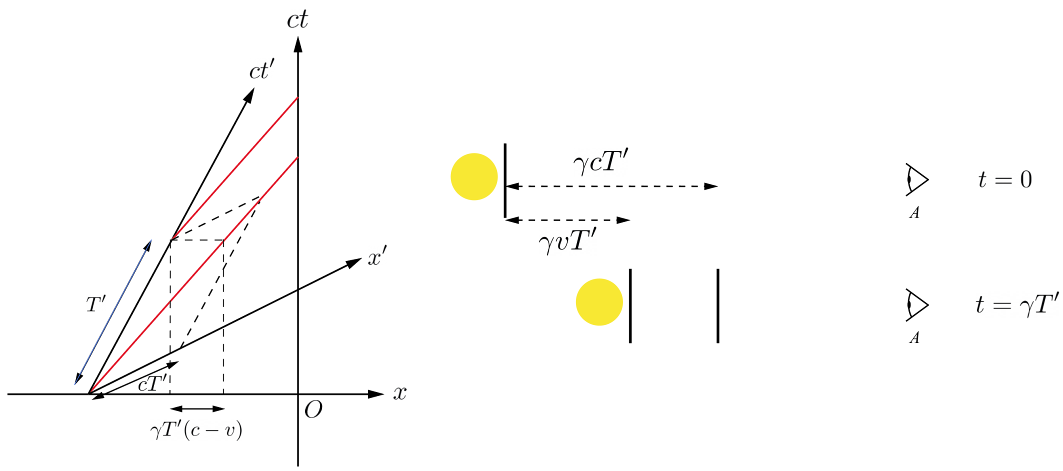

Figure 2.3. Longitudinal Doppler effect

Longitudinal Doppler effect

Here we find that:

$$\frac{\omega}{\omega'} = \sqrt{\frac{1+v/c}{1-v/c}} \tag{2.6.6}$$

We can interpret this as follows. In the source's frame $\mathcal{S}'$, the distance between two crests is $\lambda' = \frac{2\pi}{k'} = cT'$ where $T' = \frac{2\pi}{\omega'}$, so that $T = \frac{2\pi}{k'c}$. In the stationary frame $\mathcal{S}'$, we have that at time $t = 0$, a wave-front is emitted. At $t = T = \gamma T'$, then the second wave-front is emitted, but because the source is moving, the distance between the crests will be $\lambda = \gamma T'(c - v)$. Consequently:

$$k = \frac{2\pi}{\lambda} = \frac{2\pi}{\gamma c T'(1 - v/c)} = \frac{k'}{\gamma T'(1 - v/c)} \tag{2.6.7}$$

$$\implies \frac{\omega}{\omega'} = \frac{k}{k'} = \sqrt{\frac{1+v/c}{1-v/c}} \tag{2.6.8}$$

We can understand this through a helpful space-time diagram shown above.

Transverse Doppler effect

Here we find that $\cos\theta = 0$ and thus $\cos\theta' = -\frac{v_p v}{c^2}$. Consequently:

$$\frac{\omega}{\omega'} = \frac{1}{\gamma} \tag{2.6.9}$$

This follows clearly from applying time dilation, if the wave has period T' in $\mathcal{S}'$ then in $\mathcal{S}$ we have a period $T = \gamma T'$ and thus $\omega' = \gamma\omega \implies \frac{\omega}{\omega'} = \frac{1}{\gamma}$ as desired.

2.7 Thomas precession

Consider the following. In a frame $\mathcal{S}$ we have two squares, one moving upwards with speed u and another moving downwards with speed v. Two of their corners are labelled A and

B as shown. We consider two additional frames: $\mathcal{S}$ and $\mathcal{S}''$ which are the rest frames of

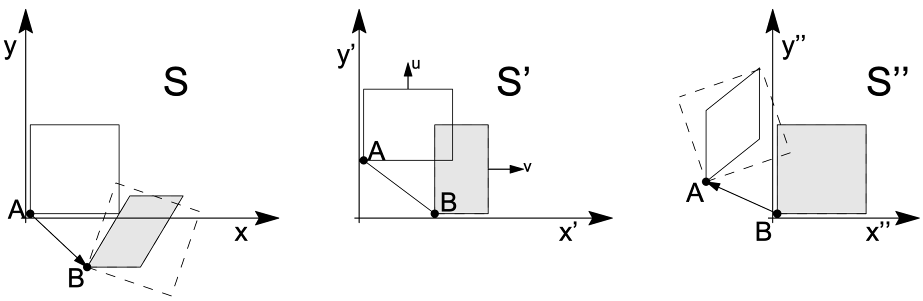

Figure 2.4. A double lorentz boost is equivalent to a single lorentz boost times a rotation. (Have to replace with my own image)

the white and gray squares respectively. We align their frames in $\mathcal{S}'$ along their respective squares.

In frame $\mathcal{S}$ velocity addition tells us that the gray square will be moving with speed $v_{\|} = u, v_{\perp} = \frac{v}{\gamma_u}$. Hence the line AB makes an angle θ with the x-axis satisfying $\tan\theta = \frac{\gamma_u u}{v}$.

Similarly, in frame $\mathcal{S}''$ velocity addition tells us that the white square will be moving with speed $u_{\|} = v, u_{\perp} = \frac{u}{\gamma_v}$. Hence the line AB makes an angle θ'' with the x-axis satisfying $\tan\theta'' = \frac{u}{\gamma_v v}$.

Clearly, these two angles are not the same. In other words, the axes of $\mathcal{S}$ and $\mathcal{S}''$ are misaligned in each other's frames but not in $\mathcal{S}'$!

We may also write that the misalignment $\Delta\theta$ satisfies:

$$\tan\Delta\theta = \frac{\frac{\gamma_u u}{v} - \frac{u}{\gamma_v v}}{1 + \frac{u}{\gamma_v v}\frac{\gamma_u u}{v}} = \frac{uv(\gamma_u\gamma_v - 1)}{\gamma_u u^2 + \gamma_v v^2} \tag{2.7.1}$$

This effect is known as Thomas precession, and the above formula applies even for non-orthogonal velocities. When we perform two successive Lorentz boosts in opposite directions, this will be equivalent to a single Lorentz boost plus an additional rotation by $\Delta\theta$.

Our rapidity statement that Lorentz boosts add up only applied because we were considering boosts in the same direction, for which $\Delta\theta = 0$.

Circular motion

Consider for example a pilot flying a plane along a circle which we model as an N sided regular polygon with internal angles $\theta = \left(1 - \frac{2}{N}\right)\pi$ with N very large. At each vertex, the pilot must therefore rotate by an angle θ', which due to Lorentz contraction satisfies:

$$\tan\theta' = \gamma\tan\theta \implies \theta' \approx \gamma\theta \tag{2.7.2}$$

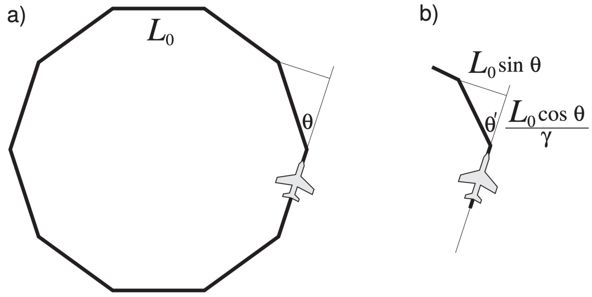

Figure 2.5. A double lorentz boost is equivalent to a single lorentz boost times a rotation. (Have to replace with my own image)

However, this means that after having gone all the way around the polygon, that is, after N rotations, the overall angle the pilot will have rotated by would be $2\pi\gamma > 2\pi$. There has been an extra rotation by $2\pi(\gamma - 1)$! This seemingly paradoxical result is of course be explained through Thomas precession.

Indeed, let us assume a momentary rest frame $\mathcal{S}'$ of the pilot. Here it is moving with velocity $\mathbf{v}$ relative to the rest frame $\mathcal{S}$ of the circle. In time $d\tau$ the pilot will be moving relative to $\mathcal{S}'$ with velocity $d\mathbf{v}_0 = \mathbf{a}_0 d\tau$ where $\mathbf{a}_0$ is the pilot's proper acceleration. Let the new instantaneous frame be $\mathcal{S}''$. It is important to note that $\mathbf{a}_0$ always points towards the center of the circle and is thus perpendicular to $\mathbf{v}_0$. Consequently, to move from time τ to

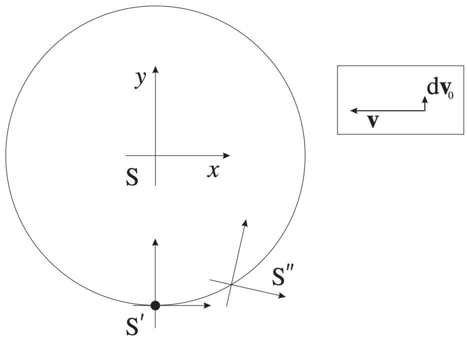

Figure 2.6. A double lorentz boost is equivalent to a single lorentz boost times a rotation. (Have to replace with my own image)

$\tau + d\tau$ we will have to perform a Lorentz boost from $\mathcal{S}$ (circle rest frame) to $\mathcal{S}'$ (pilot rest frame at τ) to $\mathcal{S}''$ (pilot rest frame at τ') along two orthogonal directions, first $\mathbf{v}_0$ and then $d\mathbf{v}$. We have already found the resulting precession angle seen from $\mathcal{S}$:

$$\tan d\theta \approx d\theta = \frac{v dv_0(\gamma_v - 1)}{\gamma_v v^2} = \left(1 - \frac{1}{\gamma_v}\right)\frac{dv_0}{v} \tag{2.7.3}$$

Finally, we substitute $dv_0 = \gamma_v dv$)by velocity addition) to find:

$$d\theta = (\gamma_v - 1)\frac{dv}{v} \implies \Delta\Theta = 2\pi(\gamma_v - 1) \tag{2.7.4}$$

as found earlier.

Relativistic dynamics

3

3.1 4-force

Transformation law

From Newton's second law, we can define the 4-force via the derivative of the 4-momentum as follows:

$$\mathsf{F} = \frac{d\mathsf{P}}{d\tau} = \left(\frac{1}{c}\frac{dE}{d\tau}, \frac{d\mathbf{p}}{d\tau}\right) \tag{3.1.1}$$

Let us define $\mathbf{f} = \frac{d\mathbf{p}}{dt}$ as the 3-force, then we find:

$$\mathsf{F} = \gamma\left(\frac{1}{c}\frac{dE}{dt}, \mathbf{f}\right) \tag{3.1.2}$$

Obviously, an invariant quantity that we can construct is:

$$\mathsf{U} \cdot \mathsf{F} = \gamma^2\left(\frac{dE}{dt} - \mathbf{u} \cdot \mathbf{f}\right) \tag{3.1.3}$$

We can calculate this quantity most easily in the particle's rest frame where $\mathbf{u} = 0$ and $E = mc^2$:

$$\mathsf{U} \cdot \mathsf{F} = \gamma^2 c^2 \frac{dm}{dt} = c^2 \frac{dm}{d\tau} \tag{3.1.4}$$

where we recast the result using invariant quantities. We see that when U and F are orthogonal, the rest mass is constant. Consequently, we get that:

$$\frac{dE}{dt} = \mathbf{u} \cdot \mathbf{f} \tag{3.1.5}$$

Such forces which go solely into changing the kinetic energy of the particle are known as **pure forces**.

Using the Lorentz transformations, it is easy to see that the 4-force transforms according to:

$$\frac{dE'}{dt'} = \frac{\frac{dE}{dt} - vf_{\parallel}}{1 - \mathbf{u}\cdot\mathbf{v}/c^2} \tag{3.1.6}$$

$$f'_{\parallel} = \frac{f_{\parallel} - \frac{v}{c^2}\frac{dE}{dt}}{1 - \mathbf{u}\cdot\mathbf{v}/c^2} \tag{3.1.7}$$

$$f'_{\perp} = \frac{f_{\perp}}{\gamma(v)(1 - \mathbf{u}\cdot\mathbf{v}/c^2)} \tag{3.1.8}$$

As we can see, the 3-force is not invariant at all. Now we have that for a pure 3-force $\mathbf{f}$:

$$\mathbf{f} = \frac{d\mathbf{p}}{dt} = \frac{d}{dt}(\gamma m_0 \mathbf{u}) = \gamma m_0 \mathbf{a} + m_0 \mathbf{u}\frac{d\gamma}{dt} \tag{3.1.9}$$

where $\mathbf{a} = \frac{d\mathbf{u}}{dt}$ is the usual acceleration. After some algebra one finds that:

$$\frac{d\gamma}{dt} = \frac{1}{m_0 c^2}\frac{dE}{dt} = \frac{\mathbf{u}\cdot\mathbf{f}}{m_0 c^2} \tag{3.1.10}$$

$$\mathbf{f} = \gamma m_0 \mathbf{a} + \frac{\mathbf{u}\cdot\mathbf{f}}{c^2}\mathbf{u} \tag{3.1.11}$$

giving the parallel and perpendicular components to $\mathbf{u}$:

$$f_{\parallel} = \gamma m_0 a_{\parallel} + \frac{u^2}{c^2} f_{\parallel} \implies \boxed{f_{\parallel} = \gamma^3 m_0 a_{\parallel}} \tag{3.1.12}$$

and similarly:

$$\boxed{f_{\perp} = \gamma m_0 a_{\perp}} \tag{3.1.13}$$

Clearly, we see that the force acting on the particle is not necessarily parallel to its acceleration. This follows from the fact that the component $\mathbf{p}^{\perp}$ perpendicular to the force cannot change. In other words, we require:

$$p_f^{\perp} = p_i^{\perp} \implies \gamma(v_f)v_f^{\perp} = \gamma(v_i)v_i^{\perp} \tag{3.1.14}$$

so we see that the perpendicular velocity component must change as a result of the $\gamma(v)$ factor changing in the acceleration process.

The great train disaster

A train with rest length L is moving relative towards a bridge with Lorentz factor $\gamma = 3$. The bridge has a rest length of L and is divided into 3 sections of equal rest length.

From the bridge's point of view, the train gets contracted by a factor of 3 so all of the train's weight is acting on just one section, so the bridge breaks and the train falls.

The bridge's architect however states that from the train's point of view the bridge is just 100 meters long so there's no way the train could have fallen. In fact each section only had to support 1/9 the train's weight.

To resolve this paradox let's consider two frames, the rest frame of the bridge $\mathcal{S}$ and the rest frame of the train $\mathcal{S}'$. We note that the a force acting on the each train particle transforms as $f' = \gamma f$ while the weight force acting on each bridge particle transforms as $W' = W/\gamma$.

The breaking force of each section is smaller than $f = nW$ in the bridge frame, where n is the number of particles the train is made up of. The breaking force in the train frame is then smaller than $f' = \gamma nW = \gamma^2 nW'$. In other words, each section can't support 1/9 of the train's rest weight W'.

3.2 Relativistic rockets

Consider a particle accelerating along a line. Suppose that in frame $\mathcal{S}$ the particle is moving with speed v at event A. In a proper time $d\tau$, the particle is now moving at a speed $v(t+d\tau)$ relative to $\mathcal{S}$:

$$v(t+d\tau) = \frac{v(t)+ad\tau}{1+v(t)ad\tau/c^2} \approx v(t) + ad\tau - \frac{v(t)^2}{c^2}ad\tau \tag{3.2.1}$$

$$\implies \frac{dv(t)}{d\tau} = a\left(1 - \frac{v(t)^2}{c^2}\right) \tag{3.2.2}$$

$$\implies \frac{v(t)}{c} = \tanh\left(\frac{1}{c}\int_0^t ad\tau\right) = \tanh\rho \tag{3.2.3}$$

implying that:

$$\frac{d\rho}{d\tau} = \frac{a}{c} \tag{3.2.4}$$

This however only applies to event A and its vicinity, but how do we know that this applies along the particle's entire world-line?

We consider another frame $\mathcal{S}'$ in which $\mathcal{S}$ has rapidity ρ_S, thus obtained through a boost which we take to be along the particle's acceleration. Since rapidities add, we have that the particle's rapidity in $\mathcal{S}'$ is $\rho' = \rho_A + \rho$ and thus:

$$\frac{d\rho'}{d\tau} = \frac{d\rho_S}{d\tau} + \frac{d\rho}{d\tau} = \frac{d\rho}{d\tau} = \frac{a}{c} \tag{3.2.5}$$

since $\mathcal{S}$ is an inertial frame. So, we see that the time evolution of the rapidity is the same in all inertial frames co-linear with the acceleration. Thus the relation

$$\frac{d\rho}{d\tau} = \frac{a}{c} \tag{3.2.6}$$

applies to the particle's entire motion in any inertial frame.

We can apply this to a rocket undergoing constant linear acceleration. Then we have that:

$$\rho(\tau) = \frac{a\tau}{c} + \text{cnst.} \tag{3.2.7}$$

We can set the constant of integration to zero by considering the particle's rest frame at time $\tau = 0$. Then we find that the particle's speed is:

$$v = c\tanh\left(\frac{a\tau}{c}\right) \tag{3.2.8}$$

Next we wish to relate τ to t in $\mathcal{S}$. We have that:

$$\frac{dt}{d\tau} = \gamma = \cosh\left(\frac{a\tau}{c}\right) \implies t = \frac{c}{a}\sinh\left(\frac{a\tau}{c}\right) \tag{3.2.9}$$

assuming clocks t, τ are synchronized at $t = \tau = 0$. Inserting this into (3.2.8) we reach:

$$\boxed{v(t) = \frac{at}{\sqrt{1 + a^2t^2/c^2}}} \tag{3.2.10}$$

Note that as $t \to \pm\infty$, $v \to \pm c$, an uniformly accelerating particle will seem to approach the speed of light in the infinite time limit. Moreover, we see that:

$$\frac{dv(t)}{dt} = \frac{a}{(1 + a^2t^2/c^2)^{3/2}} \tag{3.2.11}$$

so the acceleration in $\mathcal{S}$ approaches zero as $t \to \infty$, while in the particle's instantaneous rest frame the acceleration remains constant at a.

Finally, we may look at the particle's trajectory. We have that:

$$\frac{dx}{d\tau} = \frac{dx}{dt}\frac{dt}{d\tau} = c\sinh\left(\frac{a\tau}{c}\right) \tag{3.2.12}$$

and thus:

$$x = \frac{c^2}{a}\cosh\left(\frac{a\tau}{c}\right) \tag{3.2.13}$$

where we assume that the particle has position $x = 0$ at $t = 0$. Hence

$$x^2 = \left(\frac{c^2}{a}\right)^2(1 + \frac{a^2t^2}{c^2}) \iff \boxed{x^2 - c^2t^2 = \frac{c^4}{a^2}} \tag{3.2.14}$$

The particle undergoes hyperbolic motion.

Note that $ds^2 = x^2 - c^2t^2$ is just the space-time interval between the events $(t = 0, x = 0)$ and (t, x). This suggests that a four-vector formulation of this problem. We have that:

$$\mathsf{X} = \frac{c^2}{a}(\cosh\rho, \sinh\rho) \implies \dot{\mathsf{A}} = \frac{a^2}{c^2}\mathsf{U} \tag{3.2.15}$$

Now, for a particle moving with constant acceleration then:

$$0 = \frac{d}{d\tau}(a^2) = \frac{d}{d\tau}(\mathsf{A}\cdot\mathsf{A}) = 2\mathsf{A}\cdot\dot{\mathsf{A}} \propto \mathsf{A}\cdot\mathsf{U} \tag{3.2.16}$$

so the 4-acceleration and 4-velocity are orthogonal.

3.3 Central forces

In the case of a central force, $\mathbf{f} = f(r)\hat{\mathbf{r}}$, we can define the 3-angular momentum as:

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \tag{3.3.1}$$

As in classical mechanics, angular momentum is conserved:

$$\dot{\mathbf{L}} = \mathbf{v} \times \mathbf{p} + \mathbf{r} \times \mathbf{f} = 0 \tag{3.3.2}$$

Consequently, adopting polar coordinates so that $\mathbf{p} = \gamma m(\dot{r}, r\dot{\phi}) \equiv (p_r, \gamma m r\dot{\phi})$, we find:

$$L = \gamma m r^2 \dot{\phi} \iff \frac{L}{mr^2} = \frac{d\phi}{d\tau} \tag{3.3.3}$$

This relates the angular momentum of a particle in some frame to the derivative of the angular position of the particle with respect to proper time.

Now using the energy-momentum relation with $\mathbf{p} = (p_r, \gamma m r\dot{\phi})$, we find that:

$$p_r^2 = \frac{E^2}{c^2} - \frac{L^2}{r^2} - m^2c^2 \tag{3.3.4}$$

Now define the potential energy due to $\mathbf{f}$ as:

$$V = -\int_{\mathcal{O}}^{\mathbf{r}} \mathbf{f} \cdot d\mathbf{r} \tag{3.3.5}$$

Conservation of energy then requires that:

$$E_{tot} \equiv \gamma mc^2 + V = \text{cnst.} \iff p_r^2c^2 + \frac{c^2L^2}{r^2} + m^2c^4 = (\varepsilon - V)^2 \tag{3.3.6}$$

Now:

$$\frac{dr}{d\tau} = \frac{dr}{dt}\frac{dt}{d\tau} = \frac{p_r}{m} \tag{3.3.7}$$

can be substituted into (3.3.6) to get the radial kinetic energy:

$$\frac{1}{2}m\left(\frac{dr}{d\tau}\right)^2 = \frac{(\varepsilon - V)^2 - m^2c^4 - L^2\frac{c^2}{r^2}}{2mc^2} \tag{3.3.8}$$

$$= \varepsilon_{eff} - V_{eff} \tag{3.3.9}$$

where

$$\varepsilon_{eff} = \frac{\varepsilon^2 - m^2c^4}{2mc^2} \tag{3.3.10}$$

$$V_{eff} = \frac{2\varepsilon V - V^2}{2mc^2} + \frac{L^2}{2mr^2} \tag{3.3.11}$$

For a central potential $V(r) = -\frac{\alpha}{r}$:

$$V_{eff} = \frac{-2\alpha\epsilon/r - \alpha^2/r^2}{2mc^2} + \frac{L^2}{2mr^2} \tag{3.3.12}$$
$$= \frac{1}{2mc^2}\left(\frac{L^2c^2 - \alpha^2}{r^2} - \frac{2\alpha\epsilon}{r}\right) \tag{3.3.13}$$
$$= \frac{1}{2mc^2}\left(\frac{(L^2 - L_c^2)c^2}{r^2} - \frac{2\alpha\epsilon}{r}\right) \tag{3.3.14}$$

where we defined $L_c = \frac{\alpha}{c}$. The first term presents dominates at very small r and can be either attractive or repulsive, while the second gives an attractive potential at large r. In the regime where $L > L_c$ and $\varepsilon_{eff} > 0$, then we have stable bound orbits, and we have that:

$$m\frac{d^2r}{d\tau^2}\frac{dr}{d\tau} = -\frac{dV_{eff}}{d\tau} = -\frac{dV_{eff}}{dr}\frac{dr}{d\tau} \tag{3.3.15}$$
$$\iff m\frac{d^2r}{d\tau^2} = -\frac{dV_{eff}}{dr} \tag{3.3.16}$$
$$\iff \frac{d^2r}{d\tau^2} = \tag{3.3.17}$$

3.4 Energy and momentum relations

We begin by justifying our definitions for the energy $E = \gamma(v)m_0\mathbf{v}$ and momentum $p = \gamma(v)m_0\mathbf{v}$.

We consider a general elastic collision between two identical particles (elastic meaning that the rest masses are left unchanged). We choose a frame F such that the two particles have opposite velocities, and orient our axes so that the x-axis bisects the angle of collision, thus ensuring that P^1 is conserved.

We now consider two frames, one moving along the $-x$ direction, following the right particle, and another moving along the $+x$ direction, following the left particle. Let their relative speed be v.

From the first frame's point of view, the right particle doesn't move along the x-axis, only along the y-axis (say with speed u), while the left particle moves along the x-axis with speed v, as well as along the y-axis (say with speed u'). By symmetry, from the second frame's point of view the speeds are exactly the same, but just with reversed roles.

We propose that there is a quantity $\mathbf{p} = \alpha(v)m_0\mathbf{v}$, known as momentum, is conserved in this collision, and investigate whether or not it exists. In the first frame, we see:

$$2\alpha(u)m_0u = 2\alpha(w)m_0u' \implies \frac{\alpha(w)}{\alpha(u)} = \frac{u}{u'} \tag{3.4.1}$$

Lorenz boosting to the second frame, we get $u' = \frac{u}{\gamma(v)}$ and thus:

$$\alpha(w) = \gamma(v)\alpha(u) \tag{3.4.2}$$

Finally, we have that $w^2 = v^2 + (u')^2 = v^2 + u^2 - u^2v^2/c^2$. Setting $\alpha(v) = \gamma(v)$ in general we see that (3.4.2) is satisfied. Therefore, we should have that:

$$\mathbf{p} = \gamma(v) m_0 \mathbf{v} \tag{3.4.3}$$

We have yet to consider what happens when the collision involves photons which are massless. We begin by using Planck's relations for photons $E = h\nu$ and $p = h\nu/c$. We consider a mass decaying into two photons. In the mass' rest frame, the photons each have frequency ν, while in some frame moving with speed v to the right, the photons have frequencies ν_1 and ν_2 as shown.

If energy and momentum are to be conserved, in the rest frame:

$$E = 2h\nu, \qquad p = 0 \tag{3.4.4}$$

while in the moving frame:

$$E' = h(\nu_1 + \nu_2), \qquad p' = \frac{h}{c}(\nu_2 - \nu_1) \tag{3.4.5}$$

We now use the longitudinal Doppler equation to relate ν_1 and ν_2:

$$\nu_{2,1} = \sqrt{\frac{1 \pm v/c}{1 \mp v/c}}\nu \tag{3.4.6}$$

$$\implies \nu_1 + \nu_2 = 2\gamma\nu, \; \nu_2 - \nu_1 = 2\gamma\frac{v}{c}\nu \tag{3.4.7}$$

Plugging these into (3.4.5) gives:

$$E' = \gamma E, \qquad p' = \gamma E\frac{v}{c^2} \tag{3.4.8}$$

We now resort to the correspondence principle, our result from Special relativity should reproduce Classical results in the limit $\frac{v}{c} \to 0$. Since in classical mechanics we expect $E' - E = \frac{1}{2}m_0v^2$, we should have:

$$E(\gamma - 1) = \frac{1}{2}m_0v^2 \implies E = mc^2 \tag{3.4.9}$$

finally giving the desired relations:

$$\boxed{E = \gamma mc^2, \quad p = \gamma mv} \tag{3.4.10}$$

3.5 Conservation laws

For a system of N particles with 4-momenta P_i, we define the collective total 4-momentum to be:

$$\boxed{\mathsf{P}(t = t_0) = \sum_i \mathsf{P}_i(t = t_0)} \tag{3.5.1}$$

We have to specify the time at which the sum is taken since in general 4-vectors represent different events in different frames. Here t_0 is the time in the frame in which we are measuring the total 4-momentum. By this definition, in a different frame we must have

$$\mathsf{P}(t' = t'_0) = \sum_i \mathsf{P}_i(t' = t'_0) \tag{3.5.2}$$

However, due to the loss of simultaneity, it is not immediate that one can always find a Lorentz boost Λ such that $\mathsf{P}(t' = t'_0) = \Lambda\mathsf{P}(t = t_0)$. Indeed if the particles have different velocities and don't move as a rigid body then in general $\mathsf{P}_i(t' = t'_0) \neq \Lambda\mathsf{P}_i(t = t_0)$, the individual 4-momenta are not transforms of each other.

If we want the total 4-momentum to be an actual 4-vector that transforms accordingly, then we need a new axiom, the conservation of momentum. Let:

$$\mathsf{P}_{AA} = \text{ 4-momentum in frame A at simultaneous times in frame A} \tag{3.5.3}$$
$$\mathsf{P}_{AB} = \text{ 4-momentum in frame A at simultaneous times in frame B} \tag{3.5.4}$$
$$\mathsf{P}_{BB} = \text{ 4-momentum in frame B at simultaneous times in frame B} \tag{3.5.5}$$

If the conservation of momentum is satisfied then we must have $\mathsf{P}_{AA} = \mathsf{P}_{AB}$, and thus

$$\mathsf{P}_{BB} = \Lambda\mathsf{P}_{AB} = \Lambda\mathsf{P}_{AA} \tag{3.5.6}$$

as desired.

$$\boxed{\begin{array}{l}\text{If a sum of 4-vectors evaluated at space-like events} \quad (3.5.7a)\\ \text{is conserved, then this sum is also a 4-vector.} \quad (3.5.7b)\end{array}}$$

It immediately follows that if a 4-vector is conserved in one frame, then it is conserved in all frames.

We prove one final result:

$$\boxed{\begin{array}{l}\text{If one component of a 4-vector is conserved in} \quad (3.5.8a)\\ \text{all frames, then the entire 4-vector is conserved.} \quad (3.5.8b)\end{array}}$$

To begin, note that if a component of a 4-vector is null in zero frames, then the entire 4-vector must be zero. Indeed if one of the spatial components is zero in all frames, then by rotations we see that all spatial components must be zero. If the time component is zero in all frames, but at least one spatial component is not, then we can Lorentz boost along that component to make the time component non-zero, a contradiction. Hence all components of the four-vector must be zero.

Suppose P has a component P^μ that is conserved so that $P^\mu = P'^\mu$. Then letting $\mathsf{Q} = \mathsf{P}' - \mathsf{P}$, and applying the lemma we have proven, we see that $\mathsf{Q} = 0$, and thus the entire 4-vector P is conserved.

3.6 Relativistic collisions

We can now use the tools we have developed on conservation laws to examine a plethora of relativistic collisions.

Radioactive decay/absorption

Suppose a particle of mass M decays into two smaller particles of masses m_1 and m_2. In the rest frame of the initial particle, the four-momentum of M reads $\mathsf{P}_1 = (Mc, 0, 0, 0)$, while for the final two particles it is $\mathsf{P}_2 = (E_1/c, p_1, 0, 0)$ and $\mathsf{P}_3 = (E_2/c, p_2, 0, 0)$. Conservation of 4-momentum implies that:

$$E_1 + E_2 = Mc^2, \qquad p_1 = -p_2 \tag{3.6.1}$$

The energy-momentum equivalence relation also implies that:

$$\begin{aligned} E_1^2 - p_1^2c^2 = m_1^2c^4, &\qquad E_2^2 - p_1^2c^2 = m_2^2c^4 &(3.6.2)\\ \iff (E_1 - E_2)(E_1 + E_2) &= (m_1^2 - m_2^2)c^4 &(3.6.3)\\ \iff E_1 - E_2 &= \frac{m_1^2 - m_2^2}{M}c^2 &(3.6.4)\\ \iff &\boxed{E_1 = \frac{m_1^2 - m_2^2 + M^2}{2M}c^2} &(3.6.5) \end{aligned}$$

Suppose one of the particles is a photon so that $m_1 = 0$. Let $E_0 = Mc^2 - m_2c^2$ be the change in rest mass energy. Then:

$$E_1 = \frac{M^2 - m_2^2}{2M}c^2 = \left(1 - \frac{E_0}{2Mc^2}\right)E_0 \tag{3.6.6}$$

so the energy of the photon is slightly smaller than the rest energy change, with:

$$E_1 - E_0 = -\frac{E_0^2}{2Mc^2} \tag{3.6.7}$$

known as the recoil energy reducing the photon energy. The recoil energy is required to recoil the mass m_2 as required by conservation of momentum.

If instead we have a mass m_2 strike a mass m_1 thus forming a larger mass M, then one can easily find through the same process as the case of emission that:

$$\boxed{E_1 = \frac{-m_1^2 - m_2^2 + M^2}{2M}c^2} \tag{3.6.8}$$

Two-particle decay

Suppose a particle of mass M decays into several smaller particles. We have that:

$$\mathsf{P} = \sum_i \mathsf{P}_i \tag{3.6.9}$$

and thus

$$M^2c^4 = \left(\sum_i E_i\right)^2 - \left(\sum_i \mathbf{p}_i\right) \cdot \left(\sum_i \mathbf{p}_i\right) c^2 \tag{3.6.10}$$

If we only have two decay products then:

$$\mathsf{P} = \mathsf{P}_1 + \mathsf{P}_2 \implies M^2c^2 = m_1^2c^2 + m_2^2c^2 + 2\mathsf{P}_1 \cdot \mathsf{P}_2 \tag{3.6.11}$$

Clearly $\mathsf{P}_1 \cdot \mathsf{P}_2 = \gamma(u)m_1m_2c^2$ (evaluate this product in the rest frame of one of the particles) where u is the relative speed of one decay product relative to the other. Hence:

$$M^2 = m_1^2 + m_2^2 + 2\gamma(u)m_1m_2 \tag{3.6.12}$$

If one is able to measure the outgoing particles' masses and relative speeds, then we can trace back to the original mass.

Threshold energy and the CM frame

Suppose we take a particle of mass m with energy E, momentum $\mathbf{p}$ and collide it with another particle of mass M with the goal of creating new particles.

We can consider this from the center of mass frame where $\mathsf{P}_{CM} = (E_{CM}/c, \mathbf{0})$, while in the laboratory frame $\mathsf{P} = (E/c + Mc, \mathbf{p})$. Thus:

$$E_{CM}^2 = (E + Mc^2)^2 - p^2c^2 = m^2c^4 + M^2c^4 + 2EMc^2 \tag{3.6.13}$$

Our goal is to find the minimum E, known as **threshold energy**, such that the collision may create several particles of total rest mass $\sum_i m_i$. Clearly, this is achieved when all the particles move with momentum p in the lab frame, and thus no momentum in the CM frame. In this case $E_{CM} = \sum_i m_i c^2$ which when substituted into (3.6.13) gives the threshold energy:

$$\boxed{E_{th} = \frac{\left(\sum_i m_i\right)^2 - m^2 - M^2}{2M} c^2} \tag{3.6.14}$$

It is also useful to know what is the relative velocity between the CM frame and lab frame. Suppose we have a system with momentum $\mathbf{p}$ and energy E in the lab frame. WLOG we can align our x-axis with $\mathbf{p}$, and thus Lorentz boost to the CM frame:

$$E_{CM} = \gamma(v)(E - pv), \qquad 0 = \gamma(v)(vE/c^2 - p) \tag{3.6.15}$$

the latter of which gives $v = \frac{pc^2}{E}$ and hence $E_{CM} = \gamma\frac{E^2 - p^2c^2}{E}$.

Three-body decay

We now consider a particle of mass M decaying into three products of masses m_1, m_2, m_3. We have that:

$$\mathsf{P} = \mathsf{P}_1 + \mathsf{P}_2 + \mathsf{P}_3 \tag{3.6.16}$$

Now a useful trick when solving collisions problems is squaring both sides of the momentum conservation law.

$$(\mathsf{P} - \mathsf{P}_3)^2 = (\mathsf{P}_1 + \mathsf{P}_2)^2 \implies M^2c^2 + m_3c^2 - 2\mathsf{P}\cdot\mathsf{P}_3 = m_1^2c^2 + m_2c^2 + 2\mathsf{P}_1\cdot\mathsf{P}_2 \quad (3.6.17)$$

Note that the result is symmetric in m, M reflecting the fact that while in our derivation m was made to collide with M, the opposite picture may also be taken.

Elastic collisions

In an elastic collision the colliding particles do not undergo any change in mass. This alone allows us to derive an interesting result with a classical analogue. Suppose two particles with 4-momenta P and Q collide elastically, outgoing with 4-momenta P' and Q'. Conservation of momentum implies that

$$\mathsf{P} + \mathsf{Q} = \mathsf{P}' + \mathsf{Q}' \quad (3.6.18)$$

$$\iff \mathsf{P}^2 + \mathsf{Q}^2 + 2\mathsf{P}\cdot\mathsf{Q} = \mathsf{P}'^2 + \S Q'^2 + 2\mathsf{P}'\cdot\mathsf{Q}' \quad (3.6.19)$$

$$\iff \boxed{\mathsf{P}\cdot\mathsf{Q} = \mathsf{P}'\cdot\mathsf{Q}'} \quad (3.6.20)$$

Consequently, since $\mathsf{P}\cdot\mathsf{Q} \propto \gamma_u$ where u is the relative velocities of the particles, we see that the particles will have the same relative velocity before and after the collision. Note that the same result holds in classical mechanics.

Consider two identical particles of mass m colliding. We adopt the rest frame of one of the particles and orient our axes so that the x-axis points along the collision line.

We find that before the collision the particles have 4-momenta:

$$\mathsf{P}_1 = (\gamma_u mc, \gamma_u mu, 0, 0) \quad (3.6.21)$$

$$\mathsf{P}_2 = (mc, 0, 0, 0) \quad (3.6.22)$$

while after the collision they are:

$$\mathsf{P}_3 = (\gamma_v mc, \gamma_v mv\cos\theta_1, \gamma_v mv\sin\theta_1, 0) \quad (3.6.23)$$

$$\mathsf{P}_4 = (\gamma_w mc, \gamma_w mw\cos\theta_1, -\gamma_w mw\sin\theta_1, 0) \quad (3.6.24)$$

Conservation of momentum then yields:

$$\gamma_u + 1 = \gamma_v + \gamma_w \quad (3.6.25)$$

$$\gamma_u \mathbf{u} = \gamma_v \mathbf{v} + \gamma_w \mathbf{w} \quad (3.6.26)$$

The second gives:

$$\gamma_u^2 u^2 = \gamma_v^2 v^2 + \gamma_w^w w^2 + 2\gamma_v\gamma_w \mathbf{v}\cdot\mathbf{w} \quad (3.6.27)$$

and substituting the first into the above we find

$$(\gamma_v + \gamma_w - 1)^2 u^2 = \gamma_v^2 v^2 + \gamma_w^w w^2 + 2\gamma_v\gamma_w \mathbf{v}\cdot\mathbf{w} \quad (3.6.28)$$

and using the relation $\gamma_v^2 v^2 = (\gamma_v^2 - 1)c^2$ we find:

$$(\gamma_v + \gamma_w - 1)^2 c^2 - c^2 - \gamma_v^2 v^2 - \gamma_w^w w^2 = 2\gamma_v \gamma_w v w \cos\theta \tag{3.6.29}$$

$$\implies 2c^2(\gamma_v - 1)(\gamma_w - 1) = 2\gamma_v \gamma_w v w \cos\theta \tag{3.6.30}$$

$$\implies \boxed{\cos\theta = \frac{(\gamma_v - 1)(\gamma_w - 1)}{\gamma_v \gamma_w v w} c^2 = \sqrt{\frac{\gamma_v - 1}{\gamma_v + 1}\frac{\gamma_w - 1}{\gamma_w + 1}}} \tag{3.6.31}$$

This gives the angle between the outgoing elastically collided particles. In the low speed limit the particles leave at right angles to each other, and as we increase the speeds θ decreases.

Compton scattering

Tensors and the Lorentz groups

4

5

Covariant electromagnetism

5.1 Remarks on relativistic waves

5.2 The Continuity equation and 4-current

Electric charge is locally conserved, this is expressed using the continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0 \tag{5.2.1}$$

If it were possible to establish $\mathsf{J} = (\rho c, \mathbf{J})$ as a 4-vector, then one could neatly write the continuity equation in a Lorentz covariant form: [1] $\Box \cdot \mathsf{J} \equiv \partial_\mu J^\mu = 0$

Consider two frames $\mathcal{S}$ and $\mathcal{S}'$ moving with relative velocity $\mathbf{u}$. In frame $\mathcal{S}$ a finite region of charge density ρ moves with velocity $\mathbf{v}$ to the right as shown:

Due to the Lorentz invariance of charge, we must have that the same amount of charge must be contained within an infinitesimal volume, so that:

$$\rho d\mathbf{r} = \rho' d\mathbf{r}' \tag{5.2.2}$$

Now letting w be the speed of the charge volume in $\mathcal{S}'$ then clearly $\gamma_w = \gamma_v \gamma_u (1 + \frac{\mathbf{u} \cdot \mathbf{v}}{c^2})$ by velocity-addition. Hence:

$$d\mathbf{r} = \frac{d\mathbf{r}_0}{\gamma_v} \implies d\mathbf{r}' = \frac{\gamma_v}{\gamma_w} d\mathbf{r} = \frac{d\mathbf{r}}{\gamma_u (1 + \mathbf{u} \cdot \mathbf{v}/c^2)} \tag{5.2.3}$$

which gives:

$$\rho' = \gamma_u \left(\rho + \frac{\mathbf{J} \cdot \mathbf{u}}{c^2} \right) \tag{5.2.4}$$

as desired. We now make use of the definition $\mathbf{J} = \rho \mathbf{v}$ and $\mathbf{J}'_\parallel = \rho' \mathbf{w}_\parallel$ to get the transformation of parallel components:

$$\mathbf{J}'_\parallel = \gamma_u \left(\rho + \frac{\mathbf{J} \cdot \mathbf{u}}{c^2} \right) \mathbf{w}_\parallel = \gamma_u \left(\rho + \rho \frac{\mathbf{v} \cdot \mathbf{u}}{c^2} \right) \frac{\mathbf{u} + \mathbf{v}}{1 + \mathbf{u} \cdot \mathbf{v}/c^2} = \gamma_u \rho (\mathbf{u} + \mathbf{v}) \tag{5.2.5}$$

[1] Lorentz covariant means that it makes no reference to frame coordinates, sort of like how Newton's laws in vector form are Galilean covariant as they don't make reference to spatial coordinates

which gives:

$$\mathbf{J}'_{\parallel} = \gamma_u(\mathbf{J} + \rho\mathbf{u}) \tag{5.2.6}$$

Finally,

$$\mathbf{J}'_{\perp} = \gamma_u\left(\rho + \frac{\mathbf{J}\cdot\mathbf{u}}{c^2}\right)\mathbf{w}_{\perp} = \gamma_u\left(\rho + \rho\frac{\mathbf{v}\cdot\mathbf{u}}{c^2}\right)\frac{\mathbf{v}}{\gamma_u(1+\mathbf{u}\cdot\mathbf{v}/c^2)} = \rho\mathbf{v} \tag{5.2.7}$$

which gives:

$$\mathbf{J}'_{\perp} = \mathbf{J}_{\perp} \tag{5.2.8}$$

It follows that $(\rho c, \mathbf{J})$ transforms as a 4-vector which we call the 4-current. We could have also noted that $\mathsf{J} = \rho_0\mathsf{U}$ where ρ_0 is the rest charge density, a Lorentz scalar. The continuity equation takes the form:

$$\boxed{\Box\cdot\mathsf{J} = 0} \tag{5.2.9}$$

5.3 E and B, two sides of the same coin

Our discussion on charges and currents suggest that there is an interplay between charge distributions and current distributions, which themselves produce electric and magnetic fields. As Lorentz transforming charges produce currents and vice versa, one should expect that Lorentz transforming electric fields should produce magnetic fields too.

Consider in some frame $\mathcal{S}$ a neutral wire carrying a current I (made of moving positive charges). If we place a test charge at some radial distance r with initial speed v along the wire, then one would expect the force on it to be a purely magnetic Lorentz force:

$$F_{mag} = -\frac{qv\mu_0 I}{2\pi r} \tag{5.3.1}$$

Let's now boost to the test charge's rest frame $\mathcal{S}'$. Now the positive charge density will be $\rho_+ = \rho$ in $\mathcal{S}$ and hence $\rho'_+ = \gamma_v\rho\left(1 - \frac{uv}{c^2}\right)$ in $\mathcal{S}'$ while the negative charge density will be $\rho_- = -\rho$ in $\mathcal{S}$ and hence $\rho'_- = -\gamma_v\rho$ in $\mathcal{S}'$. The test particle will thus experience no magnetic force but an electrostatic force due to a net charge density $\rho' = \gamma_v\rho\frac{uv}{c^2}$. If the wire has cross-section A then the electric field produced will be:

$$F'_{el} = -\gamma_v\frac{q\rho uvA}{2\pi c^2\varepsilon_0 r} = -\gamma_v\frac{q\mu_0\rho uvA}{2\pi r} \tag{5.3.2}$$

We can transform this form in the original frame to find:

$$F_{el} = -\frac{q\mu_0\rho uvA}{2\pi r} \tag{5.3.3}$$

Recall that if the wire has current I and cross-section A then $I = nAeu = \rho Au$ where n is the charge carrier density and e the electron charge. Therefore the above result may be rewritten as:

$$F_{el} = -\frac{qv\mu_0 I}{2\pi r} = F_{mag} \tag{5.3.4}$$

which is precisely the magnetic force we calculated earlier! In hindsight there was no real need to define a magnetic force, all of this could be calculated using Lorentz contraction and Coulomb's law.

5.4 Gauge invariance

What is a gauge?

We now seek to find a more general law of transformation between the electric and magnetic fields. To do so we must look at the gauge invariance of Maxwell's equations.

$$\nabla \cdot \mathbf{E} = \rho \tag{5.4.1}$$
$$\nabla \cdot \mathbf{B} = 0 \tag{5.4.2}$$
$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{5.4.3}$$
$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} \tag{5.4.4}$$

From the second equation and the Hemholtz decomposition theorem we see that we may write $\mathbf{B} = \nabla \times \mathbf{A}$ where $\mathbf{A}$ is a vector potential. It then follows that:

$$\nabla \times \mathbf{E} = -\nabla \times \frac{\partial \mathbf{A}}{\partial t} \implies \nabla \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}\right) = 0 \tag{5.4.5}$$

which means the electric and magnetic field may be written as functions of the scalar and vector potentials:

$$\boxed{\mathbf{B} = \nabla \times \mathbf{A}, \qquad \mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}} \tag{5.4.6}$$

These equations have a hidden symmetry, known as a Gauge invariance, which follows from the fact that the curl of a gradient is null. Consequently, suppose we perform the transformation $\mathbf{A}' \mapsto \mathbf{A} + \nabla\chi$ for some well-behaved χ:

$$\nabla \times \mathbf{A}' = \nabla \times \mathbf{A} + \nabla \times (\nabla\chi) = \mathbf{E} \tag{5.4.7}$$

We therefore have an infinite family of possible $\mathbf{A}$ for a given $\mathbf{A}$. This is somehow reminiscent of how an indefinite integral has infinitely many possible values due to the fact that the derivative of a constant is zero. We can extend this argument to $\mathbf{E}$:

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} - \frac{\partial(\nabla\chi)}{\partial t} \tag{5.4.8}$$

so if we want this gauge invariance to apply to $\mathbf{E}$ then we need $\phi \mapsto \phi - \frac{\partial\chi}{\partial t}$. With this choice then:

$$\mathbf{E} = -\nabla\phi + \nabla\frac{\partial\chi}{\partial t} - \frac{\partial \mathbf{A}}{\partial t} - \frac{\partial(\nabla\chi)}{\partial t} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} \tag{5.4.9}$$

as desired.

To summarize, our definitions of $\mathbf{E}$ and $\mathbf{B}$ are invariant under gauge transformations:

$$\boxed{\phi \mapsto \phi + \frac{\partial\chi}{\partial t}, \qquad \mathbf{A} \mapsto \mathbf{A} - \nabla\chi} \tag{5.4.10}$$

These transformations can be written more succintly as:

$$(\phi/c, \mathbf{A}) \mapsto (\phi/c - \frac{1}{c}\frac{\partial \chi}{\partial t}, \mathbf{A} + \nabla\chi) \tag{5.4.11}$$

which suggests postulating that $A^\mu = (\phi/c, \mathbf{A})$ is a 4-vector. If this is the case then a gauge transformation can be written as:

$$A^\mu \mapsto A^\mu + \partial^\mu \chi \tag{5.4.12}$$

One very famous gauge that is often used in classical electromagnetism is the Coulomb gauge:

$$\nabla \cdot \mathbf{A} = 0 \tag{5.4.13}$$

With this gauge one obtains the homogeneous wave equations:

$$\nabla^2 \mathbf{A} - \frac{1}{c^2}\frac{\partial \mathbf{A}}{\partial t^2} = 0 \tag{5.4.14}$$

as can be easily verified. Unfortunately this gauge is incompatible with special relativity because it does not treat time and space on equal footing (it is not Lorentz covariant). It would be nice to have a gauge condition that is manifestly covariant.

The Lorentz gauge

With this in mind, we try to formulate Ampere-Maxwell's law using the vector potential:

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} = \mu_0 \mathbf{J} - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} - \frac{1}{c^2}\frac{\partial(\nabla\phi)}{\partial t} \tag{5.4.15}$$

$$\iff \nabla\left(\nabla \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t}\right) + \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla^2 \mathbf{A} = \mu_0 \mathbf{J} \tag{5.4.16}$$

Note that $\nabla \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t} \equiv \partial_\mu A^\mu$. It would be nice to set this equal to zero, so we define a new gauge known as the Lorentz gauge:

$$\boxed{\Box \cdot \mathsf{A} = 0} \tag{5.4.17}$$

Note that this finally shows that A^μ is a 4-vector, since its dot product with the 4-gradient gives a Lorentz scalar.

Also, it is always possible to find a Lorentz gauge for a given $\mathbf{E}, \mathbf{B}$. Indeed, suppose we have some 4-potential A^μ such that $\partial_\mu A^\mu = f$. Then if we perform some gauge transformation $A'^\mu = A^\mu + \partial^\mu \chi$ we find:

$$\partial_m u A'^\mu = \partial_\mu A^\mu + \Box^2 \chi \tag{5.4.18}$$

For this to be zero we require $\Box^2 \chi = -f$. Due to the existence and uniqueness theorem this can always be done so one can always use the Lorentz gauge.

With this in mind we get that:

$$\frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla^2 \mathbf{A} = \mu_0 \mathbf{J} \implies \Box^2 \mathbf{A} = \mu_0 \mathbf{J} \tag{5.4.19}$$

Knowing that $(\phi/c, \mathbf{A})$ and $(\rho c, \mathbf{J})$ are 4-vector we should expect a very similar equation to hold for ρ. We can use Gauss's law to write:

$$\nabla^2\phi + \frac{\partial}{\partial t}(\nabla \cdot \mathbf{A}) = \nabla^2\phi - \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = \rho/\varepsilon_0 \tag{5.4.20}$$

$$\iff \nabla^2(\phi/c) - \frac{1}{c^2}\frac{\partial^2\phi/c}{\partial t^2} = \mu_0\rho c \implies \Box^2\phi = \mu_0\rho \tag{5.4.21}$$

We can combine $\Box^2\mathbf{A} = \mu_0\mathbf{J}$ and $\Box^2\mathbf{A} = \mu_0\rho$ into a single, manifestly covariant equation:

$$\boxed{\Box^2\mathsf{A} = \mu_0\mathsf{J}} \tag{5.4.22}$$

5.5 Making Electromagnetism covariant

The electromagnetic field tensor

With our development of the 4-potential we now seek to write Maxwell's equations in manifestly covariant form. To do so we will need a quantity which encodes both $\mathbf{E}$ and $\mathbf{B}$ and that follows Lorentzian transformation laws.

Clearly this cannot be a 4-vector since we have a total of 6 electromagnetic field components. The next logical step is a 4-tensor $F^{\mu\nu}$ which transforms as:

$$F'^{\mu\nu} = \Lambda^\mu_\alpha\Lambda^\nu_\beta F^{\alpha\beta} \iff \mathbb{F}' = \Lambda\mathbb{F}\Lambda^T \tag{5.5.1}$$

This is easily done by We can define the following rank-2 tensor:

$$\boxed{F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu} \tag{5.5.2}$$

known as the electromagnetic field tensor. One very important property of this tensor is that it is anti-symmetric. Consequently $F^{\mu\mu} = 0$.

Note also that $A^\mu \to A^\mu + \Box\chi$ then:

$$F^{\mu\nu} \to \partial^\mu(A^\nu + \Box\chi) - \partial^\nu(A^\mu + \Box\chi) = F^{\mu\nu} \tag{5.5.3}$$

so the electromagnetic field tensor is gauge invariant as one would require for it to encode information about $\mathbf{E}$ and $\mathbf{B}$.

Now we know that $F^{\mu\nu}$ will definitely include the electric and magnetic fields as we are taking derivatives of the potentials. Indeed:

$$F^{i0} = \frac{1}{c}\partial^i\phi - \frac{1}{c}\partial^0\mathbf{A} = E^i \implies F^{0i} = -E_i/c \tag{5.5.4}$$

Similarly:

$$F^{12} = -\frac{\partial A_y}{\partial x} + \frac{\partial A_x}{\partial y} = -B_3 \tag{5.5.5}$$

We can cycle through the indices and find that $F^{13} = B_2$ and $F^{23} = -B_1$. In general it is easy to see that:

$$B_i = \frac{1}{2}\epsilon_{ijk}F^{jk}, \quad E^i = cF^{i0} \tag{5.5.6}$$

Thus:

$$\boxed{F^{\mu\nu} = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix}} \tag{5.5.7}$$

The Electromagnetic field equations

Immediately we see that:

$$\partial_\mu F^{\mu\nu} = \partial_\mu\partial^\mu A^\nu - \partial_\mu\partial^\nu A^\mu = \Box^2 A^\nu \tag{5.5.8}$$

so using (5.4.22) we find that:

$$\boxed{\partial_\mu F^{\mu\nu} = \mu_0 J^\nu} \tag{5.5.9}$$

Also, we see that due to the antisymmetry of the electromagnetic field tensor the following must also hold:

$$\partial_{[\alpha}F_{\beta\gamma]} \equiv \partial_\alpha F_{\beta\gamma} + \partial_\gamma F_{\alpha\beta} + \partial_\beta F_{\gamma\alpha} = 0 \tag{5.5.10}$$

known as the Bianchi identity. It is easy to see that this reproduces the homogeneous Maxwell equations.

We can write (5.5.10) in another way by introducing the dual electromagnetic field tensor:

$$\tilde{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\alpha\beta}F_{\alpha\beta} \tag{5.5.11}$$

It is then easy to see that due to the anti-symmetry of the Levi-Civita 4-tensor:

$$\partial_\mu \tilde{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\alpha\beta}\partial_\mu F_{\alpha\beta} \tag{5.5.12}$$

$$= \frac{1}{6}\epsilon^{\mu\nu\alpha\beta}(\partial_\mu F_{\alpha\beta} + \partial_\mu F_{\alpha\beta} + \partial_\mu F_{\alpha\beta}) \tag{5.5.13}$$

$$= \frac{1}{6}\epsilon^{\mu\nu\alpha\beta}(\partial_\mu F_{\alpha\beta} + \partial_\beta F_{\mu\alpha} + \partial_\alpha F_{\beta\mu}) \tag{5.5.14}$$

We recognize that the factor in parenthesis must vanish, so we find:

$$\boxed{\partial_\mu \tilde{F}^{\mu\nu} = 0} \tag{5.5.15}$$

Maxwell's equations have thus been reduced to two manifestly covariant equations:

$$\boxed{\partial_\mu F^{\mu\nu} = \mu_0 J^\mu, \quad \partial_\mu \tilde{F}^{\mu\nu} = 0} \tag{5.5.16}$$

5.6 Lorentz transforming the Lorentz force

Manifestly covariant Lorentz force

In classical electromagnetism we define the electric and magnetic fields as vector fields embedded in space which act on a charge q with a Lorentz force:

$$\mathbf{f} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{5.6.1}$$

We can write this as:

$$\begin{aligned} f^i &= q(E^i + \epsilon^{ijk} v_j B_k) && (5.6.2)\\ &= q(cF^{i0} + \epsilon^{ijk} v_k B_i) && (5.6.3)\\ &= q(cF^{i0} + F^{ij} v_j) = qF^{i\mu} U_\mu && (5.6.4) \end{aligned}$$

which suggests writing down more generally that:

$$\boxed{\mathsf{F} = q\mathbb{F} \cdot \mathsf{U} \iff f^\mu = qF^{\mu\nu} U_\nu} \tag{5.6.5}$$

which gives an additional equation:

$$\frac{dE_{en}}{dt} = q\mathbf{v} \cdot \mathbf{E} \tag{5.6.6}$$

where E_{en} is the energy, and not the electric field amplitude. We can make sense of this equation if the Lorentz force is a pure force (which it should be, electromagnetic fields can only accelerate particles), then we see that:

$$\frac{dE}{dt} = \mathbf{v} \cdot \mathbf{f} = \mathbf{v} \cdot \mathbf{E} \tag{5.6.7}$$

E and B transformations

We now use the fact that the electromagnetic field tensor is a tensor to derive the transformation laws of the electric and magnetic fields. We see that:

$$\begin{aligned} E'_x &= F'^{10} = \Lambda^1_\mu \Lambda^0_\nu F^{\mu\nu}\\ &= \Lambda^1_0 \Lambda^0_1 F^{01} + \Lambda^1_1 \Lambda^0_0 F^{10}\\ &= -\beta^2\gamma^2 E_x + \gamma^2 E_x\\ &= E_x \end{aligned}
\qquad
\begin{aligned} E'_y &= F'^{20} = \Lambda^2_\mu \Lambda^0_\nu F^{\mu\nu}\\ &= \Lambda^2_2 \Lambda^0_1 F^{21} + \Lambda^2_2 \Lambda^0_0 F^{20}\\ &= -\gamma\beta c B_z + \gamma E_y\\ &= \gamma(E_y - vB_z) \end{aligned}
\qquad
\begin{aligned} E'_z &= F'^{30} = \Lambda^3_\mu \Lambda^0_\nu F^{\mu\nu}\\ &= \Lambda^3_3 \Lambda^0_1 F^{31} + \Lambda^3_3 \Lambda^0_0 F^{30}\\ &= \gamma\beta c B_y + \gamma E_z\\ &= \gamma(E_z + vB_y) \end{aligned}$$

$$\begin{aligned} B'_x &= F'^{32} = \Lambda^3_\mu \Lambda^2_\nu F^{\mu\nu}\\ &= \Lambda^3_3 \Lambda^2_2 F^{32}\\ &= B_x \end{aligned}
\qquad
\begin{aligned} B'_y &= F'^{13} = \Lambda^1_\mu \Lambda^3_\nu F^{\mu\nu}\\ &= \Lambda^1_1 \Lambda^3_3 F^{13} + \Lambda^1_0 \Lambda^3_3 F^{03}\\ &= \gamma B_y + \beta\gamma E_z/c\\ &= \gamma(B_y + v/c^2 E_z) \end{aligned}
\qquad
\begin{aligned} B'_z &= F'^{21} = \Lambda^2_\mu \Lambda^1_\nu F^{\mu\nu}\\ &= \Lambda^2_2 \Lambda^1_0 F^{20} + \Lambda^2_2 \Lambda^1_1 F^{21}\\ &= -\gamma\beta E_y/c + \gamma B_z\\ &= \gamma(B_z - v/c^2 E_y) \end{aligned}$$

Consequently for boosts along the x-axis:

$$\boxed{\begin{aligned} E'_x &= E_x & B'_x &= B_x \\ E'_y &= \gamma(E_y - vB_z) & B'_y &= \gamma(B_y + v/c^2 E_z) \\ E'_z &= \gamma(E_z + vB_y) & B'_z &= \gamma(B_z - v/c^2 E_y) \end{aligned}}$$

These can be generalized to:

$$\boxed{\begin{aligned} \mathbf{E}'_\parallel &= \mathbf{E}_\parallel & \mathbf{B}'_\parallel &= \mathbf{B}_\parallel \\ \mathbf{E}'_\perp &= \gamma(\mathbf{E}_\perp + \mathbf{v} \times \mathbf{B}) & \mathbf{B}'_\perp &= \gamma(\mathbf{B}_\perp - \mathbf{v} \times \mathbf{E}/c^2) \end{aligned}}$$

As we can see, the electric field in one frame morphs into part of the magnetic field in another frame, thus explaining the phenomenon in 5.3, as well as most of the interactions in the natural world.

6 Electromagnetic radiation

Spinors

7

8 Classical field theory

8.1 Klein-Gordon field

Consider the following Lagrangian density for a set of three scalar real fields ϕ_a, $a = 1, 2, 3$:

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi_a \partial^\mu \phi_a - \frac{1}{2} m^2 \phi_a \phi_a \tag{8.1.1}$$

This lagrangian is invariant under SO(3). Indeed let us consider an infinitesimal rotation by an angle θ about the axis $\hat{\mathbf{n}}$:

$$\mathsf{R}_{\mathbf{n}}(\theta)\phi_a = \phi_a + \theta \epsilon_{abc} n_b \phi_c \tag{8.1.2}$$

The lagrangian after this rotation is given by (we can use the same a, b, c indices as all second order terms will be negligible):

$$\begin{aligned}\mathcal{L}' = &\frac{1}{2}\partial_\mu(\phi_a + \theta\epsilon_{abc} n_b \phi_c)\partial^\mu(\phi_a + \theta\epsilon_{abc} n_b \phi_c) \\ &\qquad - \frac{1}{2} m^2 (\phi_a + \theta\epsilon_{abc} n_b \phi_c)(\phi_a + \theta\epsilon_{abc} n_b \phi_c) \\ = &\mathcal{L} + \frac{1}{2}\theta\epsilon_{abc} n_b \big[(\partial_\mu\phi_c\partial^\mu\phi_a + \partial^\mu\phi_c\partial_\mu\phi_a) - 2m^2\phi_a\phi_c\big] + o(\theta^2)\end{aligned}$$

Now note that:

$$\epsilon_{abc}(\partial_\mu\phi_c\partial^\mu\phi_a + \partial^\mu\phi_c\partial_\mu\phi_a) = \epsilon_{abc}(\partial_\mu\phi_c\partial^\mu\phi_a - \partial^\mu\phi_a\partial_\mu\phi_c) = 0 \tag{8.1.3}$$

and recall that $\boldsymbol{\phi}\cdot(\mathbf{n}\times\boldsymbol{\phi}) = 0 \implies \epsilon_{abc} n_b \phi_c = 0$. Then we find that $\mathcal{L}' = \mathcal{L}$ so SO(3) is indeed a symmetry of this lagrangian.

The equations of motion are given by:

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right) = \frac{\partial\mathcal{L}}{\partial\phi_a} \tag{8.1.4}$$

where:

$$\frac{\partial\mathcal{L}}{\partial\phi_a} = -m^2\phi_a \tag{8.1.5}$$

and

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right) = \partial_\mu\partial^\mu\phi_a = \Box^2\phi_a \tag{8.1.6}$$

so we obtain:

$$\boxed{(\Box^2 + m^2)\phi_a = 0} \tag{8.1.7}$$

known as the Klein-Gordon equation. By Noether's theorem, there must be a conserved current associated to the SO(3) symmetry. It is given by:

$$J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\delta\phi_a - F^\mu \tag{8.1.8}$$

but since $\partial_\mu F^\mu = \delta\mathcal{L} = 0$ we can set $F^\mu = 0$. Then we see that since $\delta\phi_a = \epsilon_{abc}n_b\phi_c$ the conserved current is:

$$J^\mu = \epsilon_{abc}(\partial^\mu \phi_a)n_b\phi_c \tag{8.1.9}$$

giving a conserved charge:

$$Q = \int d^3x J^0 = \int d^3x \epsilon_{abc}\dot{\phi}_a n_b \phi_c \tag{8.1.10}$$

Now we can without loss of generality align our axes so that **n** points along one of the 3-axes, hence $n_b = \delta^b_n$ where $n = 1, 2, 3$. Then we see that we have three individual conserved charges:

$$Q_n = \int d^3x \epsilon_{abc}\dot{\phi}_a \delta^b_n \phi_c = \int d^3x \epsilon_{anc}\dot{\phi}_a\phi_c = -\int d^3x \epsilon_{nac}\dot{\phi}_a\phi_c \tag{8.1.11}$$

We can also check that $\partial_\mu J^\mu$ using the Klein-Gordon equation:

$$\begin{aligned}
\partial_\mu J^\mu &= \partial_\mu(\epsilon_{abc}(\partial^\mu\phi_a)n_b\phi_c) = \epsilon_{abc}n_b(\partial^\mu\phi_a\partial_\mu\phi_c + \phi_c\Box^2\phi_a) && (8.1.12)\\
&= \epsilon_{abc}n_b(\partial^\mu\phi_a\partial_\mu\phi_c - m^2\phi_a\phi_c) && (8.1.13)\\
&= 0 && (8.1.14)
\end{aligned}$$

where we used the fact that $\epsilon_{abc}\partial^\mu\phi_a\partial_\mu\phi_c n_b = g^\mu_\mu\epsilon_{abc}\partial^\mu\phi_a\partial^\mu\phi_c n_b = 0$.

8.2 Electromagnetic field

Consider the following lagrangian:

$$\mathcal{L} = -\frac{1}{4}F_{\alpha\beta}F^{\alpha\beta} - \mu_0 A_\beta J^\beta \tag{8.2.1}$$

Its equation of motion is given by:

$$\begin{aligned}
\frac{\partial\mathcal{L}}{\partial A_\nu} = -\mu_0 J^\nu,\ \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)} &= \partial_\mu\frac{\partial}{\partial(\partial_\mu A_\nu)}\Big(-\frac{1}{2}(\partial_\alpha A_\beta)F^{\alpha\beta}\Big) && (8.2.2)\\
&= -\frac{1}{2}\partial_\mu\Big[F^{\mu\nu} + \partial_\alpha A_\beta\frac{\partial}{\partial(\partial_\mu A_\nu)}(\partial^\alpha A^\beta - \partial^\beta A^\alpha)\Big] && (8.2.3)\\
&= -\frac{1}{2}\partial_\mu[F^{\mu\nu} + \partial_\alpha A_\beta(g^{\alpha\mu}g^{\beta\nu} - g^{\alpha\nu}g^{\beta\mu})] = -\partial_\mu F^{\mu\nu} && (8.2.4)\\
\implies \partial_\mu F^{\mu\nu} &= \mu_0 J^\nu && (8.2.5)
\end{aligned}$$

which reproduces the inhomogeneous Maxwell equations. It follows that (8.2.1) must be the lagrangian for an electromagnetic field.

One important property of the lagrangian is that it is **not** gauge invariant, but transforms quite nicely under gauge transformations which leads to charge conservation. Indeed, consider a general gauge transformation:

$$A_\mu \to A_\mu + \partial_\nu \partial^\nu \chi \tag{8.2.6}$$

The electromagnetic field lagrangian is gauge invariant, since

$$\mathcal{L} \to -\frac{1}{4} F_{\alpha\beta} F^{\alpha\beta} - \mu_0 (A_\beta + \tag{8.2.7}$$

The Principle of Equivalence

9

The Einstein Field Equations

10

Swarzchild's solution and Black holes

11

Part IV

Quantum Field Theory

Part V

Condensed matter physics

Part VI

Atomic physics and quantum optics

Acknowledgments

This is the most common positions for acknowledgments. A macro is available to maintain the same layout and spelling of the heading.

Note added. This is also a good position for notes added after the paper has been written.

Bibliography

[1] Author, *Title*, *J. Abbrev.* **vol** (year) pg.

[2] Author, *Title*, arxiv:1234.5678.

[3] Author, *Title*, Publisher (year).