Joe's Relatively Small Book of Special Relativity

Joseph Minahan ©

This page intentionally left blank

Special Relativity Notes

Preface

In the fall of 2006 I had been asked by several students if I knew a good book to read about tensors. I am sure that they exist, but I could not immediately think of any. So a week before the exam I quickly wrote up 8 pages of notes summarizing what was discussed in class concerning tensors. I called it *Tensors without tears*. The following year I added about 4 more pages explaining many examples of tensors.

The students seemed to like the notes, or at least that is what they claimed in the evaluations. But many students also said that they did not like the book, either because it was too compressed, or it was too expensive, or sometimes because it was too compressed *and* too expensive. In fact, my guess is that less than half the students even bothered to buy the book. For several years myself and others have been looking for a suitable substitute, but so far the books that we have found are too elementary for this particular course. In particular, the course is supposed to cover some topics in electromagnetism, but the usual introductory books on special relativity don't address this subject.

Having given up trying to persuade the students to buy the book, I have decided to bite the bullet and type up some notes relevant for the whole course. However, the notes are intended to complement the book and I still recommend its purchase.

List of changes:

- *September 2011*: Added a new section (8) about the Lagrangian and Hamiltonian formalism for relativistic particles.

1 Introduction

Special relativity is relevant in physics when the speed of an object is less than, but of the same order of magnitude as the speed of light. In this course, we will always refer to this speed as c, where $c \approx 2.99 \times 10^8$ m/sec. In fact, we will almost always approximate this as $c = 3 \times 10^8$ m/sec. If we consider a car traveling on the motorway at a speed of 100 km/hr, Galilean-Newtonian mechanics is more than adequate to describe its motion. On the other hand, for protons circling in the now completed LHC (Large Hadron Collider) at CERN, special relativity will be particularly important.

In this course we will almost always be concerned with classical physics *without* gravity. The modern definition of classical physics means that we ignore quantum effects. Introducing gravity modifies special relativity to *General Relativity*, but unfortunately that is beyond the scope of this course.

Before we see what special relativity is, let us recall how we treat classical physics in the Newtonian-Galilean fashion. Our primary concern is the measurement of positions, times, velocities momenta etc. of particles. One of the main things we need to consider is the position of a particle at a particular time. A time combined with a position will be called an *event*. An event could mean the time and position where an explosion occurred, or the time and position of a particle at some particular point on its trajectory. The trajectory itself is made up of an infinite number of events as the particle travels through time and space. In any case, an event is described by a point in *space-time*. We can express this point by the four numbers

$$(t_0, x_0, y_0, z_0), \tag{1.1}$$

where the first number refers to the time and the other three refer to the spatial coordinates. We will often measure these four quantities for various events, which can occur over a four-dimensional space-time coordinate system

$$(t, x, y, z)\,, \tag{1.2}$$

which we will sometimes write as

$$(t, \vec{x})\,. \tag{1.3}$$

These coordinates are often measured in a lab, so we will sometimes refer to a coordinate system as a *lab*. But the name we will mainly use is *reference frame*, which we will denote in bold-face as **S**. The person making these measurements in a particular reference frame will be referred to as an *observer*. The reference frame for a particular observer is often called the observer's *rest frame*. In fact, we will often refer to the *rest frames* of bodies or particles, that is, the reference frame where that particular body or particle is at rest.

In this class we will be mainly concerned with a particular type of reference frame, called an *inertial frame*. An inertial frame is a reference frame with the following properties:

1. *There is a universal time coordinate that can be synchronized everywhere in the inertial frame.* This means is that at every spatial point in the reference frame we can place a clock and that all the clocks agree with each other.

2. *Euclidean spatial components.* This means that the spatial components satisfy all axioms of Euclidean geometry.

3. *A body with no forces acting on it will travel at constant velocity according to the clocks and measuring sticks in the inertial frame.*

An important part of this course is determining what different observers measure. In particular, different observers can be in different reference frames. Let us suppose that you are on the side of a motorway in reference frame **S**. Another observer who is passing by in a Volvo traveling at velocity $\vec{v}$ is in a different reference frame $\mathbf{S}'$. The frame $\mathbf{S}'$ has a different set of coordinates (t', x', y', z') and uses these coordinates to measure the space-time positions of events. An important question is how do the coordinates in $\mathbf{S}'$ relate to those in **S**? For example suppose that as the car is passing the observer on the side of the road with velocity $\vec{v}$ it also has a constant acceleration $\vec{a}$. What we have previously learned in physics is that the times in **S** and $\mathbf{S}'$ are the same, that is

$$t' = t\,. \tag{1.4}$$

The positions are related by

$$\vec{x} = \vec{x}\,' + \frac{1}{2}\,\vec{a}\,t^2 + \vec{v}\,t + \vec{x}_0, \tag{1.5}$$

where $\vec{x}_0$ is independent of time.

Let us consider the special case where $\vec{a} = 0$, in other words, the two frames are moving at a constant velocity with respect to each other. In this case the transformation between the two sets of coordinates is known as a *Galilean transformation*. It also means that if **S** is an inertial frame then $\mathbf{S}'$ is also an inertial frame. It will often be the case that the times and positions that we measure are between two space-time points. The *displacement* between these two points $(t_1, \vec{x}_1)$ and $(t_2, \vec{x}_2)$ is given by

$$(\Delta t, \Delta\vec{x}) = (t_2 - t_1, \vec{x}_2 - \vec{x}_1)\,. \tag{1.6}$$

The displacement in $\mathbf{S}'$, is then given by the Galilean transformation

$$(\Delta t', \Delta\vec{x}\,') = (\Delta t, \Delta\vec{x} - \vec{v}\,\Delta t)\,, \tag{1.7}$$

and the *inverse transformation*, that is the transformation that takes us back to the coordinates in **S** is

$$(\Delta t, \Delta\vec{x}) = (\Delta t', \Delta\vec{x}\,' + \vec{v}\,\Delta t')\,. \tag{1.8}$$

One advantage of considering the displacement between two events as opposed to the space-time position of one event is that the constant vector $\vec{x}_0$ drops out.

Let us now consider the velocity of a body measured by two observers in different reference frames. Again suppose that the Volvo is moving with constant velocity $\vec{v}$. Hence the velocity of frame $\mathbf{S}'$ with respect to[1] **S** is $\vec{v}$. Let us further suppose that the

[1] The expression *with respect to* will be used all the time, so to save space we will often abbreviate this as *wrt*.

body inside the Volvo is moving with a velocity $\vec{u}\,'$ as measured by an observer in the car. If $\vec{u}\,'$ is constant, then in order to measure the body's velocity the observer needs to determine two space-time points (events) in the body's trajectory. Let us assume that the displacement between these two events is $(\Delta t', \Delta\vec{x}\,')$. The observer in $\mathbf{S}'$ then computes the velocity of the body to be

$$\vec{u}\,' = \frac{\Delta\vec{x}\,'}{\Delta t'}\,. \tag{1.9}$$

The observer in frame $\mathbf{S}$ (the one at the side of the road) would measure a different velocity $\vec{u}$, which is found by using the inverse Galilean transformation in (1.8). Hence the velocity this observer measures is

$$\vec{u} = \frac{\Delta\vec{x}}{\Delta t} = \frac{\Delta\vec{x}\,' + \vec{v}\,\Delta t'}{\Delta t'} = \vec{u}\,' + \vec{v}\,. \tag{1.10}$$

Hence we see that to find the velocity in $\mathbf{S}$ all we do is add $\vec{v}$ to the velocity measured in $\mathbf{S}'$. This *addition of velocities* appears to be almost trivial so why are we worrying about it? The reason is that this rule for the addition of velocities is violated in nature!

1.1 The Michelson-Morley experiment

Let's assume that the Galilean addition of velocity rule is the correct one. This means that if an observer in $\mathbf{S}'$ measures $\vec{u}\,'$ for a velocity, then an observer in $\mathbf{S}$ measures $\vec{u} = \vec{u}\,' + \vec{v}$. One velocity we could choose to measure is that of a wave, say of a sound wave, a water wave, or even a light wave.

We know that sound waves need a medium to travel in (for example, air), and water waves need water to travel in, so naturally it had been assumed that light waves would also need a medium to travel in. This medium was referred to as the *ether*. If the ether exists, then it must be everywhere in the universe because we are able to see galaxies many billions of light years away from us. Since the earth is traveling around the sun, we can also assume that the earth has some velocity through the ether. It might be a coincidence that on a particular day of the year, the earth's velocity through the ether is zero. But the earth is traveling around the sun at a rate of over a 100,000 km/hr, so even if the earth is temporarily motionless through the ether, six months from now it will be racing through it at over 200,000 km/hour (see figure 1.)

Let us call the rest frame of the ether $\mathbf{S}'$ and that of the earth $\mathbf{S}$. The earth is not exactly an inertial frame since it is accelerating around the sun and it is also rotating, but this is a small effect for what we are interested in, so we will ignore this and just assume that $\mathbf{S}$ is an inertial frame. We also assume that the rest frame of the ether is an inertial frame and that it is moving with velocity $\vec{v}$ wrt the Earth. Hence, if an observer in $\mathbf{S}'$ measures the velocity of a lightwave, he/she will measure its velocity to be $\vec{u}\,'$ where $|\vec{u}\,'| = c$. What will an observer on the earth measure for the velocity? If Galilean-Newtonian physics is correct, then they will measure a velocity $\vec{u} = \vec{u}\,' + \vec{v}$.

In 1881 Michelson and Morley set out to measure the earth's speed through the ether. Without any loss of generality, let us assume that $\vec{v} = v\hat{x}$, that is the ether is moving in

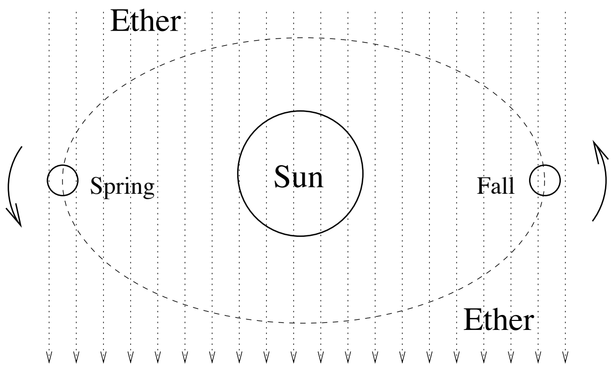


Figure 1: Earth traveling through the ether. If in spring it happens to be traveling along the direction the ether is moving, by fall it is moving in the opposite direction.

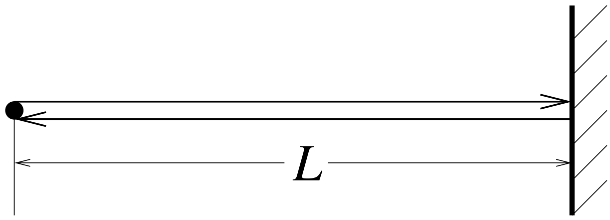


Figure 2: Light traveling a distance L along the direction to the ether before reflecting back off a mirror and returning.

the $\hat{x}$ direction with respect to the earth. Let us assume that a light pulse is sent out from position x_0 in **S** and travels along the $\hat{x}$ direction to x_1 where it hits a mirror and is reflected back. This is shown in figure 2. If we now assume the Galilean-Newtonian addition of velocities rule, then the velocity of the lightwave going out is

$$\vec{u}_{\text{out}} = (c+v)\hat{x}\,, \tag{1.11}$$

while the velocity coming back in is

$$\vec{u}_{\text{in}} = -(c-v)\hat{x}\,. \tag{1.12}$$

Let us assume that the distance between x_0 and x_1 is L. Then the time T_1 to go back and forth is

$$T_1 = \frac{L}{c+v} + \frac{L}{c-v} = \frac{2cL}{c^2-v^2}\,. \tag{1.13}$$

Now let us suppose that the light pulse is instead sent along the $\hat{y}$ direction according to an observer on the earth as shown in figure 3a. Again the pulse travels a distance L before hitting a mirror and reflecting backward toward the origin. Now the light going out has velocity

$$\vec{u}_{\text{out}} = \vec{u}\,'_{\text{out}} + \vec{v}\,, \tag{1.14}$$

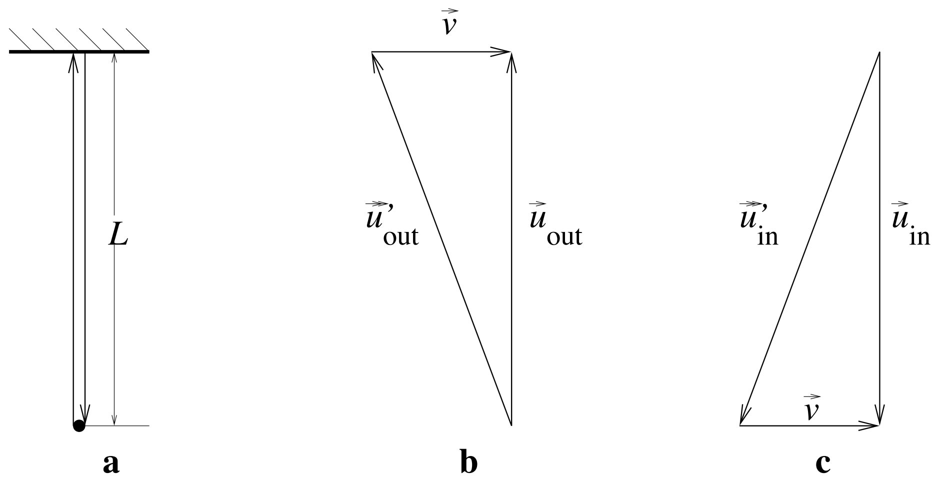


Figure 3: **a)** Light traveling a distance L orthogonal to the direction of the ether before reflecting back off a mirror and returning. **b)** Addition of velocity vectors for the outgoing light wave. **c)** Addition of velocity vectors for the incoming light wave.

as shown in figure 3b. Since $\vec{v}$ and $\vec{u}_{\text{out}}$ are at right angles to each other, we find using the Pythagoran theorem that

$$\vec{u}'_{\text{out}} \cdot \vec{u}'_{\text{out}} = \vec{u}_{\text{out}} \cdot \vec{u}_{\text{out}} + \vec{v} \cdot \vec{v}\,. \tag{1.15}$$

Since $|\vec{u}'_{\text{out}}| = c$, we find the magnitude of $\vec{u}_{\text{out}}$ is

$$|\vec{u}_{\text{out}}| = \sqrt{c^2 - v^2}\,. \tag{1.16}$$

Examing figure 3c, it is clear that the velocity vector $\vec{u}_{\text{in}}$ for the return inward, has the same magnitude as $\vec{u}_{\text{out}}$, $|\vec{u}_{\text{in}}| = |\vec{u}_{\text{out}}|$. Hence, the total time T_2, for the light wave to go to the mirror and back is

$$T_2 = \frac{L}{\sqrt{c^2 - v^2}} + \frac{L}{\sqrt{c^2 - v^2}} = \frac{2L}{\sqrt{c^2 - v^2}}\,. \tag{1.17}$$

Comparing T_2 in eq. (1.17) to T_1 in eq. (1.13), we see that they are different.

In 1887, the physicists Michelson and Morley measured the time difference between T_1 and T_2 using an interferometer. Figure 4 shows a rough sketch of an interferometer looking down from the top. Light of a certain wavelength λ is directed toward a half-silvered mirror, also known as a beam splitter, at an angle of 45 degrees. Half the light is reflected toward the left (upward in the diagram) and half is transmitted through. Both beams are reflected from mirrors and are directed back toward the half-silvered mirror. Some of the light is reflected and some is transmitted from both beams. If we consider the light that is transmitted from the first beam and reflected from the second beam, then the two beams recombine into one beam heading toward the right (down) where the light is observed as an interference pattern. The two beams can then interfere with each

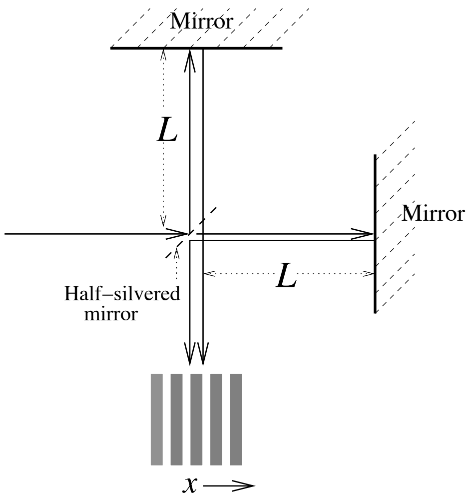


Figure 4: An interferometer splits light into two beams and then recombines them. The recombined light leads to interference patterns.

other, leading to a brighter or a dimmer beam, depending on whether the interference is constructive or destructive.

For the sake of argument, let us suppose that the ether is at rest in the sun's rest frame. Then in the earth's rest frame, the ether is traveling at 100,000 km/hr. Let us also suppose that in the earth's frame, the ether is traveling in the direction of the original light beam. If we assume that the distance between the half-silvered mirror and the other mirrors is L, then the time it takes the light that is originally transmitted to go back and forth from the half-silvered mirror is T_1 in eq. (1.13) and the time it takes the beam that was originally reflected to go back and forth is T_2 in eq. (1.17).

As the name implies, an interferometer measures interference and the interference is determined by how many wavelengths one beam differs from the other. If the difference is a a whole number then we have constructive interference and if the difference is a half number of wavelengths then there is destructive interference. Let us call this difference Δn. The number of wavelengths n that can pass by in a time T is $n = cT/\lambda$. Therefore, the difference in the number of wavelengths is

$$\Delta n = \frac{c\,\Delta T}{\lambda} = \frac{c\,(T_1 - T_2)}{\lambda}\,. \tag{1.18}$$

Now we will make an approximation. Even though 100,000 km/hr seems like a tremendous rate of speed, it is quite small compared to the speed of light c. Hence, we will make the following approximations for T_1 and T_2 using a Taylor expansion:

$$\begin{aligned} T_1 &= \frac{2\,L_1\,c}{c^2 - v^2} = \frac{2\,L_1}{c\,(1 - v^2/c^2)} \approx \frac{2\,L_1}{c}(1 + v^2/c^2) \\ T_2 &= \frac{2\,L_2}{\sqrt{c^2 - v^2}} = \frac{2\,L_2}{c\,\sqrt{1 - v^2/c^2}} \approx \frac{2\,L_2}{c}(1 + \tfrac{1}{2}v^2/c^2)\,. \end{aligned} \tag{1.19}$$

We use different lengths for the two arms, in order to incorporate interference fringes. If we let $L_1 = L + \Delta L(x)$ and $L_2 = L - \Delta L(x)$, with $\Delta L(x) << L$ and x is the position where the light hits the target, then the difference in wavelengths is

$$\Delta n \approx \frac{L\, v^2}{\lambda\, c^2} + \frac{2\, \Delta L(x)}{\lambda}\,. \tag{1.20}$$

If we assume that $\Delta L(x) = k\, x$ where k is a constant, then Δn is linear in x. Where it is a whole number we will have constructive interference giving a light band, and where it is half-integer there is destructive interference giving a dark band.

The actual experiment had the interferometer floating in a puddle of mercury, allowing the experimenters to rotate the apparatus so that the arm that was originally pointing along the direction of the ether could be made transverse and *vice versa*. Hence, under a 90 degree rotation, the difference in wavelengths is

$$\Delta n' \approx -\frac{L\, v^2}{\lambda\, c^2} + \frac{2\, \Delta L(x)}{\lambda}\,. \tag{1.21}$$

What the experimenters then measure is how this difference changes as they rotate the interferometer, in other words, they measure $\delta n \equiv \Delta n - \Delta n'$

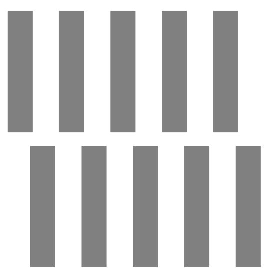
$$\delta n \approx 2\, \frac{L\, v^2}{\lambda\, c^2}\,, \tag{1.22}$$

where we see that the $\Delta L(x)$ has dropped out and δn is independent of x. This means that the interference pattern should have a constant shift to the side as the interferometer is rotated.

One wants to make L as big as possible so that δn is as big as possible, making it easier to measure the effect. Let us plug in some numbers to get an idea on its size. For Michelson and Morley, L was 11 m. Using $\lambda = 5000$ Å$= 5 \times 10^{-7}$ m, $c = 3 \times 10^5$ km/s, and $v = 100,000$ km/hr ≈ 28 km/s, we find

$$\delta n \approx 2\, \frac{11 \text{ m}}{5 \times 10^{-7} \text{ m}} \left(\frac{28 \text{ km/s}}{3 \times 10^5 \text{ km/s}}\right)^2 \approx 0.4\,. \tag{1.23}$$

This means that the interference pattern shifts over by 0.4 interference bands, as shown to the right.



The accuracy of the interferometer was good enough to measure δn as small as 0.01. And they saw no measurable effect. In other words they concluded that $\delta n < 0.01$, which means that $v < 13,000$ km/hr. Maybe they were unlucky and measured the earth's speed through the ether when it was moving in the same direction and speed as the ether. But that would mean that six months later the velocity should be $v \approx 200,000$ km/hr and they should measure δn to be 1.6. But again they found $\delta n < 0.01$. It seemed that the earth was always in the rest frame of the ether!

1.2 Einstein's postulates

The Michelson-Morley experiment is a null result; the earth does not seem to have any velocity in the ether. But we may also understand it another way – the speed that the light travels as measured by an observer on the earth does not depend on the earth's velocity. But the earth in the springtime and the earth in the fall are in different reference frames. To a very good approximation each of these frames is an inertial frame (it is not exactly inertial because of the gravitational force of the sun). In 1905, based on these observations, Einstein made the following two postulates

1. The laws of physics are identical in any inertial frame.

2. The speed of light in a vacuum, c, is the same in any inertial frame.

The first postulate does not sound particularly radical. Newton believed the same thing. It means that if you perform any sort of experiment in an inertial frame (say a laboratory in the basement of Ångström) and perform the same experiment in a different inertial frame (say a spaceship heading for Alpha Centauri), the result of the experiment will be the same.

The second postulate is the groundbreaker. We can see immediately that if we start with an inertial frame $\mathbf{S}$ and do a Galilean transformation to a new reference frame $\mathbf{S}'$, the new reference frame $\mathbf{S}'$ is *not* an inertial frame. This is because the velocity of a light wave in $\mathbf{S}'$ is $\vec{v}+\vec{c}$, where $\vec{c}$ is the velocity of the lightwave in $\mathbf{S}$, and so an observer using the coordinates of $\mathbf{S}'$ would measure a light speed different from c, violating the second postulate.

1.3 Lorentz transformations

If Galilean transformations do not transform us from one inertial frame to another, what does? To help us find the transformation, let us define some new notation. Let us write the space-time coordinates in an inertial frame $\mathbf{S}$ as

$$(ct, x, y, z) \equiv (x^0, x^1, x^2, x^3) \tag{1.24}$$

where x^i, $i = 1, 2, 3$ are the three spatial coordinates and x^0 is the time coordinate. We have defined x^0 with a factor of c so that x^0 also has dimensions of length. We will often write a space-time coordinate as x^μ, $\mu = 0, 1, 2, 3$, putting the time coordinate on equal footing with the three spatial coordinates. x^μ is an example of a "4-vector", a 4 dimensional vector with some special properties which we will describe later. We will use Greek letters, μ, ν, λ *etc.* to signify the coordinates of a 4-vector. We will use Latin letters i, j, k *etc.* to signify the three spatial coordinates.

Let us now assume that $\mathbf{S}'$ is also an inertial frame with coordinates

$$(ct', x', y', z') = (x^{0'}, x^{1'}, x^{2'}, x^{3'})\,, \tag{1.25}$$

and we will write the general space-time coordinate as $x^{\mu'}$. Einstein's first postulate tells us that the transformation between coordinates in $\mathbf{S}$ and those in $\mathbf{S}'$ is linear. To see this, suppose we have a clock in **S** moving with constant velocity. This means that no forces are acting on the clock since **S** is assumed to be an inertial frame. Let us further suppose that the clock gives a display of the time that it measures, which we call τ. An observer in **S** then uses $x^\mu(\tau)$ to be the space-time position of the clock as measured by him when the clock display is at τ. Since the clock is moving at constant velocity, this means that $x^\mu(\tau)$ is linear in τ, from which we conclude that

$$\frac{d^2x^\mu}{d\tau^2} = 0 \,. \tag{1.26}$$

If another observer in $\mathbf{S'}$ uses his coordinates to measure the space-time position of the clock $x^{\mu'}(\tau)$ as a function of τ, he too sees that the clock is moving with constant velocity, and by the same argument finds that

$$\frac{d^2x^{\mu'}}{d\tau^2} = 0 \,. \tag{1.27}$$

Now using the chain rule of calculus we have

$$\frac{dx^{\mu'}}{d\tau} = \sum_{\nu=0}^{3} \frac{dx^\nu}{d\tau}\frac{\partial x^{\mu'}}{\partial x^\nu} \equiv \frac{dx^\nu}{d\tau}\frac{\partial x^{\mu'}}{\partial x^\nu} \tag{1.28}$$

We have introduced some new notation to save us a fair amount of writing in (1.28). Notice that the index ν appears twice, once in the numerator of a derivative and once in the denominator. It will always be the case that when the same index appears twice in this fashion, it will be summed over. Since we know that it must be summed, we might as drop the sum notation $\sum$, but the sum is still there. Let us now take one more derivative on (1.28), again using the chain rule

$$\begin{aligned} \frac{d^2x^{\mu'}}{d\tau^2} = 0 &= \frac{d^2x^\nu}{d\tau^2}\frac{\partial x^{\mu'}}{\partial x^\nu} + \frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau}\frac{\partial^2 x^{\mu'}}{\partial x^\nu \partial x^\lambda} \\ &= \quad 0 \quad + \frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau}\frac{\partial^2 x^{\mu'}}{\partial x^\nu \partial x^\lambda}\,, \end{aligned} \tag{1.29}$$

and so we conclude that

$$\frac{\partial^2 x^{\mu'}}{\partial x^\nu \partial x^\lambda} = 0 \,. \tag{1.30}$$

This means that

$$x^{\mu'} = \Lambda^{\mu'}{}_\nu\, x^\nu + C^{\mu'}\,, \tag{1.31}$$

where $\Lambda^{\mu'}{}_\nu$ and $C^{\mu'}$ are constants. Notice that we are still using the repeated index notation where the index ν appears in a down position in $\Lambda^{\mu'}{}_\nu$ and an up position in x^ν, meaning that we are summing over ν from 0 to 3. The index μ' appears only once, in the up position on both sides of the equation, so it is *not* summed over. This means that (1.31) is really four equations, one for each value of μ'. It is more convenient to consider displacements, Δx^μ where Δx^μ is the difference between two events. In this case the constant $C^{\mu'}$ drops out and we find

$$\Delta x^{\mu'} = \Lambda^{\mu'}{}_{\nu}\, \Delta x^\nu\,. \tag{1.32}$$

We could also consider the possibility of differentials dx^μ, which one can think of as infinitesimal displacements. For this we have the similar relation

$$dx^{\mu'} = \Lambda^{\mu'}{}_{\nu}\, dx^\nu\,. \tag{1.33}$$

Let us now suppose that $\mathbf{S}'$ is moving with velocity $\vec{v} = v\hat{x}$ *wrt* $\mathbf{S}$. If the coordinates in $\mathbf{S}'$ were related to those in $\mathbf{S}$ by Galilean transformations we would have

$$\Delta x' = \Delta x - v\Delta t, \qquad \Delta y' = \Delta y, \qquad \Delta z' = \Delta z\,. \tag{1.34}$$

We want to modify these equations so that they are still linear. We should also modify them so that when $\Delta x = v\Delta t$, $\Delta x' = 0$. Furthermore, when $\Delta y = 0$ we have $\Delta y' = 0$ and when $\Delta z = 0$, we have $\Delta z' = 0$. The transformations that are consistent with these constraints are

$$\Delta x' = \gamma\,(\Delta x - v\Delta t), \qquad \Delta y' = \gamma_y\, \Delta y, \qquad \Delta z' = \gamma_z\, \Delta z\,. \tag{1.35}$$

where γ, γ_y and γ_z are constants to be determined.

Now let us consider the reverse transformation that takes us back from the $\mathbf{S}'$ coordinates to the $\mathbf{S}$ coordinates. $\mathbf{S}$ is moving with velocity $-v\hat{x}$ *wrt* $\mathbf{S}'$and the transformation of the coordinates is

$$\Delta x = \gamma\,(\Delta x' + v\Delta t'), \qquad \Delta y = \gamma_y\, \Delta y', \qquad \Delta z = \gamma_z\, \Delta z'\,, \tag{1.36}$$

with the γ factors being the same as in (1.35). To show that the γ factors are the same, note that (1.35) is the true for *any* Δx and Δt. Therefore, let us define $\Delta\tilde{x} = -\Delta x$ and $\Delta\tilde{x}' = -\Delta x'$. Clearly we have that

$$\Delta\tilde{x}' = \gamma\,(\Delta\tilde{x} + v\Delta t)\,, \tag{1.37}$$

and we can think of the coordinates $\Delta\tilde{x}$ and $\Delta\tilde{x}'$ as the x coordinates in two reference frames $\tilde{\mathbf{S}}$ and $\tilde{\mathbf{S}}'$ but now $\tilde{\mathbf{S}}'$ is moving with velocity $-v\hat{x}$ *wrt* $\tilde{\mathbf{S}}$, the same relation that $\mathbf{S}$ has to $\mathbf{S}'$. Hence, (1.36) follows.

To find what γ is, suppose that we have a light ray moving in the x direction in $\mathbf{S}$ and let Δx^μ refer to its displacement. By Einstein's second postulate we must have

$$\Delta x = c\,\Delta t\,. \tag{1.38}$$

By Einstein's second postulate, the same light ray as seen by an observer in $\mathbf{S}'$ must also be moving with speed c and so the displacements in the primed coordinates satisfy

$$\Delta x' = c\,\Delta t'\,. \tag{1.39}$$

Plugging these expressions into eqs. (1.35) and (1.37), we arrive at the two linear equations

$$\begin{aligned} c\,\Delta t' &= \gamma\,(c-v)\Delta t \\ c\,\Delta t &= \gamma\,(c+v)\Delta t'\,. \end{aligned} \tag{1.40}$$

Hence it follows that

$$\gamma^2 = \frac{c^2}{c^2-v^2} = \frac{1}{1-v^2/c^2} \qquad \Rightarrow \qquad \gamma = \frac{1}{\sqrt{1-v^2/c^2}}\,. \tag{1.41}$$

Notice that γ blows up as v approaches c. In fact, if $v > c$ then γ is imaginary! Hence, we might expect that c is a limiting velocity for v. We will see a more physical reason for this when we discuss causality.

We still need to find the time coordinate in $\mathbf{S}'$. Notice that (1.36) can be rewritten as

$$\Delta t' = \frac{1}{v\gamma}\Delta x - \frac{1}{v}\Delta x'\,. \tag{1.42}$$

Substituting the expression for $\Delta x'$ in (1.35) into the above leads to

$$\begin{aligned} \Delta t' &= \frac{1}{v\gamma}\Delta x - \frac{\gamma}{v}(\Delta x - v\Delta t) = \gamma\,\Delta t - \frac{\gamma}{v}(1-1/\gamma^2)\Delta x \\ &= \gamma\,\Delta t - \frac{\gamma}{v}(1-(1-v^2/c^2))\Delta x \\ \Delta t' &= \gamma\left(\Delta t - \frac{v}{c^2}\Delta x\right)\,. \end{aligned} \tag{1.43}$$

We can also easily find the reverse transformation

$$\Delta t \;=\; \gamma\left(\Delta t' + \frac{v}{c^2}\Delta x'\right)\,. \tag{1.44}$$

Unlike Galilean transformations, we see that the time measured between events by an observer in $\mathbf{S}'$ is *not* the same as the time measured between events by an observer in $\mathbf{S}$.

Let us now find γ_y and γ_z. We again assume that we have a light ray, but this time it is traveling in the y direction in $\mathbf{S}$. Therefore, $\Delta y = c\,\Delta t$. In $\mathbf{S}'$ the light ray will have a component in the x direction, namely $\Delta x' = \gamma(\Delta x - v\Delta t) = -\gamma v\Delta t$, since $\Delta x = 0$. The displacement of the time coordinate in $\mathbf{S}'$ is $\Delta t' = \gamma\,\Delta t$. Hence the speed of the light ray measured by an observer in $\mathbf{S}'$ is

$$\frac{\sqrt{(\Delta y')^2+(\Delta x')^2}}{\Delta t'} = \frac{\sqrt{c^2\gamma_y^2(\Delta t)^2+\gamma^2v^2(\Delta t)^2}}{\gamma\Delta t} = \frac{\sqrt{c^2\gamma_y^2+\gamma^2v^2}}{\gamma} = \sqrt{c^2\gamma_y^2/\gamma^2+v^2} \tag{1.45}$$

By the second postulate this must equal c, which then gives[2] $\gamma_y = 1$. By a similar argument we can also argue that $\gamma_z = 1$.

[2]Actually we get c for the speed if $\gamma_y = -1$ as well, but this is not a valid solution since it does not smoothly go to 1 as we take $v \to 0$.

Let us now collect our results in terms of the components of $\Lambda^{\mu'}{}_{\nu}$. Let us rewrite our equations in the 4-vector notation:

$$\begin{aligned}
\Delta x^{0'} &= \gamma\left(\Delta x^0 - \frac{v}{c}\Delta x^1\right) \\
\Delta x^{1'} &= \gamma\left(\Delta x^1 - \frac{v}{c}\Delta x^0\right) \\
\Delta x^{2'} &= \Delta x^2 \\
\Delta x^{3'} &= \Delta x^3 \,.
\end{aligned} \tag{1.46}$$

Comparing (1.46) with (1.32) we find that the components of $\Lambda^{\mu'}{}_{\nu}$ are

$$\begin{aligned}
\Lambda^{0'}{}_0 &= \gamma \qquad \Lambda^{0'}{}_1 = -\frac{v}{c}\gamma \qquad \Lambda^{0'}{}_2 = 0 \qquad \Lambda^{0'}{}_3 = 0 \\
\Lambda^{1'}{}_0 &= -\frac{v}{c}\gamma \qquad \Lambda^{1'}{}_1 = \gamma \qquad \Lambda^{1'}{}_2 = 0 \qquad \Lambda^{1'}{}_3 = 0 \\
\Lambda^{2'}{}_0 &= 0 \qquad \Lambda^{2'}{}_1 = 0 \qquad \Lambda^{2'}{}_2 = 1 \qquad \Lambda^{2'}{}_3 = 0 \\
\Lambda^{3'}{}_0 &= 0 \qquad \Lambda^{3'}{}_1 = 0 \qquad \Lambda^{3'}{}_2 = 0 \qquad \Lambda^{3'}{}_3 = 1 \,.
\end{aligned} \tag{1.47}$$

It is also convenient to write the transformations in matrix form

$$\begin{pmatrix} \Delta x^{0'} \\ \Delta x^{1'} \\ \Delta x^{2'} \\ \Delta x^{3'} \end{pmatrix} = \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta x^{0} \\ \Delta x^{1} \\ \Delta x^{2} \\ \Delta x^{3} \end{pmatrix} . \tag{1.48}$$

Notice that the primed indices in $\Lambda^{\mu'}{}_{\nu}$ correspond to the rows of the matrix while the unprimed indices correspond to the columns. So the entries in the first *row* go with the $0'$ index, those in the second row with $1'$ *etc.*, while the entries in the first *column* go with the 0 index, those in the second column with 1 *etc.*

This transformation from **S** to **S**$'$ is an example of a *boost*, more specifically a boost in the x direction, and boosts are part of a set of transformations called *Lorentz transformations*. We can combine two Lorentz transformations to give a third transformation. For example, suppose that we have a third inertial frame **S**$''$, with coordinates related to the coordinates in **S**$'$ by

$$\Delta x^{\mu''} = \tilde{\Lambda}^{\mu''}{}_{\nu'}\Delta x^{\nu'} \,, \tag{1.49}$$

where $\tilde{\Lambda}^{\mu''}{}_{\nu'}$ is the Lorentz transformation relating **S**$''$ to **S**$'$. Then these coordinates can be related to those in **S** by

$$\Delta x^{\mu''} = \tilde{\Lambda}^{\mu''}{}_{\nu'}\Lambda^{\nu'}{}_{\lambda}\Delta x^{\lambda} = (\tilde{\Lambda}\Lambda)^{\mu''}{}_{\lambda}\Delta x^{\lambda} \,. \tag{1.50}$$

Notice that the ν' index in $\tilde{\Lambda}^{\mu''}{}_{\nu'}$ and $\Lambda^{\nu'}{}_{\lambda}$ is repeated, hence it is summed over. In terms of matrices, we can write this as

$$\begin{pmatrix} \Delta x^{0''} \\ \Delta x^{1''} \\ \Delta x^{2''} \\ \Delta x^{3''} \end{pmatrix} = \begin{pmatrix} \tilde{\Lambda}^{0''}{}_{0'} & \tilde{\Lambda}^{0''}{}_{1'} & \tilde{\Lambda}^{0''}{}_{2'} & \tilde{\Lambda}^{0''}{}_{3'} \\ \tilde{\Lambda}^{1''}{}_{0'} & \tilde{\Lambda}^{1''}{}_{1'} & \tilde{\Lambda}^{1''}{}_{2'} & \tilde{\Lambda}^{1''}{}_{3'} \\ \tilde{\Lambda}^{2''}{}_{0'} & \tilde{\Lambda}^{2''}{}_{1'} & \tilde{\Lambda}^{2''}{}_{2'} & \tilde{\Lambda}^{2''}{}_{3'} \\ \tilde{\Lambda}^{3''}{}_{0'} & \tilde{\Lambda}^{3''}{}_{1'} & \tilde{\Lambda}^{3''}{}_{2'} & \tilde{\Lambda}^{3''}{}_{3'} \end{pmatrix} \begin{pmatrix} \Lambda^{0'}{}_0 & \Lambda^{0'}{}_1 & \Lambda^{0'}{}_2 & \Lambda^{0'}{}_3 \\ \Lambda^{1'}{}_0 & \Lambda^{1'}{}_1 & \Lambda^{1'}{}_2 & \Lambda^{1'}{}_3 \\ \Lambda^{2'}{}_0 & \Lambda^{2'}{}_1 & \Lambda^{2'}{}_2 & \Lambda^{2'}{}_3 \\ \Lambda^{3'}{}_0 & \Lambda^{3'}{}_1 & \Lambda^{3'}{}_2 & \Lambda^{3'}{}_3 \end{pmatrix} \begin{pmatrix} \Delta x^{0} \\ \Delta x^{1} \\ \Delta x^{2} \\ \Delta x^{3} \end{pmatrix} \tag{1.51}$$

where the matrices are multiplied together in the usual fashion of matrix multiplication. In particular, let us suppose that $\mathbf{S}''=\mathbf{S}$, meaning that $\tilde{\Lambda}^{\mu}{}_{\nu'} = \Lambda^{-1\mu}{}_{\nu'}$, where Λ^{-1} refers to the inverse of the matrix in (1.48). It is not hard to check that

$$\Lambda^{-1} = \begin{pmatrix} \gamma & +\frac{v}{c}\gamma & 0 & 0 \\ +\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{1.52}$$

which is consistent with the transformations in (1.36). In fact, let us explicitly check this:

$$\begin{pmatrix} \gamma & +\frac{v}{c}\gamma & 0 & 0 \\ +\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \gamma^2 - \frac{v^2}{c^2}\gamma^2 & -\frac{v}{c}\gamma^2 + \frac{v}{c}\gamma^2 & 0 & 0 \\ +\frac{v}{c}\gamma^2 - \frac{v}{c}\gamma^2 & \gamma^2 - \frac{v^2}{c^2}\gamma^2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{1.53}$$

The convention is to write $\Lambda^{-1\mu}{}_{\nu'}$ as $\Lambda^{\mu}{}_{\nu'}$, that is without the "-1" exponent, since from the position of the primed and unprimed indices it is clear that this is the inverse of $\Lambda^{\mu'}{}_{\nu}$.

We can also have boosts in directions other than the x direction. For example a boost in the y direction with velocity v would look like

$$\Lambda_{y-\mathrm{boost}} = \begin{pmatrix} \gamma & 0 & -\frac{v}{c}\gamma & 0 \\ 0 & 1 & 0 & 0 \\ -\frac{v}{c}\gamma & 0 & \gamma & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{1.54}$$

which is a different Lorentz transformation. In fact we can have Lorentz transformations that are not boosts at all. These are rotations in the spatial directions. For example, a rotation of angle ϕ in the $x-y$ plane leaves the x^0 and x^3 coordinates alone and only mixes the x^1 and x^2 coordinates. The matrix for this Lorentz transformation is given by

$$\Lambda_{\mathrm{rot}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi & 0 \\ 0 & \sin\phi & \cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{1.55}$$

In fact, there is a way of reformulating boosts so that they look like the rotations in (1.55). We can define the *rapidity*, ξ, as $\cosh\xi = \gamma$ and $\sinh\xi = \frac{v}{c}\gamma$, where cosh and sinh refers to the hyperbolic cosine and hyperbolic sine respectively. Notice that the trigonometric identity $\cosh^2\xi - \sinh^2\xi = 1$ is automatically satisfied. Then the boost matrix Λ in (1.48) becomes

$$\Lambda = \begin{pmatrix} \cosh\xi & -\sinh\xi & 0 & 0 \\ -\sinh\xi & \cosh\xi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{1.56}$$

with an obvious similarity to the transformation in (1.55).

1.4 Extra: Lorentz transformations form a group

This subsection is outside the main part of the course and may be skipped by those of you who are pressed for time or have something better to do.

Lorentz transformations make what is called the *Lorentz group*. Let's define a group. A group $\mathcal{G}$ has the following properties.

1. A group has an operation called "multiplication", which we denote by "$\cdot$".
2. If g and h are elements of $\mathcal{G}$, then $g \cdot h$ is also an element of $\mathcal{G}$.
3. A group has a unique element "1", such that $1 \cdot g = g \cdot 1 = g$.
4. Every element g has a unique inverse g^{-1} which is also an element of $\mathcal{G}$ such that $g \cdot g^{-1} = g^{-1} \cdot g = 1$.

That's it. Notice one statement we did not make is that g and h commute, namely $g \cdot h = h \cdot g$. If this statement were true for all g and h in $\mathcal{G}$, then we would say that the group is *Abelian*.

Let us now apply this to Lorentz transformations. The 4×4 matrices Λ form a *representation of the group*. There are other representations but we will not be concerned with those here. Group multiplication is just matrix multiplication, to wit if Λ is one element of the group and $\tilde{\Lambda}$ is another element of the group, then $\tilde{\Lambda}\Lambda$ is another element of the group. In general $\tilde{\Lambda}\Lambda \neq \Lambda\tilde{\Lambda}$. For example, a simple exercise shows that for the transformations in (1.56) and (1.55) that $\Lambda\Lambda_{\mathrm{rot}} \neq \Lambda_{\mathrm{rot}}\Lambda$

1.5 A relativistic invariant and proper time

An important part of this course is finding quantities that are the same for any inertial frame. Such a quantity is called a *relativistic invariant*, or simply an *invariant*. A main reason why such quantities are useful is that they are often easy to compute in one inertial frame, but not so easy in another one. So if we want to compute the invariant, we simply boost to the frame where it is easy to compute the invariant.

One such invariant is given by

$$(\Delta x^0)^2 - (\Delta x^1)^2 - (\Delta x^2)^2 - (\Delta x^3)^2 \equiv -(\Delta s)^2 \equiv c^2(\Delta\tau)^2\,. \tag{1.57}$$

Δs is called the *invariant length* and τ is the *proper time*. Notice that this invariant length is similar looking to a length of a vector in 4 dimensions. The only difference is the relative minus sign between the x^0 component and the three spatial components.

Let's now show that this is an invariant under Lorentz transformations. Let us assume that we have the Lorentz boost in (1.48). Then

$$
\begin{aligned}
&(\Delta x^{0'})^2 - (\Delta x^{1'})^2 - (\Delta x^{2'})^2 - (\Delta x^{3'})^2 \\
&= \left(\gamma \Delta x^0 - \frac{v}{c}\gamma \Delta x^1\right)^2 - \left(\gamma \Delta x^1 - \frac{v}{c}\gamma \Delta x^0\right)^2 - (\Delta x^2)^2 - (\Delta x^3)^2 \\
&= \gamma^2 \left(1 - \frac{v^2}{c^2}\right) ((\Delta x^0)^2 - (\Delta x^1)^2) - (\Delta x^2)^2 - (\Delta x^3)^2 \\
&= (\Delta x^0)^2 - (\Delta x^1)^2 - (\Delta x^2)^2 - (\Delta x^3)^2 \,. \qquad (1.58)
\end{aligned}
$$

Hence, it is the same in both frames. It is clear that we could just as easily boost in a different direction and still find it invariant. A spatial rotation is also invariant since the spatial parts of (1.57) are the negative length squared of a three dimensional vector, and this length is invariant under three-dimensional rotations.

Let us now explain the term "proper time". Suppose that Δx^μ is the displacement vector for a body with constant velocity. If $\mathbf{S}'$ is the rest-frame of the particle, then in this frame the spatial components of the displacement are zero, since the body is not moving in this frame. Hence, $\Delta\tau = \Delta t$, so the proper time is the elapsed time as measured by a clock in its rest-frame.

Note that the invariant for differentials is

$$
(dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 \equiv -(ds)^2 \equiv c^2 (d\tau)^2 \,. \qquad (1.59)
$$

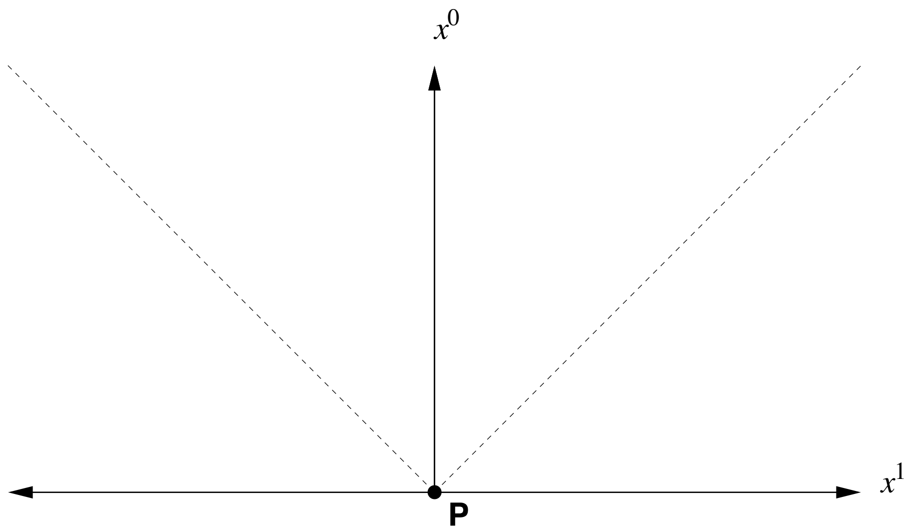


Figure 5: A space-time diagram with the origin labeled by event **P** and light-like trajectories shown with dashed lines.

2 Relativistic Physics

2.1 Space-time diagrams

A useful tool for analyzing relativistic physics is the *space-time diagram.* A space-time diagram is a two-dimensional plot of the time coordinate on the vertical axis and one of the spatial coordinates along the horizontal axis. We will almost always choose this coordinate to be $x = x^1$.

Such a diagram is shown in figure 5. The point at the origin is labeled **P**. Since this is a point in space-time, we call **P** an event. We have also shown two trajectories (dashed lines) emanating from **P** at 45 degree angles. On these trajectories, we have that $x^1 = \pm x^0$. Hence, if these are trajectories of particles we see that they have velocity $u = x/t = \pm c$. Therefore, these particles are traveling at the speed of light and we call the trajectories *light-like* trajectories. They are also called *null* trajectories since for a displacement along this trajectory, $(\Delta s)^2 = (\Delta x^1)^2 - (\Delta x^0)^2 = 0$.

The x^0 axis is itself the trajectory for a body at rest in frame **S**. Hence this is called a *time-like* trajectory. A time-like trajectory has $(\Delta s)^2 < 0$. The x^1 axis is called a *space-like* trajectory. These trajectories have $(\Delta s)^2 > 0$. Trajectories for bodies are also known as *world-lines*.

We can also include the axes for a different frame on our space-time diagram. Notice that for frame **S** the x^1 axis is defined by the line $x^0 = 0$, while the x^0 axis is defined by the line $x^1 = 0$. Hence, we can draw axes for a different frame **S′** by finding the lines $x^{0'} = 0$ and $x^{1'} = 0$. Let us assume that the origin is the same for both frames. Then the line $x^{0'} = 0$ is $\gamma x^0 - \frac{v}{c}\gamma x^1 = 0$. Hence, this is the line $x^0 = \frac{v}{c}x^1$. Assuming that $v < c$, then the slope of this line is less than 1. Likewise, the line $x^{1'} = 0$ is $\gamma x^1 - \frac{v}{c}\gamma x^0 = 0$ which leads to the line $x^0 = \frac{c}{v}x^1$ which has a slope greater than one. The space-time diagram showing both sets of axes is shown in figure 6. Notice that as v gets closer to c,

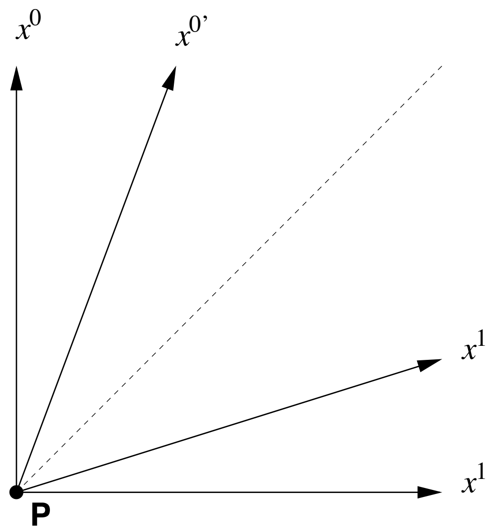


Figure 6: A space-time diagram showing the axes for reference frames **S** and **S**$'$.

the slopes of the $x^{0'}$ axis and the $x^{1'}$ axis get closer to one. In other words they approach the null trajectory from opposite sides. If **S**$'$ is the rest frame of a body, then the $x^{0'}$ axis is the body's trajectory, assuming that the trajectory goes through the origin at **P**.

2.2 The relativity of simultaneity and causality

Figure 7 shows the space-time diagram with events **R** and **Q** included. Notice that event **R** occurs at the same time as event **P** according to an observer in **S** since both events sit on the x^1 axis which has $x^0 = 0$. But an observer in **S**$'$ would see something different. According to this observer, the event **R** happened *before* **P** since **R** is below the $x^{1'}$ axis and hence occurred for some time when $t' < 0$. On the other hand, event **Q** is simultaneous with **P** according to the **S**$'$ observer, but occurs after **P** according to the **S** observer. Thus we see that the notion of simultaneous events is a relative concept.

We can now ask if event **P** can *cause* event **Q**. We can say that this causality would occur if a signal can emanate from **P** and travel to **Q**, thus causing it. Let us suppose that the signal is some body whose world-line is the $x^{1'}$ axis. One consequence of this is that the body's speed is greater than the speed of light. We have already seen that trouble occurs if we try to boost to a frame moving faster than the speed of light, in that γ is imaginary. But there is a more fundamental problem. According to the third reference frame in figure 7, **S**$''$, event **Q** happened *before* event **P**. Hence, there is no way that **P** could cause **Q**. Hence, we must conclude that there is no way to send a signal along $x^{1'}$. In fact, we cannot send a signal along *any* space-like trajectory, only along time-like or null trajectories. Looking again at figure 7, we see that event **W** is connected to **P** by a time-like trajectory. Hence, **P** can cause **W**. When this occurs for two events we say that they are *causally connected*. It is also true that **W** and **P** are not simultaneous in *any* inertial frame.

Figure 8 is known as a light-cone diagram. Any event in the shaded region below **P**, including the boundary can cause **P**, because there is a time-like or light-like world-

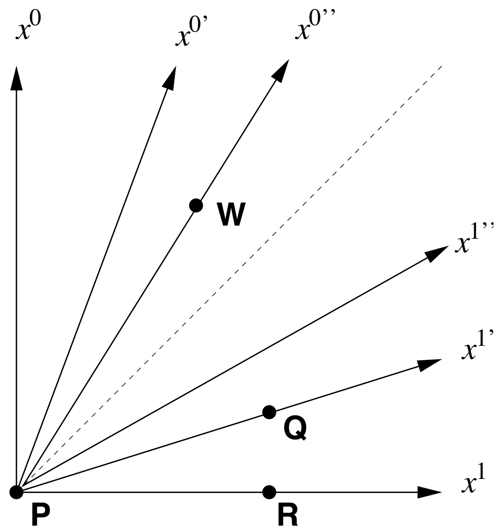


Figure 7: Event **R** is simultaneous with **P** according to an observer in **S** but not according to an observer in $\mathbf{S}'$. Similarly, event **Q** is simultaneous with event **P** according to an observer in $\mathbf{S}'$, but not according to an observer in **S**. Event **W** is causally connected to **P**.

line that connects this event to **P**. This region is called the *past light-cone*. Similarly, any event in shaded region *above* **P** can be caused by **P** because there is a time-like or light-like world-line connecting **P** to the event.

2.3 Length contraction

Suppose we have a bar of length L that is stationary in $\mathbf{S}'$ and is aligned along the x axis. What length would an observer in **S** measure?

In problems of this type, the key to solving this question is properly setting up the equations. Firstly, how would an observer in **S** make the measurement? A reasonable thing to do is to measure the positions of the front and back of the bar simultaneously, that is find their positions at the same time t, and then measure the displacement, Δx. Since the observer is making a simultaneous measurement according to his clocks, we have that $\Delta t = 0$. We also know that the displacement in $\mathbf{S}'$ is $\Delta x' = L$ since the bar is stationary in this frame.

We now use the relations in (1.36) and (1.44) to write

$$\begin{aligned} \Delta t = 0 &= \gamma\left(\Delta t' + \frac{v}{c^2}\Delta x'\right) \qquad \Rightarrow \qquad \Delta t' = -\frac{v}{c^2}\Delta x' \\ \Delta x &= \gamma\left(\Delta x' + v\Delta t'\right) = \gamma\left(1 - \frac{v^2}{c^2}\right)\Delta x' = \frac{L}{\gamma}\,. \end{aligned} \tag{2.1}$$

Hence the observer in **S** measures a contracted length since $\gamma > 1$.

This is particularly clear if we look at the corresponding space-time diagram shown in figure 9. In this diagram we have shown the trajectories for the front and back of the bar. We are still assuming that $\mathbf{S}'$ is the rest-frame. The length of the bar in the rest-frame is the length between the intersection points of the trajectories with the $x^{1'}$

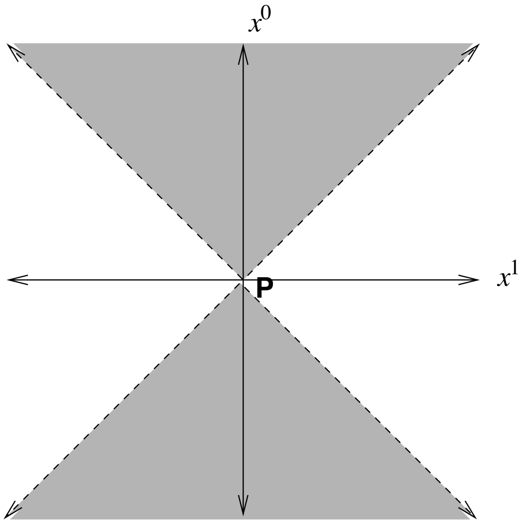


Figure 8: Light-cone diagram for event **P**. Those events in the past light-cone can cause **P** and those events in the future light-cone can be caused by **P**.

axis, while the length in the **S** frame is the length between the intersection points on the x^1 axis. It seems clear that the distance is longer on the $x^{1'}$ axis. However, one has to be a little careful. Just because something looks longer does not necessarily make it so because of the minus sign that appears in the invariant length. In any case, $\Delta x^1 < L$ because $(\Delta x^1)^2 = (\Delta x^{1'})^2 - (\Delta x^{0'})^2$.

2.4 Time dilation

Another interesting phenomenon is *time dilation*. Suppose we have a clock at a fixed point in $\mathbf{S}'$ and it measures a time interval $\Delta t'$. What is the time interval measured by an observer in **S**? Note that $\Delta t' = \Delta\tau$ is the elapsed proper time on the clock.

Again, the key to solving this question is setting up the problem correctly. Since the clock is stationary in $\mathbf{S}'$ we have that there is no spatial displacement in this frame, thus $\Delta x' = 0$. Then we just use (1.44) to obtain

$$\Delta t = \gamma \Delta t' \,. \tag{2.2}$$

Since $\gamma > 1$, the observer in **S** measures a longer elapsed time than the proper time of the clock. Therefore, the clock in $\mathbf{S}'$ seems to be running slow according the observer's clocks in **S**.

Time dilation is a real measurable effect. For example consider the case of particle decay. The lifetime of a particle τ is statistical, but we can treat it as the internal clock of the particle. If we have have many particles of the same type at rest and we measure how long it takes for them to decay, then τ will be the statistical average.

Now suppose that the particles are moving with velocity v *wrt* to an observer in **S**. Then this observer would measure the lifetime to be $\gamma\tau$. As an example, consider the elementary particle called the "muon". The muon has a relatively long life-time of 10^{-8}

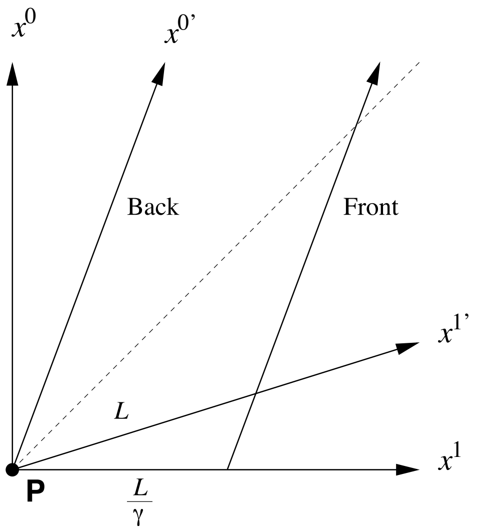


Figure 9: Space-time diagram for a bar stationary in $\mathbf{S}'$. The world-lines for the front and back of the bar are shown.

seconds. If the muon were traveling at a velocity v where $v^2 = 0.99c^2$, then $\gamma = 10$ and the measured lifetime would be around 10^{-7} seconds.

2.5 The twin paradox

In our discussion of time dilation, you might have noticed the possibility of a contradiction. We argued that an observer in **S** would measure the stationary clocks in $\mathbf{S}'$ to be running slow. But an observer in $\mathbf{S}'$ would also think the clocks in **S** are running slow. We can then spin this into a paradox as follows:

Suppose there are two clocks **A** and **B** and the world-lines of the clocks both pass through the space-time point **P**. We assume that **A** is at rest in **S** and **B** is at rest in $\mathbf{S}'$. **B** then travels to a space-time point **Q** at which point it accelerates into a new frame with velocity $-v$ *wrt* **S** and travels back to meet up with **A** again at the space-time point **R**. Which clock has more elapsed time? The paradox is that naive thinking would say that they both think the other clock is slower than their own, since at all times one clock was moving with velocity $\pm v$ *wrt* the other clock. But that can't be right. However, a bit more careful thinking shows that the problem is not entirely symmetric, since **B** undergoes an instantaneous acceleration halfway through its journey, while **A** does not. This is especially clear if we look at the space-time diagram in figure 10, where obviously the world-line for **A** looks different than the world-line for **B**.

We can then compare the times of the clocks by measuring the invariant lengths along the trajectories. From the diagram, it is clear that **B**'s invariant length is twice that in going from **P** to **Q**. Letting Δt_A be the elapsed time along **A**'s trajectory between **P** and **R**, then the elapsed time in going on **B**'s trajectory is

$$\Delta t_B = 2\sqrt{\left(\frac{\Delta t_A}{2}\right)^2 - \frac{v^2}{c^2}\left(\frac{\Delta t_A}{2}\right)^2} = \frac{\Delta t_A}{\gamma}\,. \tag{2.3}$$

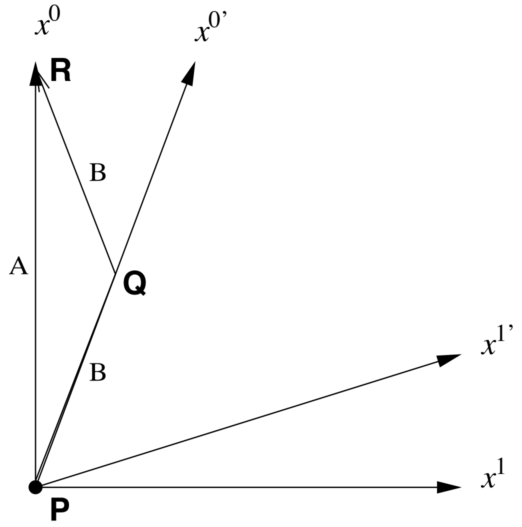


Figure 10: Space-time diagram for two clocks **A** and **B**. **B** has an instaneous acceleration at point **Q**.

Hence, **B**'s clock has less elapsed time and both sides would agree. Thus, there is no paradox.

Notice that while **B**'s trajectory looks longer in the diagram, the elapsed time is shorter. This is because of the relative minus sign that appears in the invariant in (1.57).

We can also obtain our result another way. Suppose that the distance that **B** travels away from **A** is L according to an observer in **S**. Then **A** would think the total time for **B**'s journey is $\Delta t_A = 2vL$. But **B** would see this length contracted to L/γ, so according to his clock the time for the journey is $\Delta t_B = 2vL/\gamma = \Delta t_A/\gamma$.

2.6 Velocity transformations

Suppose that a body is moving with velocity $\vec{u}\,'$ in inertial frame $\mathbf{S}'$. What is the velocity $\vec{u}$ measured by an observer in **S**? To measure a velocity, one measures the spatial displacement and divides by the elapsed time. Thus, the velocity components in $\mathbf{S}'$ are given by

$$u'_x = \frac{\Delta x'}{\Delta t'}\,, \qquad u'_y = \frac{\Delta y'}{\Delta t'}\,, \qquad u'_z = \frac{\Delta z'}{\Delta t'}\,, \tag{2.4}$$

while the components in **S** are

$$u_x = \frac{\Delta x}{\Delta t}\,, \qquad u_y = \frac{\Delta y}{\Delta t}\,, \qquad u_z = \frac{\Delta z}{\Delta t}\,, \tag{2.5}$$

We again make use of the transformations in (1.36) and (1.44) to write the velocity components in (2.5) as

$$\begin{aligned} u_x &= \frac{\Delta x' + v\Delta t'}{\Delta t' + \frac{v}{c^2}\Delta x'} = \frac{u'_x + v}{1 + \frac{u'_x v}{c^2}}\,, \\ u_y &= \frac{\Delta y'}{\gamma(\Delta t' + \frac{v}{c^2}\Delta x')} = \frac{u'_y}{\gamma(1 + \frac{u'_x v}{c^2})}\,, \\ u_z &= \frac{\Delta z'}{\gamma(\Delta t' + \frac{v}{c^2}\Delta x')} = \frac{u'_z}{\gamma(1 + \frac{u'_x v}{c^2})}\,. \end{aligned} \tag{2.6}$$

Notice that these components do *not* satisfy $\vec{u} = \vec{u}\,' + \vec{v}$, in other words, the Galilean addition of velocities is no longer true. However, for very small velocities, where $|\vec{u}| << c$ and $v << c$, we have that $\frac{u'_x v}{c^2} << 1$. Thus the denominators in (2.6) are very close to 1, and so addition of velocities is approximately true.

We can also show that if $|\vec{u}\,'| < c$ and $v < c$, then $|\vec{u}| < c$. To see this, consider the combination

$$c^2 - (\vec{u}\,')^2 = \frac{c^2(\Delta t')^2 - (\Delta x')^2 - (\Delta y')^2 - (\Delta z')^2}{(\Delta t')^2} > 0\,. \tag{2.7}$$

The numerator is the invariant in (1.57) and since $(\Delta t')^2 > 0$, the numerator must also be greater than zero. Therefore

$$c^2 - (\vec{u})^2 = \frac{c^2(\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2}{(\Delta t)^2} > 0\,. \tag{2.8}$$

2.7 Acceleration

Let us now consider the case of acceleration and how an observer in a frame different from the accelerating object would see this. The *proper acceleration*, $\vec{\alpha}$ is defined as the acceleration in the rest frame of the accelerated body. Now since this rest frame is accelerating with respect to the observer's rest frame, which we are assuming is an inertial frame, it means that the body's rest frame is *not* an inertial frame. However, at any time there is an inertial frame where the body is at rest. We call this inertial frame its *instantaneous rest frame*. This can also be called the *momentary rest frame*. At a later time, the instantaneous inertial frame is different from the one it is in now.

To avoid too many complications, let us simplify the problem a bit and assume that the velocities and accelerations are in the x direction only, so $\vec{\alpha} = \alpha\hat{x}$. We then let $\mathbf{S}$ be the frame of the observer and $\mathbf{S}'(t)$ be the instantaneous rest frame at time t. If $\vec{u} = u\,\hat{x}$ is the velocity of the body as seen by the observer, then $v = u$ at t. In $\mathbf{S}'(t)$ we have that $u' = 0$. At a slightly later time $t + dt$, where dt is assumed to be an infinitesimal displacement, we have that u' has an infinitesimal change, du', since $\mathbf{S}'(t)$ is no longer the instantaneous rest frame, $\mathbf{S}'(t+dt)$ is. But from the definition of the proper acceleration, we have that

$$du' = \alpha\, dt'\,. \tag{2.9}$$

Since $\mathbf{S}'(t)$ is the instantaneous rest frame, dt is related to dt' by time dilation:

$$dt = \gamma_u \, dt' \,, \qquad \gamma_u = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}} \,. \tag{2.10}$$

The change in velocity du as seen by the observer in $\mathbf{S}$ is then found using the velocity transformation in (2.6). Given the velocity in $\mathbf{S}'(t)$ is du' and the velocity in $\mathbf{S}$ is $u + du$ then the transformation formula leads to

$$du = \frac{du' + u}{1 + \frac{u\,du'}{c^2}} - u \approx (du' + u)\left(1 - \frac{u\,du'}{c^2}\right) - u \approx du'\left(1 - \frac{u^2}{c^2}\right) \,, \tag{2.11}$$

where we used that $|u| < c$ and $|du'| << c$ to justify the approximations.

The acceleration, a, as measured by the observer in $\mathbf{S}$ is then

$$a = \frac{du}{dt} = \frac{du'}{dt'}\left(1 - \frac{u^2}{c^2}\right)^{3/2} = \frac{\alpha}{{\gamma_u}^3} \,. \tag{2.12}$$

Notice that a is approaching 0 as u approaches c, which is reasonable since we should not be able to accelerate the body through the speed of light.

Let us take the further special case that the proper acceleration α is constant. We then observe that

$$\frac{d}{dt}(\gamma_u \, u) = \gamma_u \frac{du}{dt} + u \frac{d}{dt}\frac{1}{\sqrt{1 - \frac{u^2}{c^2}}} = \left(\gamma_u + \frac{u^2}{c^2}{\gamma_u}^3\right)\frac{du}{dt} = {\gamma_u}^3 \, \frac{du}{dt} \,. \tag{2.13}$$

Hence, we can write (2.12) as

$$\frac{d}{dt}(\gamma_u \, u) = \alpha \tag{2.14}$$

which has the solution

$$\gamma_u \, u = \alpha t + u_0 \tag{2.15}$$

where u_0 is a constant. If we take the initial condition that $u = 0$ at $t = 0$, then $u_0 = 0$. Squaring both sides of the above equation gives

$$\frac{u^2}{1 - \frac{u^2}{c^2}} = \alpha^2 t^2 \,, \tag{2.16}$$

which has the solution

$$u = \frac{\alpha t}{\sqrt{1 + \alpha^2 t^2 / c^2}} \,. \tag{2.17}$$

Now using $u = \frac{dx}{dt}$ we get

$$dx = \frac{\alpha t \, dt}{\sqrt{1 + \alpha^2 t^2 / c^2}} \,, \tag{2.18}$$

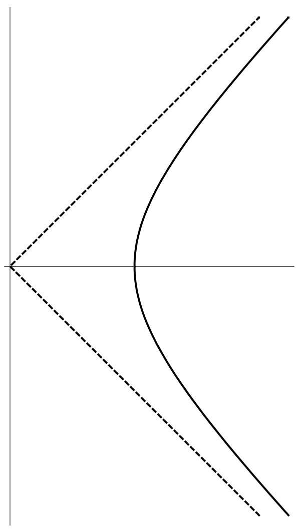


Figure 11: Graph of the hyperbolic trajectory in (2.20). The horizontal axis is the x coordinate and the vertical axis is the $x^0 = ct$ coordinate. The dashed lines are light-like lines that define the limits of the hyperbola. The hyperbola intersects the x-axis at $x = c^2/\alpha$.

which can be integrated to give

$$x = \frac{c^2}{\alpha}\sqrt{1 + \frac{\alpha^2 t^2}{c^2}} + x_0\,, \tag{2.19}$$

where x_0 is a constant. Since the constant is just a shift in x we drop it. Hence, this solution can be written as

$$x^2 - (ct)^2 = \frac{c^4}{\alpha^2}\,, \tag{2.20}$$

which is the equation for a hyperbola. This solution is graphed in figure 11, where the horizontal axis is the x axis and the vertical axis is the $x^0 = ct$ axis. The hyperbola crosses the x axis at $x = c^2/\alpha$. Notice that as t becomes large, the hyperbola approaches the light-like trajectory which is the dashed line in the plot. Interestingly, if at $t = 0$ a light ray is emitted at $x = 0$, we see that it will never catch up with the accelerating body since the dashed line never intersects the hyperbola.

2.8 Doppler shifts

If you are standing at a railroad crossing and an approaching train blows its whistle, the whistle will suddenly drop in pitch as the train passes. This phenomenon is due to the Doppler shift of the train's sound waves. When the train is moving toward us, the pitch is shifted upward, but when it is moving away it is shifted downward. We can then ask what will happen to lightwaves of wavelength λ, which is related to the frequency ν by $\lambda = c/\nu$.

Suppose we have a light source whose rest-frame is $\mathbf{S}'$ and is shining light at an observer in $\mathbf{S}$. We can think of the light source as a clock which is sending a signal at a regular time interval, $\Delta t' = 1/\nu$. In other words, this is the time for one wavelength of light to be emitted. The signal travels at the speed of light, c. This clock is time dilated in $\mathbf{S}$ to $\Delta t = \gamma \Delta t'$. But the question is more involved that just time dilation. The light being emitted is being observed by an observer at a fixed position in $\mathbf{S}$. We are *not* comparing the clock in $\mathbf{S'}$ with a series of clocks it passes in $\mathbf{S}$ as the light source moves along.

Let us say that at $t = 0$ the light source is at $x = 0$ moving with velocity v in the x direction. The observer is fixed at $x = 0$. If the light source sends a signal at $t = 0$ then the observer receives it instantaneously because the signal has zero distance to travel. The source then sends another signal at time $t = \Delta t$. But the light source is moving away from the observer and this signal is sent from position $x = v\Delta t$, and so the observer does not receive it immediately, but at a later time

$$t = \Delta T = \Delta t + \frac{v\Delta t}{c} = \gamma\, \Delta t' \left(1 + \frac{v}{c}\right) = \Delta t' \sqrt{\frac{1 + v/c}{1 - v/c}}\,. \tag{2.21}$$

Hence the time it takes the observer to see one wavelength is ΔT and so the light frequency as seen by the observer is

$$\nu_{\text{obs}} = \frac{1}{\Delta T} = \nu \sqrt{\frac{1 - v/c}{1 + v/c}} = \nu \sqrt{\frac{c - v}{c + v}}\,, \tag{2.22}$$

and the wavelength is

$$\lambda_{\text{obs}} = \frac{c}{\nu_{\text{obs}}} = \lambda \sqrt{\frac{c + v}{c - v}}\,. \tag{2.23}$$

Since the wavelength increases, we call this a *red-shift*.

To find the shift if the source is moving toward the observer, all we need to do is replace v with $-v$, hence we have

$$\lambda_{\text{obs}} = \lambda \sqrt{\frac{c - v}{c + v}}\,. \tag{2.24}$$

Since the wavelength is smaller, we call this a *blue-shift*.

It is instructive to look at the space-time diagram for this process. Figure 12 shows the world-line of the source as the $x^{0'}$ axis. Included are the world-lines for light signals sent a time $\Delta t'$ apart according to the source's clock and sent back toward the observer at $x = 0$. The diagram shows the difference between Δt and ΔT.

We can also generalize this to include a light source that is not moving directly away or toward the observer, but at an angle θ, as shown in figure 13. In this case, during the time interval Δt the source has moved a distance $v \cos\theta \Delta t$ away from the observer, so we have that ΔT is

$$\Delta T = \Delta t + \frac{v \cos\theta\, \Delta t}{c} = \gamma\, \Delta t' \left(1 + \frac{v \cos\theta}{c}\right)\,, \tag{2.25}$$

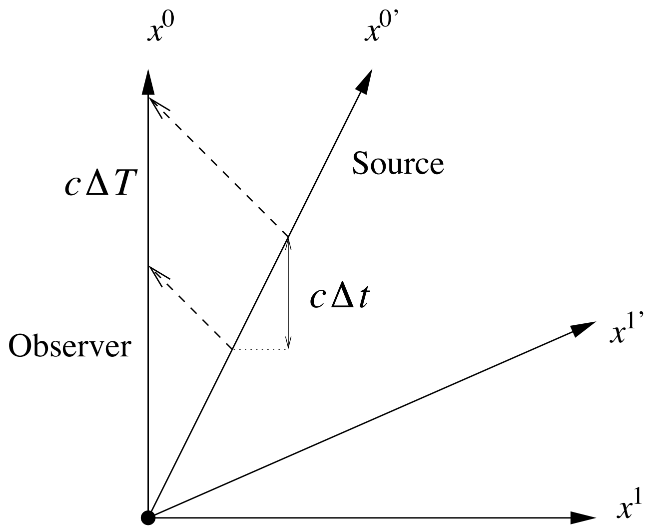


Figure 12: A space-time diagram for a light source emitting light toward an observer. The source's world-line is the $x^{0'}$ axis, while the observer's world-line is the x^0 axis.

and so the frequency and wavelength seen by the observer is

$$\begin{aligned}\nu_{\rm obs} = \frac{1}{\Delta T} &= \nu\,\frac{\sqrt{1-v^2/c^2}}{1+\frac{v}{c}\cos\theta} \\ \lambda_{\rm obs} &= \lambda\,\frac{1+\frac{v}{c}\cos\theta}{\sqrt{1-v^2/c^2}}\,. \end{aligned} \tag{2.26}$$

2.9 A little cosmology

Our universe is expanding. How do we know this? We can see distant galaxies being redshifted. The further away the galaxy the bigger the redshift. To describe the redshift, cosmologists define a quantity Z, given by

$$Z \equiv \frac{\lambda_{\rm obs}}{\lambda_{\rm galaxy}} - 1\,, \tag{2.27}$$

where $\lambda_{\rm galaxy}$ is the wavelength in the rest frame of the galaxy and $\lambda_{\rm obs}$ is the galaxy's wavelength as measured by an observer on earth. If the galaxy is redshifted, then $Z > 0$. Using the result for the redshift in (2.23), we have

$$Z = \sqrt{\frac{c+v}{c-v}} - 1\,. \tag{2.28}$$

We can then invert this equation to find the receding velocity of the galaxy given the value of Z:

$$\frac{c+v}{c-v} = (Z+1)^2 \qquad \Rightarrow \qquad v = \frac{Z(2+Z)}{2+Z(2+Z)}\,c\,. \tag{2.29}$$

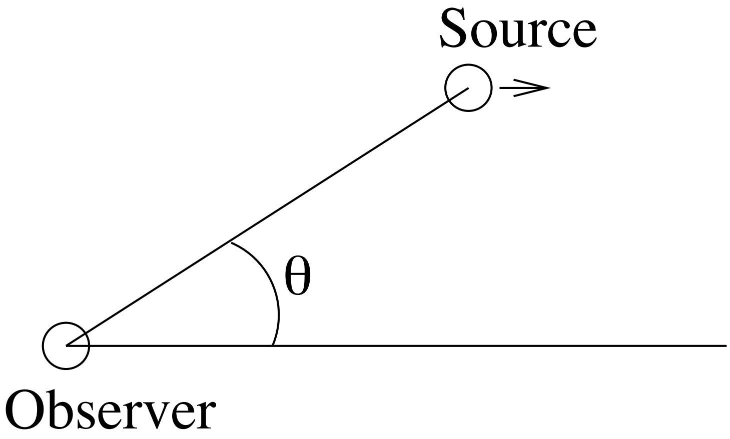


Figure 13: A light source moving away from an observer at an angle θ.

It has been observed that the the further away the galaxy, the faster it is receding. In fact the relation between the velocity and the distance is very close to linear, given by

$$v = H_0 D\,, \tag{2.30}$$

where D is the distance and H_0 is a constant called the *Hubble constant*. It turns out that it is much easier to measure v than D. But over the last ten or fifteen years, an effective way to measure D was found by not observing galaxies, but supernovae. A supernova is an exploding star and those of a certain type were found to have very uniform properties. So the brightness of the supernova during the course of its explosion (which could be observed for several weeks) could be used to judge its distance.

At present, the farthest away supernova had $Z = 1.7$, which translates into a velocity of $v = 0.76\,c$. The present value for H_0 is approximately $H_0 \approx 70$ km/s-Mpsc. Mpsc stands for Megaparsec which equals 3.26×10^6 light years, where a light year is the distance light travels in one year. Hence, the distance of this far away supernova is

$$\begin{aligned} D = \frac{v}{H_0} &= \frac{(0.76)(3\times 10^5 \text{ km/s})}{70 \text{ km/s}} \times 3.26 \times 10^6 \text{ light years} \\ &= 10.6 \times 10^9 \text{ light years}\,. \end{aligned} \tag{2.31}$$

This can be compared to the size of the visible universe which is 13.7×10^9 light years. Hence, this supernova was a significant distance away compared to the overall size of the visible universe.

2.10 The drag effect

An interesting phenomenon was known to 19th century physicists. They had observed that in a liquid moving with velocity v along the x direction, the speed of light along the same direction is

$$u = u' + kv\,, \tag{2.32}$$

where u' is the velocity when the fluid is at rest and

$$k = 1 - \frac{1}{n^2}\,, \tag{2.33}$$

where n is the index of refraction. Hence by definition, $u' = c/n$. In the middle of the century Fresnel came up with an ether based explanation for why this should happen, basically arguing that the fluid pushed the ether along with it.

But we can explain this using velocity transformations. We let $\mathbf{S}'$ be the rest frame of the liquid and $\mathbf{S}$ be the rest frame of the observer measuring the light's speed. We then assume that $v << c$, since afterall, we don't expect any liquids to be traveling anywhere near the speed of light. Using (2.6) and Taylor expanding to first order in v we have

$$u = \frac{u' + v}{1 + u'v/c^2} \approx (u' + v)\left(1 - \frac{u'v}{c^2}\right) \approx u' + v - \frac{u'^2 v}{c^2} = u' + v\left(1 - \frac{1}{n^2}\right)\,, \tag{2.34}$$

which is the result we are looking for. Incidentally, notice that the speed of the light is not the same in different reference frames. But this does not contradict Einstein's second postulate because the light is not traveling in a vacuum.

2.11 Aberration

Aberration is the effect that observers in different inertial frames can measure different angles for a light source. For this problem we can restrict the spatial dimensions to the x and y coordinates. We again assume that the two observers are in different frames $\mathbf{S}$ and $\mathbf{S}'$, and that $\mathbf{S}'$ is moving with velocity $\vec{v} = v\hat{x}$ *wrt* $\mathbf{S}'$. An observer at the spatial origin of $\mathbf{S}$ sees a light source at an angle θ away from the x-axis. Figure 14 shows a diagram of this. The observer in $\mathbf{S}'$ sees the same light source at a different angle θ'. The problem is to relate the two angles.

Again we can use the velocity transformations in (2.6) to find the relation. From figure 14 we see that the velocity components of the light coming from the source as seen by the observer in $\mathbf{S}$ are

$$u_x = -c\cos\theta \qquad u_y = -c\sin\theta\,, \tag{2.35}$$

while those seen by the observer in $\mathbf{S}'$ are

$$u'_x = -c\cos\theta' \qquad u'_y = -c\sin\theta'\,. \tag{2.36}$$

Substituting these expressions into (2.6) we find

$$-\cos\theta = \frac{-\cos\theta' + v/c}{1 - (v/c)\cos\theta'}\,, \qquad -\sin\theta = \frac{-\sin\theta'}{\gamma(1 - (v/c)\cos\theta')}\,. \tag{2.37}$$

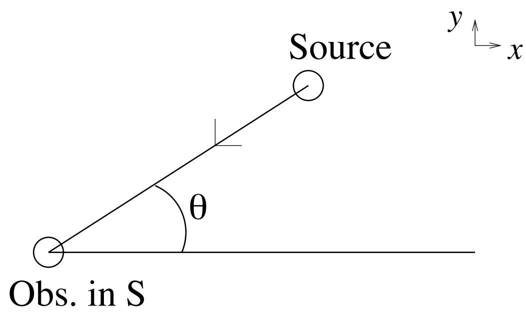


Figure 14: A light source at an angle θ as seen by an observer in **S**.

We now use the trigonometric identity

$$\tan \tfrac{1}{2}\theta = \frac{\sin\theta}{1+\cos\theta}\,. \tag{2.38}$$

Inserting the values of $\cos\theta$ and $\sin\theta$ in (2.37) into this identity gives

$$\begin{aligned} \tan \tfrac{1}{2}\theta &= \frac{\frac{\sin\theta'}{\gamma(1-(v/c)\cos\theta')}}{1+\frac{\cos\theta'-v/c}{1-(v/c)\cos\theta'}} = \frac{1}{\gamma}\,\frac{\sin\theta'}{1-(v/c)\cos\theta'+\cos\theta'-v/c} \\ &= \frac{1}{\gamma}\,\frac{1}{1-v/c}\,\frac{\sin\theta'}{1+\cos\theta'} = \sqrt{\frac{c+v}{c-v}}\,\tan\tfrac{1}{2}\theta'\,. \end{aligned} \tag{2.39}$$

We see that the observer who is moving faster toward the source (the one in $\mathbf{S}'$) sees a smaller angle. Notice that the source's velocity does not matter as far as the relation of the angles is concerned, although it will affect the absolute angles.

3 Tensors

Tensors are very useful because they have nice transformation properties under Lorentz transformations[3]. So if physical quantities can be written in terms of tensors, then we know how to find these quantities in different inertial frames. In fact, we can even generalize this to transformations from any frame (not necessarily an inertial frame) to any other frame, but we won't worry about that here. So if we can put physical quantities into tensor form, then we have a recipe for finding these quantities in any inertial frame.

Tensors can come with two types of indices. These are "upper" indices and "lower" indices. The distinction between the two is important because the two types of indices transform differently under Lorentz transformations.

Let us start with one of the simplest tensors which is one we have already introduced, the displacement "4-vector" Δx^μ. Since this has an upper index, we say that this is a *contravariant* vector. The upper index μ refers to one of 4 possible values 0, 1, 2 or 3. The 0 component is the component along the time direction. The other three components are called *spatial* components. The 4 components of Δx^μ can now be written as

$$\Delta x^\mu : \quad (\Delta x^0, \Delta x^1, \Delta x^2, \Delta x^3) = (c\Delta t, \Delta x, \Delta y, \Delta z)\,. \tag{3.1}$$

Notice the first coefficient when written with Δt has a factor of c so that the Δx^0 has units of length, just like the spatial components. We also sometimes write $\Delta x^\mu : \ (\Delta x^0, \Delta\vec{x})$. Lower case Greek letters are used for space-time indices, while lower case Roman letters (i, j, k etc.) are used for spatial indices only. The indices are associated with a particular inertial frame **S**, and the variables x^μ are known as the *coordinates* of that frame.

Now that we have a tensor, in this case a contravariant vector, let us transform it to a new frame. Suppose we start with the 4-vector A^μ in frame **S**. The goal is to find the 4-vector in the new frame $\mathbf{S}'$. This is done through a Lorentz transformation. There are 6 independent transformations. Three of these are the *boosts* along the three independent spatial directions. The boost is completely determined by the relative velocity $\vec{v}$ that $\mathbf{S}'$ is moving *with respect to* (wrt) **S**. The other 3 independent transformations are *spatial rotations.* A rotation takes place in a plane, so the independent transformations are the 3 ways of choosing 2 spatial directions, (x, y), (y, z) and (z, x). Along with the plane, we should also specify the angle through which we rotate. In any case, any Lorentz transformation can be made with some combination of these types of transformations.

A Lorentz transformation can be given by a matrix $\Lambda^{\mu'}{}_\mu$. The index μ' refers to the index in the frame $\mathbf{S}'$ and μ refers to the index in the frame **S**. You should not assume that the two indices are the same (for instance, we could have μ be 0 while μ' is $1'$). We can use this matrix to transform our contravariant 4-vector. A typical example for $\Lambda^{\mu'}{}_\mu$ is a boost in the x direction

$$\Lambda^{\mu'}{}_\mu = \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{3.2}$$

[3]The contents in this section have appeared previously in "Tensors without Tears", so parts of this section are repetitive with earlier sections of the notes.

where v is the velocity in the x direction and γ is

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}} \,. \tag{3.3}$$

Another example is the rotation in the $x - y$ plane

$$\Lambda^{\mu'}{}_{\mu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \tag{3.4}$$

while a general form for a rotation is

$$\Lambda^{\mu'}{}_{\mu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & R_{11} & R_{12} & R_{13} \\ 0 & R_{21} & R_{22} & R_{23} \\ 0 & R_{31} & R_{32} & R_{32} \end{pmatrix} , \tag{3.5}$$

The index μ' refers to the row of the matrix (with $0'$ corresponding to the first row) and μ refers to the column. So for example, the component $\Lambda^{0'}{}_1$ refers to the entry in the first row and second column of the matrix, which is $-\frac{v}{c}\gamma$ in (3.2) and 0 in (3.5). We can now express our 4-vector in $\mathbf{S}'$ as

$$A^{\mu'} = \Lambda^{\mu'}{}_{\mu} A^{\mu} \equiv \sum_{\mu=0}^{3} \Lambda^{\mu'}{}_{\mu} A^{\mu} \,. \tag{3.6}$$

We notice that the index μ in (3.6) is repeated, with it appearing once with the index down and once with the index up. When you see such a repeated index, you should assume that it is summed over the 4 components of space-time. A repeated index is also called a *dummy* index. The index μ' in (3.6) is not repeated. Furthermore, note that it appears on both the left and the right hand side of the equation, and in both cases, it is up. An index that is not repeated is called a *free index.* The free indices need to match exactly on the left and right hand sides of the equations, otherwise the equation is nonsense. This is not the case for dummy indices. We can think of the equation in (3.6) as a matrix multiplying a vector. In other words, we can write the 4-vector as

$$A^{\mu} = \begin{pmatrix} A^0 \\ A^1 \\ A^2 \\ A^3 \end{pmatrix} \,. \tag{3.7}$$

Then the transformation can be written as

$$A^{\mu'} = \begin{pmatrix} A^{0'} \\ A^{1'} \\ A^{2'} \\ A^{3'} \end{pmatrix} = \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} A^0 \\ A^1 \\ A^2 \\ A^3 \end{pmatrix} \,. \tag{3.8}$$

Summing over the repeated index μ corresponds to the usual sum that appears in matrix multiplication. So for example, $A^{1'}$ is found by taking the second row of Λ and multiplying the entry in column μ with the entry in row μ of A^μ. More explicitly, using the transformation in (3.2),

$$\begin{aligned} A^{1'} &= \Lambda^{1'}{}_0 A^0 + \Lambda^{1'}{}_1 A^1 + \Lambda^{1'}{}_2 A^2 + \Lambda^{1'}{}_3 A^3 \\ &= -\frac{v}{c}\gamma A^0 + \gamma A^1 + 0 \cdot A^2 + 0 \cdot A^3 = \gamma(-\frac{v}{c}A^0 + A^1)\,. \end{aligned} \tag{3.9}$$

Now we should also be able to make a transformation from $\mathbf{S}'$ back to $\mathbf{S}$. Since $\mathbf{S}$ is moving with velocity $-v$ with respect to $\mathbf{S}'$, we should get the correct transformation by replacing v with $-v$ in our transformation. Hence we have

$$\Lambda^\mu{}_{\mu'} = \begin{pmatrix} \gamma & +\frac{v}{c}\gamma & 0 & 0 \\ +\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\,. \tag{3.10}$$

Then we can write A^μ as

$$A^\mu = \Lambda^\mu{}_{\mu'} A^{\mu'}\,. \tag{3.11}$$

Now if we first transform from $\mathbf{S}$ to $\mathbf{S}'$ and then transform back again, that is the same as doing nothing. If we put together our transformations we have

$$A^\mu = \Lambda^\mu{}_{\mu'} A^{\mu'} = \Lambda^\mu{}_{\mu'} \Lambda^{\mu'}{}_\nu A^\nu\,. \tag{3.12}$$

The right hand side should equal the left hand side, so we see that

$$\Lambda^\mu{}_{\mu'} \Lambda^{\mu'}{}_\nu = \delta^\mu{}_\nu, \tag{3.13}$$

where $\delta^\mu{}_\nu$ is the identity matrix with

$$\delta^\mu{}_\nu = 1 \;\text{ if } \mu = \nu, \qquad \delta^\mu{}_\nu = 0 \;\text{ if } \mu \neq \nu, \tag{3.14}$$

in other words, $\delta^\mu{}_\nu A^\nu = A^\mu$. One can also check that $\Lambda^\mu{}_{\mu'}$ is the inverse matrix of $\Lambda^{\mu'}{}_\mu$. Notice that the distinction is made between the matrix and its inverse depending on whether the primed index is up or down.

The second type of tensor we can consider is a *covariant* 4-vector, B_μ. As you can see, the index is down in this case. It is important to distinguish between the two types of indices, because they transform differently. The transformation for B_μ from $\mathbf{S}$ to $\mathbf{S}'$ is given by

$$B_{\mu'} = \Lambda^\mu{}_{\mu'} B_\mu\,, \tag{3.15}$$

Notice that in this transformation, the inverse transformation is used to go from $\mathbf{S}$ to $\mathbf{S}'$. Likewise, the transformation from $\mathbf{S}'$ back to $\mathbf{S}$ is

$$B_\mu = \Lambda^{\mu'}{}_\mu B_{\mu'} \tag{3.16}$$

We can also combine 4-vectors into more general types of tensors. For example, we can create an object $T^{\mu\nu} = A^\mu C^\nu$ out of the two contravariant 4-vectors A^μ and C^ν. Each 4-vector comes with its own index, so $T^{\mu\nu}$ naturally comes with two indices. Clearly,

$$T^{\mu'\nu'} = \Lambda^{\mu'}{}_{\mu} A^\mu \Lambda^{\nu'}{}_{\nu} C^\nu = \Lambda^{\mu'}{}_{\mu} \Lambda^{\nu'}{}_{\nu} T^{\mu\nu} \,. \tag{3.17}$$

We can also combine a contravariant and a covariant vector to form $T^\mu{}_\nu = A^\mu B_\nu$. This then transforms as

$$T^{\mu'}{}_{\nu'} = \Lambda^{\mu'}{}_{\mu} A^\mu \Lambda^{\nu}{}_{\nu'} B_\nu = \Lambda^{\mu'}{}_{\mu} \Lambda^{\nu}{}_{\nu'} T^\mu{}_\nu \,, \tag{3.18}$$

where the up index transforms with the Lorentz transformation matrix and the down index with inverse matrix. We can combine even more contravariant or covariant vectors to make a tensor with even more indices. In fact we can drop the condition that the tensor we have constructed is a product of vectors. In general we will define an $\begin{pmatrix} n \\ m \end{pmatrix}$ tensor as having n up indices and m down indices. Each up index transforms with a Lorentz transformation matrix and the down indices with the inverse. So we would have for a general tensor

$$T^{\mu'_1\mu'_2\ldots\mu'_n}_{\nu'_1\nu'_2\ldots\nu'_m} = \Lambda^{\mu'_1}{}_{\mu_1} \Lambda^{\mu'_2}{}_{\mu_2} \ldots \Lambda^{\mu'_n}{}_{\mu_n} \Lambda^{\nu_1}{}_{\nu'_1} \Lambda^{\nu_2}{}_{\nu'_2} \ldots \Lambda^{\nu_m}{}_{\nu'_m} T^{\mu_1\mu_2\ldots\mu_n}_{\nu_1\nu_2\ldots\nu_m} \,, \tag{3.19}$$

where all repeated indices are summed over. Note that these multiple copies of Lorentz transformations should *not* be thought of as multiplying many matrices together using the usual rules of matrix multiplication. In matrix multiplication, say where we multiply one 4×4 matrix M_1 with another 4×4 matrix M_2 to get a third matrix $M_3 = M_1 M_2$, the way the indices would be arranged are (for example)

$$M_3{}^{\mu'}{}_{\nu'} = M_1{}^{\mu'}{}_{\nu} M_2{}^{\nu}{}_{\nu'} \,. \tag{3.20}$$

There is a repeated index connecting M_1 and M_2 that corresponds to taking the column from M_1 and pairing it with the row of M_2. But in (3.19), no Lorentz matrix has a common index with another Lorentz matrix.

However, matrix multiplication is useful for thinking about multiple Lorentz transformations. So suppose that we first transform from $\mathbf{S}$ to $\mathbf{S}'$ and then transform to $\mathbf{S}''$. We should be able to write this as a single transformation from $\mathbf{S}$ to $\mathbf{S}''$. The way this works is through matrix multiplication, namely

$$\Lambda^{\mu''}{}_{\mu'} \Lambda^{\mu'}{}_{\mu} = \Lambda^{\mu''}{}_{\mu} \,, \tag{3.21}$$

The index μ' corresponding to the $\mathbf{S}'$ frame is repeated and the resulting matrix is the Lorentz transformation matrix that would be used in (3.19) if we wanted to transform $T^{\mu_1\mu_2\ldots\mu_n}_{\nu_1\nu_2\ldots\nu_m}$ to $\mathbf{S}''$.

One thing we should stress and that you should keep in mind is that the Lorentz transformation matrix is *not* a two index tensor. This is because each index is associated with a different frame, while a general two index tensor would have both indices associated with the same frame.

Now suppose that we combine a contravariant vector and a covariant vector as $A^\mu B_\mu$. The index μ is repeated, so it means that we should sum over it. Setting the indices equal like this is called *contracting* an index. Let us now see what this quantity is in $\mathbf{S}'$. When we have more than one index, each one is transformed with a Lorentz transformation matrix. Given these transformation properties, we find

$$A^{\mu'} B_{\mu'} = \Lambda^{\mu'}{}_{\nu} A^{\nu} \Lambda^{\mu}{}_{\mu'} B_{\mu} = \Lambda^{\mu}{}_{\mu'} \Lambda^{\mu'}{}_{\nu} A^{\nu} B_{\mu} \,. \tag{3.22}$$

In this last step we just rearranged the order of the terms in the equation. These are just numbers so we are allowed to do this. Now the $\Lambda^{\mu}{}_{\mu'}\Lambda^{\mu'}{}_{\nu}$ is precisely what appears in (3.13), so we can replace it by $\delta^{\mu}{}_{\nu}$, so we find

$$A^{\mu'} B_{\mu'} = \delta^{\mu}{}_{\nu} A^{\nu} B_{\mu} = A^{\mu} B_{\mu} \,. \tag{3.23}$$

Observe that $\delta^{\mu}{}_{\nu}$ is nonzero only when μ is the same as ν, and in this case it equals 1. Hence summing over ν will end up replacing the ν in A^{ν} with a μ. However, since both ν and μ are dummy indices, we could have equally replaced μ with ν. So you can see that you are free to relabel the dummy indices in which they appear. In other words

$$A^{\mu} B_{\mu} = A^{\nu} B_{\nu} \,. \tag{3.24}$$

We can see from the equation in (3.23) that the quantity $A^{\mu} B_{\mu}$ does not change when you transform to another frame.

A quantity that does not change under Lorentz transformations is known as a *Lorentz invariant*, or a *Lorentz scalar* or sometimes just *invariant* or *scalar*. The multiplication of two 4-vectors to make a Lorentz scalar is called a *scalar product*, or also *inner product*. This expression has no free indices, so it is like having a tensor T with no indices. According to the rules in (3.19), we would have the equation $T' = T$ where the left hand side refers to the tensor in $\mathbf{S}'$. In other words, it is an invariant.

An important tensor is the two index tensor $\eta_{\mu\nu}$. This is defined as[4]

$$\eta_{\mu\nu} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}_{\mu\nu} \,. \tag{3.25}$$

We can also write this as $\eta_{\mu\nu} = \mathrm{diag}(+1,-1,-1,-1)_{\mu\nu}$. Under the Lorentz transformation

$$\eta_{\mu'\nu'} = \Lambda^{\mu}{}_{\mu'} \Lambda^{\nu}{}_{\nu'} \eta_{\mu\nu} = \Lambda^{T\,\mu}_{\mu'} \eta_{\mu\nu} \Lambda^{\nu}{}_{\nu'} \,, \tag{3.26}$$

where we define $\Lambda^{T\,\mu}_{\mu'} \equiv \Lambda^{\mu}{}_{\mu'}$. In the last step in (3.26) we have arranged the indices so that it has the form of multiplying 3 matrices together. By rewriting the first Lorentz transformation this way we are transposing the rows and the columns, but are otherwise doing nothing else. If $\Lambda^{\mu'}{}_{\mu}$ is the general rotation in (3.5) then the matrix equation in (3.26) becomes

$$\Lambda^{T\,\mu}_{\mu'} \eta_{\mu\nu} \Lambda^{\nu}{}_{\nu'} = \left(\Lambda^{T} \eta \Lambda\right)_{\mu'\nu'} = \left(\eta \Lambda^{T} \Lambda\right)_{\mu'\nu'} \,, \tag{3.27}$$

[4] Rindler uses the symbol g instead of η.

where Λ^T is the transpose of Λ. Now for a rotation, it turns out that $\Lambda^T = \Lambda^{-1}$. Therefore, we have for rotations

$$\eta_{\mu'\nu'} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}_{\mu'\nu'} . \tag{3.28}$$

If $\Lambda^{\mu'}{}_{\mu}$ corresponds to the boost in (3.2), then $\Lambda^T = \Lambda$ and we find

$$\begin{aligned} \eta_{\mu'\nu'} &= \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ +\frac{v}{c}\gamma & -\gamma & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} \gamma^2\left(1-\frac{v^2}{c^2}\right) & 0 & 0 & 0 \\ 0 & -\gamma^2\left(1-\frac{v^2}{c^2}\right) & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}_{\mu'\nu'} \end{aligned} \tag{3.29}$$

We would get the same result if we were to boost in any other direction because of the rotational symmetry of $\eta_{\mu\nu}$. Therefore, $\eta_{\mu'\nu'}$ has the same form as $\eta_{\mu\nu}$. But it is important to note that $\eta_{\mu\nu}$ is *not* a Lorentz invariant. This is because its indices changed under the transformation.

We can use the η-tensor to change contravariant indices to covariant indices. So for example consider the transformation of $A^\mu\eta_{\mu\nu}$

$$\begin{aligned} A^{\mu'}\eta_{\mu'\nu'} &= \Lambda^{\mu'}{}_{\mu}A^\mu\Lambda^\lambda{}_{\mu'}\Lambda^\nu{}_{\nu'}\eta_{\lambda\nu} = \Lambda^\nu{}_{\nu'}\Lambda^\lambda{}_{\mu'}\Lambda^{\mu'}{}_{\mu}A^\mu\eta_{\lambda\nu} \\ &= \Lambda^\nu{}_{\nu'}\delta^\lambda{}_{\mu}A^\mu\eta_{\lambda\nu} = \Lambda^\nu{}_{\nu'}A^\mu\eta_{\mu\nu} . \end{aligned} \tag{3.30}$$

Hence $A^\mu\eta_{\mu\nu}$ transforms as if there is only the down index ν. Therefore, we can define a covariant 4-vector from a contravariant vector by $A_\nu \equiv A^\mu\eta_{\mu\nu}$. This process is known as *lowering* an index. We can also consider the inverse of $\eta_{\mu\nu}$ with two raised indices, such that $\eta^{\mu\lambda}\eta_{\lambda\nu} = \delta^\mu{}_\nu$. But we can also think of this as $\eta^\mu{}_\nu$ (in other words, $\eta^\mu{}_\nu = \delta^\mu{}_\nu$) where the η-tensor with the down indices lowered one of the indices of $\eta^{\mu\nu}$. We can then use $\eta^{\mu\nu}$ to raise indices; $A^\mu = \eta^{\mu\nu}A_\nu$. Obviously, this is called *raising* an index. We can raise or lower several indices by using more than one η-tensor. If we go back to our invariant $A^\mu B_\mu$, we now see that we can write this as

$$A^\mu B_\mu = A^\mu B^\nu \eta_{\mu\nu} = A_\nu B^\nu = A_\mu B^\mu \tag{3.31}$$

In other words, if we have a repeated index with one raised and the other lowered, we can switch the lowered and raised indices. Notice that the last step is just a relabeling of the dummy index. We could have used any greek letter we like, although not one that is already being used as a free index. Also, it is not wise to use the same dummy index within a product of tensors, say like, $T^\mu{}_\mu R^\mu{}_\mu$ since it is not clear which index is being contracted with which. In general $T^\mu{}_\mu R^\nu{}_\nu \neq T^\mu{}_\nu R^\nu{}_\mu$.

Just to be clear about which combinations of tensors are equal to each other, the following set of equalities holds

$$T^{\mu\nu} R_\mu{}^\lambda = T^{\rho\nu} R_\rho{}^\lambda = T_\mu{}^\nu R^{\mu\lambda}\,. \tag{3.32}$$

Notice that the free indices ν and λ stay the same.

Sometimes a tensor has some extra symmetry. For example, we can have the symmetric tensor $T^{\mu\nu} = T^{\nu\mu}$. It is easy to show that this is symmetric in all inertial frames, to wit

$$T^{\mu'\nu'} = \Lambda^{\mu'}{}_\mu \Lambda^{\nu'}{}_\nu T^{\mu\nu} = \Lambda^{\mu'}{}_\mu \Lambda^{\nu'}{}_\nu T^{\nu\mu} = \Lambda^{\nu'}{}_\nu \Lambda^{\mu'}{}_\mu T^{\nu\mu} = T^{\nu'\mu'}\,. \tag{3.33}$$

Likewise for an antisymmetric tensor $G^{\mu\nu} = -G^{\nu\mu}$, this too is antisymmetric in all frames:

$$A^{\mu'\nu'} = \Lambda^{\mu'}{}_\mu \Lambda^{\nu'}{}_\nu A^{\mu\nu} = -\Lambda^{\mu'}{}_\mu \Lambda^{\nu'}{}_\nu A^{\nu\mu} = -\Lambda^{\nu'}{}_\nu \Lambda^{\mu'}{}_\mu A^{\nu\mu} = -A^{\nu'\mu'}\,. \tag{3.34}$$

One other useful property of tensors is that if all components are zero in one frame, then all components are zero in any other frame. This is clear from the linear transformation in (3.19), where every term on the right hand side has one component of the tensor in frame **S**. Hence the sum of the terms is also zero, so every component in $\mathbf{S}'$ is zero.

4 Particular tensors

In this section we will construct particular tensors corresponding to physical quantities.

4.1 Invariant length and proper time

The displacement vector Δx^μ has already been mentioned as an example of a contravariant 4-vector. From this and the η-tensor we can construct a Lorentz invariant we discussed previously, $-\Delta s^2$,

$$\begin{aligned} -\Delta s^2 &= \Delta x^\mu \Delta x^\nu \eta_{\mu\nu} = \Delta x^\mu \Delta x_\mu = \Delta x_\mu \Delta x^\mu \\ &= \Delta x^0 \Delta x^0 - \Delta x^1 \Delta x^1 - \Delta x^2 \Delta x^2 - \Delta x^3 \Delta x^3 \\ &= c^2 \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 = c^2 \Delta t^2 - \Delta \vec{x} \cdot \Delta \vec{x} \\ &= c^2 \Delta \tau^2 \,. \end{aligned} \tag{4.1}$$

As you can see, I am trying to make the point that all of the expressions in the above equation are the same. This invariant is loosely speaking a length squared, although not exactly because of the minus signs. Instead one can think of it as the square of the *proper time* τ between two events, multiplied by a factor of c^2 to get the dimensions right.

Since the tensor $\eta_{\mu\nu}$ is used to construct a length squared, it is often called the *metric*. In particular it is the metric of a particular type of space known as *Minkowski* space. Minkowski space is a flat space, meaning that it is not curved, with the metric given in (3.25) and in particular, one element on the diagonal has the opposite sign of the other elements[5].

Since Δs^2 is an invariant, then its sign is also invariant. If $\Delta s^2 < 0$, then we call the displacement vector *time-like*, while if $\Delta s^2 > 0$ we call it *space-like*. If $\Delta s^2 = 0$, then the vector is *light-like*.

Instead of displacements, we can also consider differentials dx^μ and construct an invariant out of these

$$ds^2 = c^2 d\tau^2 = dx^\mu dx^\nu \eta_{\mu\nu} = dx^\mu dx_\mu \tag{4.2}$$

The differentials dx^μ are themselves contravariant 4-vectors and so transform as

$$dx^{\mu'} = \Lambda^{\mu'}{}_{\mu} dx^\mu \tag{4.3}$$

But we also know from elementary calculus that when one changes variables, the differentials of the new variables (in this case $x^{\mu'}$) are related to the differentials of the old variables (x^μ) by

$$dx^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu} dx^\mu \tag{4.4}$$

Comparing these two equations, we see that

$$\frac{\partial x^{\mu'}}{\partial x^\mu} = \Lambda^{\mu'}{}_{\mu} \,, \tag{4.5}$$

[5] Another example of a metric is $g_{\mu\nu} = \text{diag}(+1, +1, +1, +1)$, which is the metric for four-dimensional flat Euclidean space, where clearly one has $\Delta s^2 = g_{\mu\nu} \Delta x^\mu \Delta x^\nu$. With this difference in signs, we say that the Euclidean metric has a different *signature* than the Minkowski metric.

a result we saw in a previous lecture. It also follows that by transforming back that

$$\frac{\partial x^{\mu}}{\partial x^{\mu'}} = \Lambda^{\mu}{}_{\mu'} \,. \tag{4.6}$$

4.2 Derivatives are covariant vectors

We can now see that there is a natural covariant vector that we can build. Consider a scalar function $f(x^\nu)$ which depends on the coordinates. Since this function is a scalar, when it is transformed to a new frame, the function becomes

$$f'(x^{\mu'}) = f(x^{\mu}) \tag{4.7}$$

Notice that $f'(x^\mu) \neq f(x^\mu)$. The function is different, but it is different so that the *new* function as a function of the *new* variables is equal to the *old* function as a function of the *old* variables. We can then consider a derivative of $f(x^\mu)$ with respect to one of the coordinates x^ν

$$\frac{\partial}{\partial x^{\nu}} f(x^{\mu}) \,. \tag{4.8}$$

Now change variables, in other words, compute the derivative of the new variables acting on the new function. We then find

$$\frac{\partial}{\partial x^{\nu'}} f'(x^{\mu'}) = \frac{\partial x^{\nu}}{\partial x^{\nu'}} \frac{\partial}{\partial x^{\nu}} f(x^{\mu}) = \Lambda^{\nu}{}_{\nu'} \frac{\partial}{\partial x^{\nu}} f(x^{\mu}) \,. \tag{4.9}$$

In other words $\frac{\partial}{\partial x^\nu} f(x^\mu)$ transforms as a covariant 4-vector. In fact, it does not matter what scalar function the derivative is acting on, it transforms as a covariant 4-vector. We can also generalize this to a derivative acting on a $\begin{pmatrix} n \\ m \end{pmatrix}$ tensor function, $\frac{\partial}{\partial x^\lambda} T^{\mu_1\mu_2\ldots\mu_n}_{\nu_1\nu_2\ldots\nu_m}$ where we see that the new object is now a $\begin{pmatrix} n \\ m+1 \end{pmatrix}$ tensor with an extra down index λ. Since the derivative adds a covariant index, it is common to write this as

$$\partial_\nu \equiv \frac{\partial}{\partial x^{\nu}} \,. \tag{4.10}$$

Another common terminology is to write the derivative of a tensor as

$$T^{\mu_1\mu_2\ldots\mu_n}_{\nu_1\nu_2\ldots\nu_m,\lambda} = \frac{\partial}{\partial x^{\lambda}} T^{\mu_1\mu_2\ldots\mu_n}_{\nu_1\nu_2\ldots\nu_m} \,, \tag{4.11}$$

where the comma (,) indicates that this index is coming from a derivative. If you go on to study general relativity, you will see that these ideas have to be modified somewhat, but in a very interesting way.

4.3 4-velocity

We can see from (4.2) that $d\tau$, the differential of the proper time, is an invariant[6]. Since dx^μ is a 4-vector, then so is

$$u^\mu \equiv \frac{dx^\mu}{d\tau}\,, \tag{4.12}$$

which we call the *velocity 4-vector*, or simply *4-velocity*. In the rest frame of the particle, $d\tau = dt$, while $dx^i = 0$. Hence, in this frame $u^\mu = (c, 0, 0, 0)$. We can then exploit the fact that $u^\mu u_\mu$ is an invariant, and so

$$u^\mu u_\mu = c^2 \tag{4.13}$$

in all frames. In a frame where the particle is moving with velocity $\vec{v}$, we have that $dt = \gamma(\vec{v})d\tau$ by time dilation. Therefore

$$\frac{dx^0}{d\tau} = \frac{d(ct)}{d\tau} = c\gamma \qquad \frac{d\vec{x}}{d\tau} = \gamma\frac{d\vec{x}}{dt} = \gamma\vec{v}, \tag{4.14}$$

and so

$$u^\mu = \big(\gamma(v)c, \gamma(v)\vec{v}\,\big)\,. \tag{4.15}$$

Note that since $u^\mu u_\mu > 0$, the 4-velocity is time-like. Notice further that the 4-velocity is a tangent vector for the particle's world-line, that is its space-time trajectory (see Figure 15).

The above construction depends on the existence of a rest frame. For light rays, there is clearly no rest frame, since the speed of light has to be the same in all inertial frames. Instead, for light we can define the 4-velocity as

$$u^\mu = \frac{dx^\mu(\lambda)}{d\lambda}\,, \tag{4.16}$$

where λ parameterizes the light's trajectory and $x^\mu(\lambda)$ is the space-time position of the light as a function of λ. We can use any parameterization we choose, so long as λ is always increasing along the entire trajectory. But however we choose λ, it will always be the case that $u^\mu u_\mu = 0$, that is the 4-velocity is light-like. Furthermore, u^μ is a tangent vector for the light's world-line.

4.4 4-acceleration

Now that we have the 4-velocity, it is not hard to see how to find the acceleration 4-vector, or more simply called the *4-acceleration*. We just take the derivative of u^μ with respect to the proper time, namely

$$a^\mu = \frac{du^\mu}{d\tau}\,. \tag{4.17}$$

[6]Why should the differential of the proper time be an invariant? The answer is that independent of what frame we start in, we always find the proper time by going to one particular frame, the particle's rest frame.

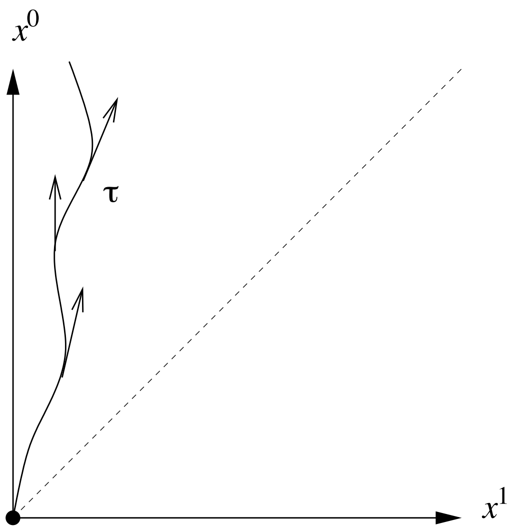


Figure 15: A world-line of a particle parameterized by its proper time τ. The tangent vectors are the 4-velocity vectors.

In the instantaneous rest frame of the particle we have that $u^\mu = (c, 0, 0, 0)$ at a time t, where t is the time coordinate in the instantaneous rest-frame. After an infinitesimal time $dt = d\tau$ elapses, the new 4-velocity in this frame is $du^\mu = (c\gamma(d\vec{v}) - c, \gamma(d\vec{v})d\vec{v})$. But $\gamma(d\vec{v}) = 1 + \mathrm{O}((d\vec{v})^2)$. Thus, to lowest order we have that $a^\mu = (0, \vec{\alpha})$ in the instantaneous rest frame, where

$$\vec{\alpha} = \frac{d\vec{v}}{dt} \tag{4.18}$$

and where $\vec{\alpha}$ is the *proper acceleration*. Again, constructing the invariant, we see that $a^\mu a_\mu = -(\vec{\alpha})^2 < 0$. Hence, the 4-acceleration is space-like (again assuming there is a rest-frame). We can also see that by going to the rest-frame that $a^\mu u_\mu = 0$, hence the 4-acceleration is a normal vector to the particle world-line.

Figure 16 shows a particle with constant proper acceleration in the x direction. The tangents to the world-line are the 4-velocities. The vectors that are normal to the 4-velocity are the 4-acceleration. Notice that the 4-acceleration and 4-vectors don't look orthogonal, but they really are because the inner product of the vectors is using the relative signs between the time and spatial components. Notice further that the arrow lengths are getting longer as the velocity gets closer to the speed of light because of the γ factors that appear in the 4-velocity.

For light we can define the 4-acceleration as

$$a^\mu = \frac{du^\mu}{d\lambda} = \frac{d^2x^\mu}{d\lambda^2}\,. \tag{4.19}$$

We still have that $a^\mu u_\mu = 0$. To see this, we can take a derivative on $u^\mu u_\mu$, to get

$$\frac{d}{d\lambda}(u^\mu u_\mu) = a^\mu u_\mu + u^\mu a_\mu = 2a^\mu u_\mu\,. \tag{4.20}$$

However $u^\mu u_\mu$ is a constant, hence its derivative is zero so $a^\mu u_\mu = 0$. Note that we could have used the same argument for time-like trajectories, just using τ instead of λ.

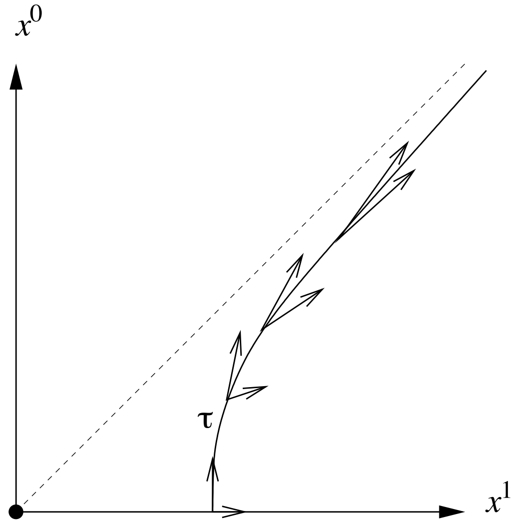


Figure 16: A world-line of a particle with constant proper acceleration. The tangent vectors are the 4-velocity vectors, while the normal vectors are the 4-acceleration vectors.

4.5 Wave 4-vector

Consider a wave-train (the waves could be water waves or sound waves or even light waves) given by an amplitude $A(\Phi(t,\vec{x}))$, where $\Phi(t,\vec{x})$ is the phase of the wave which satisfies the periodicity condition

$$A(\Phi(t,\vec{x})+2\pi)=A(\Phi(t,\vec{x}))\,. \tag{4.21}$$

Let us write the space-time coordinates in the 4-vector form, $\Phi(t,\vec{x})=\Phi(x^\mu)$. The phase is a Lorentz scalar, in other words if we go to a new frame, the new phase is related to the old phase by

$$\Phi'(x^{\mu'})=\Phi(x^\mu)\,. \tag{4.22}$$

Notice that the function Φ' is different from Φ, that is, $\Phi'(x^\mu)\neq\Phi(x^\mu)$ (observe the difference between this statement and the one in (4.22)).

The angular frequency ω is found from the phase by taking its time derivative, giving

$$\omega=\frac{\partial\Phi}{\partial t}=c\,\frac{\partial\Phi}{\partial x^0} \tag{4.23}$$

while the wave-vector $\vec{k}$ is

$$\vec{k}=-\vec{\nabla}\Phi\,. \tag{4.24}$$

Note that the wave-vector is related to the wavelength λ by $|\vec{k}|=2\pi/\lambda$ and the angular frequency is related to the frequency ν by $\omega=2\pi\nu$. We also note that the *dispersion relation* for the wave is

$$\omega=v_w|\vec{k}| \tag{4.25}$$

where v_w is the speed of the wave. Since Φ is a scalar, if we compare (4.23) and (4.24) with (4.8) we see that we can express everything in terms of a covariant 4-vector k_μ whose components are

$$k_\mu=\left(\frac{\omega}{c},-\vec{k}\right)=\partial_\mu\Phi\,. \tag{4.26}$$

It is very important that we write the index for k_μ as down. Otherwise we would find the wrong transformations when going to a different inertial frame. Solving (4.26), we find

$$\Phi(x^\mu) = \omega t - \vec{k}\cdot\vec{x} + \Phi_0 = k_\mu x^\mu + \Phi_0 \,. \tag{4.27}$$

where Φ_0 is a constant.

We can also see that k_μ must be either space-like or light-like. The scalar product of k_μ with itself gives

$$k_\mu k^\mu = k_\mu k_\nu \eta^{\mu\nu} = \frac{\omega^2}{c^2} - \vec{k}\cdot\vec{k} = \left(\frac{v_w^2}{c^2} - 1\right)\,\vec{k}\cdot\vec{k} \le 0\,, \tag{4.28}$$

since $v_w \le c$ and $\vec{k}\cdot\vec{k}$ is positive definite. Obviously for light-waves k_μ is light-like. Recall from quantum mechanics that there is a particle wave duality, so light waves can also be thought of as a collection of photons. Under this duality, it turns out that the wave 4-vector of the light wave is parallel to the 4-velocity of the photon.

In figure 17 we show the space-time diagram for a wave-train moving in the x direction. The lines correspond to fixed values of Φ, in this case multiples of 2π. We can find these lines using the solution in (4.27) with $\Phi_0 = 0$ ($\Phi_0 \ne 0$ shifts the wave-fronts by a constant amount). We then find that

$$ct = \frac{ck}{\omega}x - 2\pi n\,\frac{c}{\omega}, \tag{4.29}$$

with n an integer. Using the dispersion relation in (4.25), we see that the slope of the line is given by $c/v_w \ge 1$. Since the wave 4-vector is $k_\mu = \partial_\mu\Phi$, k_μ is a normal vector to the wave-front (along the wave-front, Φ is fixed, while along the k_μ direction, Φ is changing.)

With the wave 4-vector we have a simple way of finding the doppler shift derived previously. Assume that the light source is at an angle θ as in Figure 13. In frame **S** the wave 4-vector is then

$$k^\mu = \left(\frac{\omega}{c}, -k\cos\theta, -k\sin\theta, 0\right)\,, \tag{4.30}$$

where $\omega = ck$ (notice that we have written the 4-vector with the index up.) Then in frame $\mathbf{S}'$ we have

$$k^{\mu'} = \Lambda^{\mu'}{}_{\nu}\, k^\nu\,, \tag{4.31}$$

and so we have

$$\begin{aligned} c\,k^{\mu'} &= (\gamma(\omega - vk^x), \gamma(ck^x - v\,\omega/c), ck^y, 0) \\ &= (\gamma(1+(v/c)\cos\theta)\omega, -\gamma(\cos\theta + v/c)\omega, -\sin\theta\,\omega, 0)\,, \end{aligned} \tag{4.32}$$

Therefore,

$$\nu' = \frac{\omega'}{2\pi} = \gamma(1+(v/c)\cos\theta)\nu\,. \tag{4.33}$$

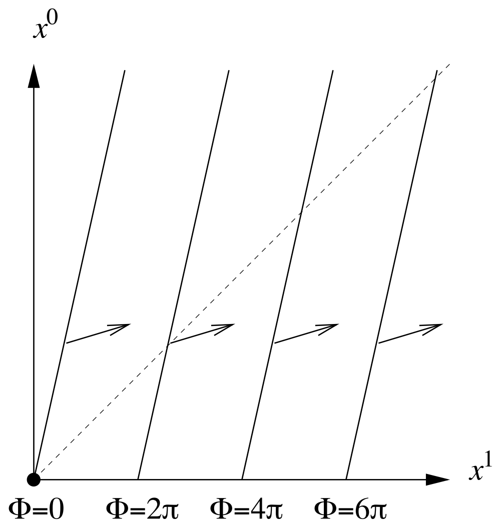


Figure 17: A wave-train moving at constant velocity in the x direction. The parallel lines are the world-lines for the wave-fronts. The wave-vector k_μ are the normals to the wave-fronts.

Here, ν' is the frequency of the source and ν is the frequency observed by the observer, so this result matches (2.26).

We can also use (4.32) to find the aberration. The cosine of the angle is just the negative of the first spatial component divided by the time component, while the sine is the negative of the second spatial component divided by the time component. Hence, in $\mathbf{S}'$ this leads to

$$\begin{aligned} \cos\theta' = -\frac{ck^{x'}}{\omega'} &= \frac{\cos\theta + v/c}{1+(v/c)\cos\theta} \\ \sin\theta' = -\frac{ck^{y'}}{\omega'} &= \frac{\sin\theta}{\gamma(1+(v/c)\cos\theta)}\,, \end{aligned} \tag{4.34}$$

which is the inverse transformation of (2.39).

4.6 4-momentum

Recall that for nonrelativistic physics, the momentum of a particle is related to its velocity by

$$\vec{p} = m\vec{v} \tag{4.35}$$

where m is the particle's mass. So it is clear that we can make a *4-momentum*, which we express as p^μ, by taking the 4-velocity and multiplying by a scalar quantity which has the units of mass,

$$p^\mu = m_0 u^\mu\,. \tag{4.36}$$

m_0 is called the *rest mass* of the particle, the mass of the particle when it is at rest. Let us suppose that the particle itself has velocity $\vec{v}$. In this case we have

$$p^\mu = (m_0\gamma(v)c, m_0\gamma(v)\vec{v})\,. \tag{4.37}$$

If we look at the spatial components of p^μ for very small velocities, $v << c$, then we see that $p^i \approx m_0 v^i$. The spatial components of the 4-momentum reduce to the ordinary Newtonian momentum in the nonrelativistic limit.

Now how should one interpret p^0? Again let us consider the nonrelativistic limit and let us consider the Taylor expansion of $\gamma(v)$,

$$\gamma(v) = \frac{1}{\sqrt{1 - v^2/c^2}} \approx 1 + \frac{1}{2}\frac{v^2}{c^2}\,. \tag{4.38}$$

If we then consider cp^0 in this limit, then we find

$$cp^0 \approx m_0 c^2 + \frac{1}{2} m_0 v^2\,. \tag{4.39}$$

The second term in (4.39) is simply the ordinary Newtonian kinetic energy of a particle with mass m_0. So $p^0\,c$ is interpreted as an energy. This means that when the particle is at rest, it has energy

$$E = m_0 c^2\,, \tag{4.40}$$

which is known as its rest energy.

Sometimes one uses the *relativistic mass*, m, where

$$m = \gamma(v) m_0\,. \tag{4.41}$$

In this case, the spatial components of the 4-vector have the Newtonian form $p^i = mv^i$. Note that while the velocity of a particle is limited by the speed of light, its momentum is not because the relativistic mass approaches ∞ as v approaches c.

Finally, from the 4-momentum we can construct a very useful invariant which will be used all the time when we consider particle mechanics. From its definition in (4.36) we see that

$$p_\mu p^\mu = m_0^2 u_\mu u^\mu = m_0^2 c^2\,. \tag{4.42}$$

5 Particle Mechanics

We learned in our classes on Newtonian mechanics that the rate of change of momentum is equal to the force acting on the system, that is

$$\vec{f} = \frac{d\vec{p}}{dt}\,. \tag{5.1}$$

Hence, if there is no external force then the total momentum is conserved. We would now like to find a corresponding statement in special relativity.

Let us suppose that the spatial momentum is given as in (4.35) and that we have a situation where two particles collide with each other. Their individual momenta will not be conserved, but their total momentum *will* be conserved. For the sake of argument, let us suppose that the two particles have the same mass, m_0 and that they undergo an elastic collision. If the incoming velocities in frame **S** are $\vec{u}_1$ and $\vec{u}_2$ and the outgoing velocites are $\vec{u}_3$ and $\vec{u}_4$, then conservation of momentum leads to

$$m_0\vec{u}_1 + m_0\vec{u}_2 = m_0\vec{u}_3 + m_0\vec{u}_4 \qquad \Rightarrow \qquad \vec{u}_1 + \vec{u}_2 = \vec{u}_3 + \vec{u}_4\,. \tag{5.2}$$

However, given these velocities, an observer in $\mathbf{S}'$ would measure

$$\begin{aligned} \vec{u}\,'_{1,||} &= \frac{\vec{u}_{1,||} - \vec{v}}{1 - \vec{u}_1\cdot\vec{v}/c^2} & \qquad \vec{u}\,'_{,2,||} &= \frac{\vec{u}_{2,||} - \vec{v}}{1 - \vec{u}_2\cdot\vec{v}/c^2} \\ \vec{u}\,'_{3,||} &= \frac{\vec{u}_{3,||} - \vec{v}}{1 - \vec{u}_3\cdot\vec{v}/c^2} & \qquad \vec{u}\,'_{4,||} &= \frac{\vec{u}_{4,||} - \vec{v}}{1 - \vec{u}_4\cdot\vec{v}/c^2}\,, \end{aligned} \tag{5.3}$$

where $||$ refers to the parallel part. Given that the denominator factors are in general different, this observer would find

$$\vec{u}\,'_1 + \vec{u}\,'_2 \neq \vec{u}\,'_3 + \vec{u}\,'_4\,, \tag{5.4}$$

and so momentum does not seem to be conserved.

Of course, the problem is that we are not using the correct momentum! Instead we should consider the 4-momentum introduced in the last section. To see why this will work, we could compute the actual velocities and show that they are consistent with momentum conservation. But there is a much simpler way to proceed. We note that if this momentum is conserved we would find $\Delta P^\mu = 0$, where ΔP^μ is the difference between the total outgoing and the total incoming momenta. ΔP^μ is a contravariant vector, which is a particular type of tensor. And as we already argued in the previous section, if all components of a tensor are zero in one frame, then all components are zero in *any other* frame. Hence, any observer would find that ΔP^μ is zero, so the total momentum is conserved for every observer.

The equation $\Delta P^\mu = 0$ is actually four equations, one for each value of μ. For the spatial components, this is conservation of spatial momentum. But what does $\Delta P^0 = 0$ mean? In the last section we saw that p^0 for a nonrelativistic particle could be expanded as

$$c\,p^0 \approx m_0c^2 + \tfrac{1}{2}\,m_0\,v^2\,, \tag{5.5}$$

where the second term is the Newtonian kinetic energy for a particle of mass m_0. Hence, it seems reasonable to identify $c\,p^0$ with the particle's energy and so the equation $\Delta P^0 = 0$ is the statement that energy is conserved. If a particles velocity is 0 then its energy is $m_0 c^2$. Hence, we call this its rest energy, which we can also think of as the internal energy of the particle. We then define its kinetic energy T to be

$$T = c\,p^0 - m_0\,c^2\,, \tag{5.6}$$

that is, its total energy minus the rest energy.

Now the rest energy is a *real* energy and not just a constant that goes along for the ride. To see this, suppose that we have an elastic collision. This means that the particles going in are the same as the particles going out. Let's say that particle 3 has the same rest mass as particle 1 and particle 4 has the same rest mass as particle 2. Then the conservation of energy law $\Delta P^0 = 0$ would lead to

$$T_1 + T_2 = T_3 + T_4\,. \tag{5.7}$$

This is conservation of kinetic energy, which we know is true for elastic collisions in Newtonian mechanics. But we could also consider the case where the outgoing particles are different from the incoming particles. In this case we no longer would have the conservation rule in (5.7). Hence, this is not an elastic collision. But the total energy is still conserved, its just that some of the kinetic energy is converted to rest energy (that is internal energy) or vice versa. So mass can become kinetic energy! This is the true meaning of Einstein's famous equation $E = m\,c^2$.

So the particle's mass contributes to the total energy and can be transmuted into other forms of energy, say kinetic energy. But just like mass is energy, we can also have that energy is mass. Suppose we have two particles, say a proton and a neutron. They can bind together to form a *deuteron*. If we measure the rest mass of the deuteron, M_d, we find

$$M_d < m_p + m_n\,, \tag{5.8}$$

where m_p and m_n are the rest masses of the proton and neutron respectively. We see that the rest energy of the deuteron is less than the sum of the rest energy of the proton and the rest energy of the neutron. This difference is due to the binding energy of the proton to the neutron. In order to separate the two out to spatial infinity one needs an amount of energy equal to the binding energy $E_{\rm bind}$ which is given by

$$E_{\rm bind} = (m_p + m_n - M_d)\,c^2\,, \tag{5.9}$$

Hence the deuteron mass has less internal energy than the internal energy of the proton and neutron because of the binding energy.

Since mass is equivalent to energy, particle physicists often state the rest masses of particles as energies. Some examples are:

$$\begin{aligned} \text{proton}: \ m_p c^2 &= 938 \text{ MeV} \\ \text{neutron}: \ m_n c^2 &= 939 \text{ MeV} \\ \text{electron}: \ m_e c^2 &= 0.511 \text{ MeV} \\ Z \text{ boson}: \ m_Z c^2 &= 92 \text{ GeV} \end{aligned}$$

where "MeV" stands for Mega electron volts (10^6 eV) and "GeV" stands for Giga electron volts (10^9 eV). (Recall that 1 eV $= 1.6 \times 10^{-19}$ joules).

You might have heard that the LHC, a large accelerator at CERN in Geneva Switzerland, is now operational after a disastrous startup. The accelerator has two beams of protons going in opposite directions around the loop. At the present time, each proton in the accelerator's rest frame has 3.5 TeV $= 3.5 \times 10^3$ GeV. The original plan was for each proton to have 7 TeV and this will eventually happen in 2016. Assuming this higher energy and using that $c\, p^0 = \gamma\, m_0\, c^2$, we have that the protons will each have $\gamma \approx 7.5 \times 10^3$, which translates into a speed of

$$v = c\sqrt{1 - 1/\gamma^2} \approx (1 - 0.9 \times 10^{-8})c\,. \tag{5.10}$$

In other words the protons will be traveling at a rate of 3 m/s less than the speed of light!

5.1 Particle Decay

In this section we consider the decay of one particle of mass m_1 to two particles of mass m_2 and m_3. Let us assume that we are in the rest frame of the first particle and we want to find the outgoing velocities of the decay products. In this frame, the momentum 4-vector for the first particle is

$$p_1^\mu = (m_1 c, 0, 0, 0)\,, \tag{5.11}$$

and conservation of momentum tells us that

$$p_1^\mu = p_2^\mu + p_3^\mu\,. \tag{5.12}$$

Hence, the momentum 4-vectors for particles 2 and 3 are

$$p_2^\mu = (m_2\, c\, \gamma_2, m_2\, \vec{v}_2\, \gamma_2) \qquad p_3^\mu = (m_3\, c\, \gamma_3, m_3\, \vec{v}_3\, \gamma_3)\,, \tag{5.13}$$

where conservation of momentum tells us that $\vec{v}_2$ is in the opposite direction of $\vec{v}_3$. Now we can do a simple trick. We rewrite (5.12) as

$$p_2^\mu = p_1^\mu - p_3^\mu\,, \tag{5.14}$$

and then "take the square" of both sides, to give

$$\begin{aligned} p_2^\mu p_{2\mu} &= (p_1^\mu - p_3^\mu)(p_{1\mu} - p_{3\mu}) \\ p_2^\mu p_{2\mu} &= p_1^\mu p_{1\mu} + p_3^\mu p_{3\mu} - 2\, p_1^\mu p_{3\mu}\,. \end{aligned} \tag{5.15}$$

Using (4.42) and the explicit values in the above momenta, we find

$$m_2^2 = m_1^2 + m_3^2 - 2\, m_1 m_3 \gamma_3\,, \tag{5.16}$$

and so we reach the result for γ_3

$$\gamma_3 = \frac{m_1^2 + m_3^2 - m_2^2}{2 m_1 m_3}\,. \tag{5.17}$$

Using the same logic, we also have

$$\gamma_2 = \frac{m_1^2 + m_2^2 - m_3^2}{2m_1 m_2}\,. \tag{5.18}$$

Hence, the outgoing speeds of the particles are

$$\begin{aligned} |\vec{v}_2| = c\sqrt{1 - 1/\gamma_2^2} &= c\sqrt{\frac{(m_1^2 + m_2^2 - m_3^2)^2 - 4m_1^2 m_2^2}{(m_1^2 + m_2^2 - m_3^2)^2}} \\ &= c\,\frac{\sqrt{(m_1^2 - m_2^2)^2 - 2(m_1^2 + m_2^2)m_3^2 + m_3^4}}{m_1^2 + m_2^2 - m_3^2}\,. \end{aligned} \tag{5.19}$$

Likewise,

$$|\vec{v}_3| = c\,\frac{\sqrt{(m_1^2 - m_3^2)^2 - 2(m_1^2 + m_3^2)m_2^2 + m_2^4}}{m_1^2 + m_3^2 - m_2^2}\,. \tag{5.20}$$

Finally, notice that if $m_2 > m_1 - m_3$ then

$$\gamma_3 < \frac{m_1^2 + m_3^2 - (m_1 - m_3)^2}{2m_1 m_3} = \frac{2m_1 m_3}{2m_1 m_3} = 1\,, \tag{5.21}$$

which is impossible since $\gamma_3 \geq 1$. Hence we conclude that $m_1 \geq m_2 + m_3$ for the decay to occur. This is sensible, since it means that the total energy of the original particle has to be greater than or equal to the internal energy of the decay products because the kinetic energy of the decay products cannot be negative.

5.2 Two body elastic scattering

The next process we consider is the elastic collision of two particles into the same two particles. To make the calculations easier, we assume that the particles have the same rest mass. Normally when one considers two body scattering, there are two distinct reference frames that are considered. One is known as the *lab frame*, where the incoming second particle is at rest. The other frame is the *center of momentum frame* (COM). This is often mistakenly called the "center of mass frame". In the COM frame, the total spatial momentum for the incoming particles is zero. Since the momentum is conserved, the total spatial momentum for the outgoing particles is also zero.

Without any loss of generality, let us assume that the incoming particle motion is all in the x direction and that the outgoing particle motion is in the $x - y$ plane. Let us assume that in the lab frame, **S**, the incoming velocity of the first particle is $\vec{u}_1 = u\hat{x}$ and that the outgoing particles are coming out at angles θ and ϕ. This process is shown in figure 18 a. In the COM frame, **S′**, the total spatial momentum is zero, meaning that the trajectories of the outgoing particles are in opposite directions (see figure 18 b). Assuming that each particle has velocity v in the COM frame, the 4-momentum are

$$\begin{aligned} p_1^{\mu'} &= (mc\,\gamma(v), vm\,\gamma, 0, 0), \qquad p_2^{\mu} = (mc\,\gamma(v), -vm\,\gamma(v), 0, 0), \\ p_3^{\mu'} &= (mc\,\gamma(v), vm\,\gamma(v)\cos\theta', vm\,\gamma(v)\sin\theta', 0) \\ p_4^{\mu'} &= (mc\,\gamma(v), -vm\,\gamma(v)\cos\theta', -vm\,\gamma(v)\sin\theta', 0)\,. \end{aligned} \tag{5.22}$$

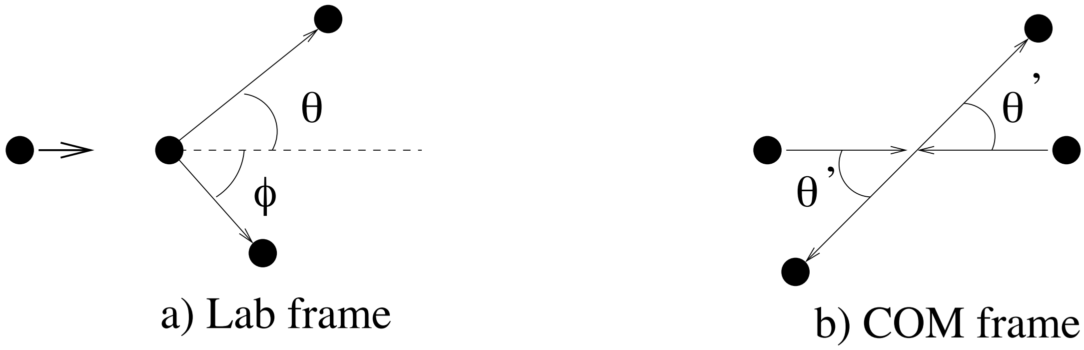


Figure 18: Two body elastic scattering in the lab frame (a) and in the COM frame (b).

Note that the velocity of $\mathbf{S}'$ is $\vec{v} = v\hat{x}$ *wrt* $\mathbf{S}$. Let us find u in terms of v. To do this we consider the invariant $(p_1+p_2)^2 \equiv (p_1^\mu + p_2^\mu)(p_{1\mu} + p_{2\mu})$. In the COM frame this is

$$(p_1+p_2)^2 = 2\,m^2c^2 + 2\,p_1^{\mu'}p_{2\mu'} = 2\,m^2c^2 + 2\,(\gamma(v))^2m^2(c^2+v^2)\,. \tag{5.23}$$

In the lab frame the same invariant is

$$(p_1+p_2)^2 = 2\,m^2c^2 + 2\,p_1^{\mu}p_{2\mu} = 2\,m^2c^2 + 2m^2c^2\gamma(u)\,. \tag{5.24}$$

Hence, we have that

$$\gamma(u) = (\gamma(v))^2\,\frac{c^2+v^2}{c^2} = \frac{c^2+v^2}{c^2-v^2}\,, \tag{5.25}$$

and so

$$u = c\sqrt{1-1/\gamma(u)^2} = \frac{2vc^2}{c^2+v^2}\,. \tag{5.26}$$

To find the angles θ and ϕ in terms of θ', we can use addition of velocities. The velocities in the COM frame $\mathbf{S}'$ are

$$u'_{3x} = v\cos\theta' \qquad u'_{3y} = v\sin\theta' \qquad u'_{4x} = -v\cos\theta' \qquad u'_{4y} = -v\sin\theta'\,. \tag{5.27}$$

Hence, using (2.5) we get

$$\begin{aligned} u_{3x} &= \frac{v(1+\cos\theta')}{1+(v^2/c^2)\cos\theta'} \qquad u_{3y} = \frac{v\sin\theta'}{\gamma(v)(1+(v^2/c^2)\cos\theta')} \\ u_{4x} &= \frac{v(1-\cos\theta')}{1-(v^2/c^2)\cos\theta'} \qquad u_{4y} = \frac{-v\sin\theta'}{\gamma(v)(1-(v^2/c^2)\cos\theta')}\,. \end{aligned} \tag{5.28}$$

The tangents of the angles are then

$$\begin{aligned} \tan\theta &= \frac{u_{3y}}{u_{3x}} = \frac{\sin\theta'}{\gamma(v)(1+\cos\theta')} = \frac{1}{\gamma(v)}\tan\tfrac{1}{2}\theta' \\ \tan\phi &= \frac{u_{4y}}{u_{4x}} = \frac{-\sin\theta'}{\gamma(v)(1-\cos\theta')} = \frac{1}{\gamma(v)}\tan\tfrac{1}{2}(\theta'+\pi)\,. \end{aligned} \tag{5.29}$$

5.3 Threshold energies

Suppose that two particles collide, one initially at rest, and a third particle is created. What is the minimum kinetic energy needed? The conservation of momentum equation is

$$p_1^\mu + p_2^\mu = p_3^\mu + p_4^\mu + p_5^\mu \,. \tag{5.30}$$

Let us assume that the two incoming particles and the first two outgoing particles have the same mass m and that the mass of the third outgoing particle is M. The best thing to do is to go into the COM frame. In the COM frame we have that the *threshold* occurs when the final particles all have zero spatial momentum. The incoming particles also have equal and opposite spatial momentum. Hence, the conservation of energy equation is

$$m\gamma(v)c^2 + m\gamma(v)c^2 = mc^2 + mc^2 + Mc^2 \,. \tag{5.31}$$

Hence, we have

$$\gamma(v) = 1 + \frac{M}{2m} \tag{5.32}$$

The kinetic energy for each incoming particle T is then

$$T = \gamma(v)mc^2 - mc^2 = \frac{M}{2}c^2 \,, \tag{5.33}$$

in other words the extra energy for each particle is used to create half the rest energy of the third particle.

Let us now find how much energy is required in the lab frame. In the lab frame we let the velocity of the first particle be $\vec{u}$ and the second particle is at rest. Hence, we have that

$$p_1^\mu = (mc\gamma(u), m\vec{u}\gamma(u)) \qquad p_2^\mu = (mc, \vec{0}) \,. \tag{5.34}$$

We then construct the invariant

$$p_1 \cdot p_2 \equiv p_1^\mu p_{2\mu} = m^2c^2\gamma(u) = m(T_{\text{lab}} + mc^2) \,, \tag{5.35}$$

where T_{lab} is the kinetic energy of the first particle in the lab frame. Back in the COM frame, this invariant is

$$p_1 \cdot p_2 = m^2c^2\gamma(v)^2 + m^2v^2\gamma(v)^2 = 2m^2c^2\gamma(v)^2 - m^2c^2 = m^2c^2\left(1 + \frac{2M}{m} + \frac{M^2}{2m^2}\right) \tag{5.36}$$

where we used the result in (5.33). Comparing the invariants in the two frames, we find

$$T_{\text{lab}} = Mc^2\left(2 + \frac{M}{2m}\right) \,, \tag{5.37}$$

In particular, notice that if $M >> m$, then

$$T_{\text{lab}} \approx \frac{M^2c^2}{2m}\,, \tag{5.38}$$

which is much bigger than the rest energy of the third particle. This means that it will take a tremendous amount of kinetic energy to produce the third particle. The reason for the big difference is that in order to conserve the spatial momentum, the new particle will also have to a lot of kinetic energy, so most of the energy of the incoming particle goes toward the creation of the third particle's kinetic energy.

This is why it is much more efficient to build particle colliders where two beams of particles smash into each other with equal and opposite momenta, than it is to build an accelerator where one particle hits another particle at rest. For example, the aforementioned LHC collider at CERN is built so that the accelerator is in the COM frame of the two colliding beams of protons. Each beam will eventually have an energy of 7 TeV (a TeV is a trillion eV=1000 GeV) so there is 14 TeV of energy available to create particles. The beam should be powerful enough to create particles with masses perhaps in excess of 1 TeV (The reason why it won't make a single particle with mass 14 TeV is because the protons themselves are not point particles but are actually a collection of quarks and gluons). To make a TeV particle when one proton is at rest would require that the other proton have a kinetic energy that is in excess of 10^5 TeV.

5.4 The 4-momentum for photons

Light comes in discrete quanta called photons. Each photon is an individual particle. It might seem that constructing a 4-momentum for a photon as we did for other particles is problematic since photons do not have a rest frame to measure a rest energy in. Nonetheless, previously we saw that the 4-momentum is proportional to the 4-velocity u^μ, so we can try the same thing and assume that the photon's 4-momentum is proportional to its 4-velocity.

We also saw that the definition for a light ray's 4-velocity depends on how we parameterize the light's trajectory. However, one fact is indisputable, namely that the 4-velocity is light-like, that is, $u^\mu u_\mu = 0$. Therefore, what we will do is let the photon's 4-momentum be p^μ with $p^\mu p_\mu = 0$. We saw previously for a particle with a rest mass m_0 that $p^\mu p_\mu = m_0 c^2$. Hence, one can say that a photon has zero rest mass, or that is a massless particle.

Actually quantum mechanics lets us make a more definite statement. The DeBroglie wavelength is given by

$$\lambda = \frac{h}{|\vec{p}|}\,, \tag{5.39}$$

where $\vec{p}$ is the spatial part of the 4-momentum and h is Planck's constant. A better way to write this relation is with the wave-vector $\vec{k}$, in which case we have

$$\vec{p} = \hbar\,\vec{k}\,, \tag{5.40}$$

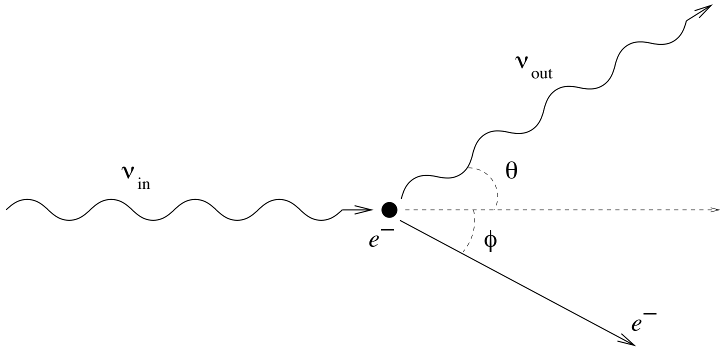


Figure 19: A photon with frequency $\nu_{\rm in}$ scattering off an electron at rest. The outgoing photon has frequency ν_{out} and an angle θ off of the x-axis.

where $\hbar = h/2\pi$ and $h = 6.63 \times 10^{-34}$ joule-sec is Planck's constant. We can now extend this to the complete 4-vector giving us

$$p^\mu = \hbar\, k^\mu\,. \tag{5.41}$$

Recall that we previously found for light waves that $k^\mu k_\mu = 0$, so obviously p^μ will satisfy the same relation.

In any case, we can express the 4-momentum for a photon as

$$p^\mu = (|\vec{p}\,|, \vec{p}) = \hbar(\omega/c, \vec{k})\,. \tag{5.42}$$

From this we see that the photon's energy is $\hbar\omega = h\nu$.

5.5 Compton scattering

Suppose we consider a photon with frequency $\nu_{\rm in}$ scattering off of a charged particle at rest, such as an electron. This is usually an elastic process, although there is a small but nonzero probability that more photons will be created, making the process inelastic. In any case, we assume that the process is elastic. Without any loss of generality, we assume that the photon is heading in the x direction and that the photon scatters off with angle θ in the $x-y$ plane. Momentum conservation then implies that the electron also scatters off in the $x-y$ plane. The outgoing photon will have frequency $\nu_{\rm out}$. This process is shown in figure 19.

We now show how to find the outgoing frequency in terms of the incoming frequency and the outgoing angle θ. Conservation of 4-momentum gives us

$$p_1^\mu + p_2^\mu = p_3^\mu + p_4^\mu\,, \tag{5.43}$$

where p_1^μ and p_3^μ are the momenta for the photons and p_2^μ and p_4^μ are the momenta for the electrons. The first three momenta are

$$\begin{aligned}
p_1^\mu &= \left(\frac{h\nu_{\rm in}}{c}, \frac{h\nu_{\rm in}}{c}, 0, 0\right)\\
p_2^\mu &= (m_e c, 0, 0, 0)\\
p_3^\mu &= \left(\frac{h\nu_{\rm out}}{c}, \frac{h\nu_{\rm out}}{c}\cos\theta, \frac{h\nu_{\rm out}}{c}\sin\theta, 0\right)\,,
\end{aligned} \tag{5.44}$$

where m_e is the electron mass. We then bring p_3^μ on the other side of the conservation equation and square both side,

$$(p_1 + p_2 - p_3)^2 = (p_4)^2 \,, \tag{5.45}$$

where $(p)^2 \equiv p^\mu p_\mu$. This then gives

$$(p_1)^2 + (p_2)^2 + (p_3)^2 + 2p_2 \cdot (p_1 - p_3) - 2p_1 \cdot p_3 = (p_4)^2 \,, \tag{5.46}$$

where $p \cdot q \equiv p^\mu q_\mu$. Using that $(p_2)^2 = (p_4)^2$ and $(p_1)^2 = (p_3)^2 = 0$ we have

$$2m_e h(\nu_{\rm in} - \nu_{out}) = 2\frac{\nu_{\rm in}\nu_{\rm out}h^2}{c^2}(1 - \cos\theta) \,. \tag{5.47}$$

Dividing both sides of the equation by $m_e\nu_{\rm in}\nu_{\rm out}h$, multiplying by c and using the trigonometric identity $(1 - \cos\theta) = 2\sin^2\frac{1}{2}\theta$, we arrive at the expression

$$\frac{c}{\nu_{\rm out}} - \frac{c}{\nu_{\rm in}} = 2\sin^2\tfrac{1}{2}\theta\,\frac{h}{m_e c} \tag{5.48}$$

This is then rewritten as

$$\lambda_{\rm out} - \lambda_{\rm in} = 2\sin^2\tfrac{1}{2}\theta\,\lambda_C \geq 0 \,. \tag{5.49}$$

The quantity $\lambda_C = h/m_e c$ is called the *electron's Compton wavelength*. We can also see that the outgoing photon has a longer wavelength than the incoming photon. It is easy to understand this: the photon has to give up some of its energy to impart kinetic energy to the electron. The longer wavelength has lower energy. One should also observe that for scattering in the forward direction where the deflection angle θ is small, the shift in wavelength is also small because of the $\sin^2\frac{1}{2}\theta$ factor. In fact the shift is largest when $\theta = \pi$, which corresponds to a photon being scattered straight back.

Plugging in the values for h, c and m_e, we find

$$\begin{aligned}\lambda_C = \frac{hc}{m_e c^2} &= \frac{(6.6\times10^{-34}\ \text{joule-sec})(3\times10^8\ \text{m/sec})}{(0.511\times10^6\ \text{eV})(1.6\times10^{-19}\ \text{joules/eV})} \\ &= 2.4\times10^{-12}\ \text{m} = .024\ \text{Å}\,,\end{aligned} \tag{5.50}$$

so we see that the Compton wavelength of an electron is quite a bit shorter than the typical atomic scale. In fact for ordinary visible light where $\lambda \sim 5000$ Å, Compton scattering will have a miniscule effect on the wavelength. Photons in these wavelengths have energies on the order of an eV. However, for photons with energies of an MeV the shift in wavelength can be significant. These photons are known as gamma rays and are typical products in nuclear decays.

5.6 Forces

In this section we generalize the Newtonian force relation in (5.1) to 4-vectors. As we have done before when we encountered a time derivative of a 3-vector, we replace the

3-vector with a 4-vector and the time derivative with a derivative with respect to the proper time τ. Hence, we define the 4-force as

$$F^\mu \equiv \frac{dp^\mu}{d\tau} \tag{5.51}$$

Using time dilation $dt = \gamma(v)d\tau$, we find that the components of F^μ are

$$F^\mu = \left(\gamma(v)\frac{dp^0}{dt}, \gamma(v)\frac{d\vec{p}}{dt}\right) = \gamma(v)\left(\frac{1}{c}\frac{dE}{dt}, \vec{f}\right) . \tag{5.52}$$

Hence, the first component is related to the rate of change of the energy, while the spatial parts is the usual force $\vec{f}$ multiplied by a γ factor.

Now notice further that

$$F^\mu = \frac{d}{d\tau}(m_0 u^\mu) = m_0 a^\mu + u^\mu \frac{dm_0}{d\tau} . \tag{5.53}$$

If we now take the scalar product of F^μ with u^μ we get

$$u \cdot F = u^\mu F_\mu = m_0 u \cdot a + u \cdot u\,\frac{dm_0}{d\tau} = 0 + \frac{d(m_0 c^2)}{d\tau} . \tag{5.54}$$

If we now take the scalar product using the explicit components in (4.15) and (5.52), we get

$$u \cdot F = \gamma(v)^2 \left(\frac{dE}{dt} - \vec{v} \cdot \vec{f}\right) . \tag{5.55}$$

Thus, comparing eqs. (5.54) and (5.55) we derive the relation

$$\frac{dE}{dt} = \vec{v} \cdot \vec{f} + \frac{1}{\gamma(v)}\frac{d(m_0 c^2)}{dt} . \tag{5.56}$$

In other words, the rate of change of the energy is given by the usual $\vec{v} \cdot \vec{f}$ term we have in Newtonian mechanics, plus a term corresponding to the rate of change of the rest energy of the body.

A *pure force* is one that does not change the rest energy. We can see from (5.54) that a pure force has $u \cdot F = 0$. A *heatlike force* is one which does not change the particle's velocity. In the case of the pure force, $\frac{dE}{dt}$ can be expressed as

$$\frac{dE}{dt} = \frac{d\vec{x}}{dt} \cdot \vec{f} = \frac{d}{dt}(\vec{x} \cdot \vec{f}) = \frac{dW}{dt} , \tag{5.57}$$

where W is the work done on the system and we have used that $\vec{f}$ has no explicit t dependence.

In Newtonian mechanics, a conservative force (that is, one with no frictional forces) can be written as the derivative of a potential Φ as follows:

$$\vec{f} = -\vec{\nabla}\Phi . \tag{5.58}$$

Let us now show that there is no equivalent for a pure 4-force, in other words, it is not possible to write $F_\mu = \partial_\mu \Phi$ if the force is nonzero. To show this, assume that F_μ *does* take this form. Then we have

$$0 = u^\mu F_\mu = u^\mu \partial_u \Phi = \frac{dx^\mu}{d\tau}\frac{\partial \Phi}{\partial x^\mu} = \frac{d\Phi}{d\tau}\,. \tag{5.59}$$

This only has a solution if Φ is a constant, and so the only solution has $F_\mu = 0$.

Newtonian mechanics is of course not complete without Newton's second law,

$$\vec{f} = m\,\vec{a}\,. \tag{5.60}$$

Hence the acceleration is always parallel to the force. The same thing does not hold true for the 4-force and 4-acceleration. Let us again assume that we have a pure force. Using (5.1) we have

$$\frac{d}{dt}(m_0\gamma\vec{v}) = m_0\gamma\,\vec{a} + \vec{v}\frac{d}{dt}(m_0\gamma) = m_0\gamma\,\vec{a} + \frac{\vec{v}}{c^2}\frac{dE}{dt}\,. \tag{5.61}$$

For a pure force we also have

$$\frac{dE}{dt} = \vec{v}\cdot\vec{f}, \tag{5.62}$$

so combining this with the previous equation we get

$$\vec{f} = m_0\gamma\,\vec{a} + \frac{\vec{v}}{c^2}(\vec{v}\cdot\vec{f})\,. \tag{5.63}$$

Consequently, $\vec{f}$ is parallel to $\vec{a}$ only if, *either* $\vec{f}$ is parallel to $\vec{v}$, *or* $\vec{v}\cdot\vec{f} = 0$. Let us write $\vec{f}$ as

$$\vec{f} = \vec{f}_{||} + \vec{f}_\perp\,, \tag{5.64}$$

where $\vec{f}_{||}$ is the part of $\vec{f}$ that is parallel to $\vec{v}$ and $\vec{f}_\perp$ is the part that is perpendicular to $\vec{v}$. For the parallel part we have

$$\vec{f}_{||} = \vec{f}_{||} = m_0\gamma\,\vec{a}_{||} + \frac{v^2}{c^2}\vec{f}_{||}\,, \tag{5.65}$$

where $\vec{a}_{||}$ is the component of $\vec{a}$ that is parallel to $\vec{v}$. From (5.65) it follows that

$$\vec{f}_{||} = \frac{m_0\gamma}{1 - v^2/c^2}\vec{a}_{||} = m_0\gamma^3\,\vec{a}_{||}\,. \tag{5.66}$$

For the perpendicular part, we clearly have

$$\vec{f}_\perp = m_0\gamma\,\vec{a}_\perp \tag{5.67}$$

Since the proportionality constant is different between the parallel and perpendicular parts of $\vec{f}$ and $\vec{v}$, $\vec{f}$ and $\vec{a}$ are in general not parallel to each other. Of course in the nonrelativistic limit $v^2/c^2 << 1$, then $\gamma \approx 1$ and the proportionality factors become the same, in which case $\vec{f}$ and $\vec{a}$ are parallel.

6 Electromagnetism (Introduction)

After making a few minimal assumptions and invoking Occam's razor we can derive the laws of electrodynamics, and in particular Maxwell's equations, more or less from first principles. We will start by trying to find a force law for charged particles. We make three assumptions:

1. The electromagnetic force $\vec{f}$ is covariant. In other words it is part of a 4-force F^μ where $F^\mu = \gamma\,(\frac{dE/c}{dt}, \vec{f})$.
2. The force is proportional to the particle's charge q.
3. The electrodynamic force should not change the rest energy of the particles, afterall, an electron in an electromagnetic field stays an electron with the same rest energy. Thus, the force is a pure force and so $u^\mu F_\mu = 0$.

The first thing we observe is that the force must be velocity dependent, otherwise we could choose an arbitrary timelike u^μ such that $u^\mu F_\mu \neq 0$. We now make one further assumption: we want the velocity dependence to be as simple as possible (this is the Occam's razor part), so we choose it to be linear in u^μ. Hence, a covariant 4-force satisfying all of these properties can be written as

$$F_\mu = \frac{q}{c}\, E_{\mu\nu}\, u^\nu \tag{6.1}$$

where the c is put in for later convenience and $E_{\mu\nu}$ has no q dependence. $E_{\mu\nu}$ has to have two lower indices in order to map the contravariant vector u^μ to the covariant vector F_μ. The pure force condition gives

$$E_{\mu\nu}\, u^\mu\, u^\nu = 0 \tag{6.2}$$

for all u^μ. But this is only possible if $E_{\mu\nu}$ is antisymmetric. It is straightforward to show that an antisymmetric tensor does give zero:

$$E_{\mu\nu}\, u^\mu\, u^\nu = -E_{\nu\mu}\, u^\mu\, u^\nu = -E_{\nu\mu}\, u^\nu\, u^\mu = -E_{\mu\nu}\, u^\mu\, u^\nu\,, \tag{6.3}$$

where in the last step we simply relabeled the dummy indices. Hence, $E_{\mu\nu}\, u^\mu\, u^\nu$ is equal to minus itself, which is possible only if it is zero. We have previously seen at the end of section 3 that an antisymmetric tensor is antisymmetric in all frames, so (6.1) is indeed a covariant equation. We will call $E_{\mu\nu}$ the electromagnetic field[7].

6.1 The field equations

We next assume that the fields couple to currents. The electric current is written as $\vec{j}$ and is equal to the charge flow per unit time and unit area along the direction of its flow.

[7]Most textbooks use the symbol $F_{\mu\nu}$ for the field, but we are following Rindler's notation. Besides, F_μ is reserved for the force, so to minimize confusion we will stick to $E_{\mu\nu}$.

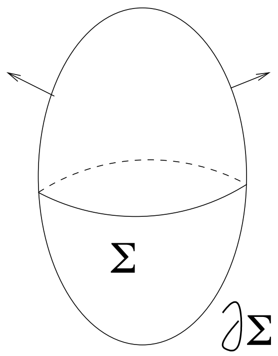


Figure 20: A volume Σ containing a charge Q where current can flow out of the boundary $\partial\Sigma$.

But we still want to construct a 4-vector, and we can do this by including the charge density ρ. To this end, we let

$$j^\mu = (\rho c, \vec{j}\,)\,. \tag{6.4}$$

To see that this makes sense, suppose we are in a frame **S** where the charges are all at rest and that the charge density is uniform, $\rho = \rho_0$. Since they are at rest, the 4-velocity for each charged particle is $u^\mu = (c, 0, 0, 0)$. Hence, we get that $j^\mu = \rho_0 u^\mu$. If we boost to a new frame $\mathbf{S}'$, j^μ becomes

$$j^\mu = (\rho_0\gamma\, c, \rho_0\gamma\,\vec{v})\,. \tag{6.5}$$

Hence the charge density in this frame is $\rho = \rho_0\gamma$ and the current is $\vec{j} = \rho_0\gamma\,\vec{v}$. The factor of γ is easy to understand. In boosting to a new frame, lengths along the boost direction are contracted by a factor of γ, so the *charge density* goes up by a factor of γ. The current $\vec{j}$ is then this charge density multiplied by the velocity that the charges are moving.

The current also satisfies the conservation law

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\vec{j} = 0\,, \tag{6.6}$$

which we can rewrite as an invariant equation

$$\partial_\mu\, j^\mu = 0\,. \tag{6.7}$$

To see why this current conservation law is true, suppose we have a volume Σ containing a total charge Q and we also suppose that charge can flow in or out through the boundary of the volume $\partial\Sigma$, as shown in figure 20. If we now integrate the conservation law in (6.6) over the volume Σ, we get

$$\begin{aligned}\int_\Sigma d^3x\left[\frac{\partial}{\partial t}\rho + \vec{\nabla}\cdot\vec{j}\right] &= \frac{\partial}{\partial t}\int_\Sigma d^3x\,\rho + \int_\Sigma d^3x\,\vec{\nabla}\cdot\vec{j} \\ &= \frac{\partial Q}{\partial t} + \int_{\partial\Sigma} d\vec{S}\cdot\vec{j} = 0\,,\end{aligned} \tag{6.8}$$

where we have used Gauss' law to relate the volume integral of a divergence to a surface integral. The meaning of this equation is that the rate of change of the charge Q inside Σ is minus the rate that the charge is flowing out. If no charge is created or destroyed, then this will be true. So current conservation is basically a statement that the total charge is conserved.

We now look for a field equation that satisfies current conservation. The simplest such equation one can write down is

$$\partial_\mu E^{\mu\nu} = \kappa\, j^\nu\,, \tag{6.9}$$

where κ is a constant to be determined later[8]. To see that this is consistent with current conservation, let us take a ∂_ν derivative on both sides of (6.9). This gives

$$\partial_\mu \partial_\nu E^{\mu\nu} = \kappa\, \partial_\nu j^\nu\,. \tag{6.10}$$

The right hand side of (6.10) is zero. Using the commutativity of the derivatives (their order does not matter) and the antisymmetry of $E^{\mu\nu}$, the left hand side of (6.10) satisfies

$$\partial_\mu \partial_\nu E^{\mu\nu} = \partial_\nu \partial_\mu E^{\mu\nu} = -\partial_\nu \partial_\mu E^{\nu\mu} = -\partial_\mu \partial_\nu E^{\mu\nu}\,, \tag{6.11}$$

where the last step is just a relabeling of the dummy indices. Just like in (6.3), we get an expression that is minus itself, and so it is zero. Hence, (6.9) is consistent with current conservation.

6.2 Maxwell's equations: part 1

The field equation (6.9) is actually four equations since there is one free index in the equation. We now show that this is equivalent to half of Maxwell's equations.

We first note that because of its antisymmetry, the field $E_{\mu\nu}$ has six independent components. Let's count them: If we set $\mu = \nu$ then $E_{\mu\nu} = -E_{\mu\nu}$ and thus we have to choose $\mu \neq \nu$ in order to get a nonzero component. In this case we still have $E_{\mu\nu} = -E_{\nu\mu}$, so we can choose $\mu < \nu$ to list all the components, since the $\mu > \nu$ case is related to the $\mu < \nu$ case by the antisymmetry relation. Hence there are 3+2+1=6 six different choices.

We define the components as follows: Three of the components are

$$E_{0i} = e_i\,, \qquad i = 1, 2, 3\,. \tag{6.12}$$

The other three are

$$E_{ij} = -c\,\varepsilon_{ijk}\, b_k\,, \tag{6.13}$$

where ε_{ijk} is completely antisymmetric under the exchange of any two indices and has $\varepsilon_{123} = 1$. The repeated k index is assumed to be summed over, even though I have shown

[8]Notice that the field has raised indices in (6.9), which one gets by using two η tensors: $E^{\mu\nu} = \eta^{\mu\lambda}\eta^{\nu\rho}E_{\lambda\rho}$

this with both indices down. Writing $E_{\mu\nu}$ in matrix form, these definitions look like

$$E_{\mu\nu} = \begin{pmatrix} 0 & e_1 & e_2 & e_3 \\ -e_1 & 0 & -c\,b_3 & c\,b_2 \\ -e_2 & c\,b_3 & 0 & -c\,b_1 \\ -e_3 & -c\,b_2 & c\,b_1 & 0 \end{pmatrix}_{\mu\nu} . \tag{6.14}$$

If we raise the indices, this becomes[9]

$$E^{\mu\nu} = \eta^{\mu\lambda}\eta^{\nu\sigma}E_{\lambda\sigma} = \begin{pmatrix} 0 & -e_1 & -e_2 & -e_3 \\ e_1 & 0 & -c\,b_3 & c\,b_2 \\ e_2 & c\,b_3 & 0 & -c\,b_1 \\ e_3 & -c\,b_2 & c\,b_1 & 0 \end{pmatrix}^{\mu\nu} . \tag{6.15}$$

Let us now write the field equations using e_i and b_i. We start with the free index $\nu = 0$. Then the left hand side of (6.9) is

$$\partial_\mu E^{\mu 0} = \partial_i E^{i0} = \frac{\partial}{\partial x^i} e_i = \vec{\nabla} \cdot \vec{e}. \tag{6.16}$$

The right hand side is $\kappa\, j^0 = \kappa\, c\, \rho$, so this component of the equation becomes

$$\vec{\nabla} \cdot \vec{e} = \kappa\, c\, \rho\,. \tag{6.17}$$

Gauss' law for an electric field, $\vec{e}$, in SI units (Standard International) is

$$\vec{\nabla} \cdot \vec{e} = \frac{1}{\epsilon_0}\,\rho\,, \tag{6.18}$$

where ϵ_0 is the dielectric constant of the vacuum. Hence, (6.17) is Maxwell's first equation (Gauss' Law) if we identify[10] $\kappa = 1/(c\epsilon_0)$.

Next, let us consider $\nu = i$ where $i = 1, 2$ or 3. Then the left hand side of (6.9) is

$$\partial_\mu E^{\mu i} = \partial_0 E^{0i} + \partial_j E^{ji} = -\frac{1}{c}\frac{\partial}{\partial t} e_i - c\,\frac{\partial}{\partial x^j}\varepsilon_{jik}\, b_k = -\frac{1}{c}\frac{\partial}{\partial t} e_i + c\,\frac{\partial}{\partial x^j}\varepsilon_{ijk}\, b_k\,. \tag{6.19}$$

If we now recall that the curl of a vector is

$$(\vec{\nabla} \times \vec{b}\,)_i = \varepsilon_{ijk}\frac{\partial}{\partial x^j}\, b_k\,, \tag{6.20}$$

where j and k are summed over, and we use the value of κ we found from Gauss' law, then the equation (6.9) with $i = 1, 2, 3$ are equivalent to the three components of the equation

$$-\frac{1}{c^2}\frac{\partial}{\partial t}\,\vec{e} + \vec{\nabla} \times \vec{b} = \mu_0\,\vec{j}\,, \tag{6.21}$$

where μ_0 is the *vacuum permitivity* (or the *magnetic constant*). Eq. (6.21) is Ampere's law for a magnetic field $\vec{b}$ with the Maxwell improvement term. This is the second set of Maxwell equations. Note that the product of the dielectric constant and the vacuum permitivity is

$$\epsilon_0\, \mu_0 = 1/c^2\,. \tag{6.22}$$

[9] e_i and b_i are the components of 3-vectors, so we will always leave the indices down.

[10] In cgs units we would have $\kappa = 4\pi/c$.

6.3 The field potential, gauge invariance and the Bianchi identity

To help find the rest of Maxwell's equations, we introduce a new field, the *gauge potential* A_μ. The field $E_{\mu\nu}$ can be expressed in terms of A_μ as

$$E_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \,. \tag{6.23}$$

Notice that this expression is manifestly antisymmetric under the exchange of μ and ν. Plugging this expression into the field equations we find the new equation

$$\Box A^\nu - \partial^\nu \partial \cdot A = \frac{1}{c\epsilon_0}\, j^\nu \,, \tag{6.24}$$

where $\Box = \partial^\mu \partial_\mu$ and $\partial \cdot A = \partial_\mu A^\mu$.

While we have introduced the gauge potential A_μ, the physical fields are still the electric and magnetic fields, $\vec{e}$ and $\vec{b}$. In fact, if we make the transformation $A_\mu \to A_\mu + \partial_\mu \Phi$, where Φ is an arbitrary scalar field, then

$$E_{\mu\nu} \to \partial_\mu A_\nu - \partial_\nu A_\mu + \partial_\mu \partial_\nu \Phi - \partial_\nu \partial_\mu \Phi = E_{\mu\nu} \,. \tag{6.25}$$

The field $E_{\mu\nu}$, and hence $\vec{e}$ and $\vec{b}$, are invariant under this transformation. Such a transformation is called a *gauge transformation*. Unlike, say, a Lorentz transformation, under a gauge transformation, *all* physical quantities are unchanged. Therefore, we can make a gauge transformation to simplify equations without affecting the physics. For example, we can choose a Φ, such that after the transformation the gauge potential satisfies the equation $\partial \cdot A = 0$. Such a condition is called a *gauge choice*. This particular gauge choice is called the *Lorenz gauge* [11].

Assuming (6.23), we have the following relation:

$$\partial_\lambda E_{\mu\nu} + \partial_\nu E_{\lambda\mu} + \partial_\mu E_{\nu\lambda} = 0 \,. \tag{6.26}$$

To see this explicity, we substitute in the gauge potential into the left hand side of (6.26), after which we find

$$\begin{aligned} &\partial_\lambda(\partial_\mu A_\nu - \partial_\nu A_\mu) + \partial_\nu(\partial_\lambda A_\mu - \partial_\mu A_\lambda) + \partial_\mu(\partial_\nu A_\lambda - \partial_\lambda A_\nu) \\ &= \partial_\lambda \partial_\mu A_\nu - \partial_\mu \partial_\lambda A_\nu + \partial_\nu \partial_\lambda A_\mu - \partial_\lambda \partial_\nu A_\mu + \partial_\mu \partial_\nu A_\lambda - \partial_\nu \partial_\mu A_\lambda = 0 \,. \end{aligned} \tag{6.27}$$

The equation in (6.26) is called the *Bianchi identity*. Actually, this is four equations, since if any of the three free indices are the same then the equation is manifestly zero strictly from the antisymmetry of $E_{\mu\nu}$. Hence, we only have new relations if all three indices are different and there are four ways to choose three distinct indices from four possible values.

[11] Notice that Lorenz is different than Lorentz. Hendrik Lorentz (as in Lorentz transformation) was a Dutch phyiscist, while Ludvig Lorenz was Danish. Nevertheless, a common mistake, even in many famous textbooks, is to call this the *Lorentz* gauge.

6.4 Maxwell's equations: part 2

We now show that the Bianchi identity gives the rest of Maxwell's equations. Let us start by choosing λ, μ, ν =1,2,3 respectively in (6.26). Hence this gives

$$\partial_1 E_{23} + \partial_2 E_{31} + \partial_3 E_{12} = -c \left(\frac{\partial}{\partial x^1} b_1 + \frac{\partial}{\partial x^2} b_2 + \frac{\partial}{\partial x^3} b_3 \right) = 0 \,. \tag{6.28}$$

This then gives the Maxwell equation

$$\vec{\nabla} \cdot \vec{b} = 0 \,, \tag{6.29}$$

which means that there are no magnetic monopole charges in nature.

Now choose λ, μ, ν =0,1,2 respectively in (6.26). This results in

$$\partial_0 E_{12} + \partial_2 E_{01} + \partial_1 E_{20} = \frac{1}{c}\frac{\partial}{\partial t}(-c\, b_3) + \frac{\partial}{\partial x^2} e_1 - \frac{\partial}{\partial x^1} e_2 = -\left(\frac{\partial}{\partial t}\vec{b} + \vec{\nabla} \times \vec{e} \right)_3 = 0 \tag{6.30}$$

This, along with similar equations for λ, μ, ν =0,2,3 and λ, μ, ν =0,3,1 result in Faraday's law for the electric field resulting from a time dependent magnetic field,

$$\frac{\partial}{\partial t}\vec{b} + \vec{\nabla} \times \vec{e} = 0 \,. \tag{6.31}$$

These are the last three Maxwell equations.

6.5 The electromagnetic wave

With the Lorenz gauge, the field equations in (6.24) simplify to

$$\Box A^\nu = \frac{1}{c\,\epsilon_0}\, j^\nu \,. \tag{6.32}$$

In a vacuum there are no charges or currents, thus $j^\mu = 0$. In this case, this equation becomes a simple wave equation

$$\Box A^\nu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} A^\nu - \vec{\nabla} \cdot \vec{\nabla} A^\nu = 0 \,. \tag{6.33}$$

Each component of A^ν is a solution to the same equation. One particular solution to this equation is

$$\begin{aligned} A^\nu(t, \vec{x}) &= \tilde{A}^\nu(\vec{k})\, e^{i\omega(k)t - i\vec{k}\cdot\vec{x}} \\ &= \tilde{A}^\nu(\vec{k})\, e^{ik_\mu x^\mu} \,, \end{aligned} \tag{6.34}$$

where $\omega(\vec{k}) = c|\vec{k}|$ and in the second line we used that $k_\mu = (\omega/c, -\vec{k}\,)$. Hence, this is a wave traveling at the speed of light. In fact, *it is* a light wave! The choice of $\vec{k}$ gives the wavelength and direction for the wave, while $A^\nu(\vec{k})$ is the amplitude. Since the wave equation is linear, a general solution is a sum over all possible values of the wave-vector $\vec{k}$. Since $\vec{k}$ can be anything, the sum is actually an integral and has the form

$$A^\nu(t,\vec{x}) = \int \frac{d^3\vec{k}}{(2\pi)^3}\,\tilde{A}^\nu(\vec{k})\,e^{ik_\mu x^\mu}\,, \tag{6.35}$$

where the factor of $1/(2\pi)^3$ is a standard normalization factor when integrating over wave vectors.

It might seem that there are four independent waves for a given $\vec{k}$ since there are four choices for ν. But the Lorenz gauge leads to the condition

$$\begin{aligned} &\partial_\nu \tilde{A}^\nu(\vec{k})e^{ik_\mu x^\mu} = 0 = ik_\nu \tilde{A}^\nu(\vec{k})e^{ik_\mu x^\mu} \\ \Rightarrow\quad & k_\nu \tilde{A}^\nu(\vec{k}) = k\cdot\tilde{A}(\vec{k}) = 0\,. \end{aligned} \tag{6.36}$$

A general solution to this is

$$\tilde{A}^\nu(\vec{k}) = \Phi(\vec{k})k^\nu + A^\nu_\perp(\vec{k})\,, \tag{6.37}$$

where $\Phi(\vec{k})$ is arbitrary, $A^0_\perp(\vec{k}) = 0$ and $\vec{k}\cdot\vec{A}_\perp(\vec{k}) = 0$. (Note that $A^\mu_\perp = (A^0_\perp, \vec{A}_\perp)$.) The first term on the right hand side is a pure gauge term – it makes no contribution to $\vec{e}$ and $\vec{b}$. This is easy to check:

$$\partial_\mu k_\nu \Phi(\vec{k})e^{ik_\lambda x^\lambda} - \partial_\nu k_\mu \Phi(\vec{k})e^{ik_\lambda x^\lambda} = i(k_\mu k_\nu - k_\nu k_\mu)\Phi(\vec{k})e^{ik_\lambda x^\lambda} = 0\,. \tag{6.38}$$

Hence, the only physical components are the two independent components in $\vec{A}_\perp(\vec{k})$ that are orthogonal to $\vec{k}$. The direction of $\vec{A}_\perp(\vec{k})$ is called the *polarization* of the light wave.

Let us now find the $\vec{e}$ and $\vec{b}$ fields in terms of $\vec{A}_\perp(\vec{k})$. If we let

$$\begin{aligned} \vec{e}\,(t,\vec{x}) &= \int \frac{d^3\vec{k}}{(2\pi)^3}\,\vec{e}\,(\vec{k})\,e^{ik_\mu x^\mu}\,, \\ \vec{b}\,(t,\vec{x}) &= \int \frac{d^3\vec{k}}{(2\pi)^3}\,\vec{b}\,(\vec{k})\,e^{ik_\mu x^\mu}\,, \end{aligned} \tag{6.39}$$

then equation (6.23) and the identifications in (6.14) leads to the relations

$$\begin{aligned} \vec{e}\,(\vec{k}) &= -i\,\frac{\omega}{c}\vec{A}_\perp(\vec{k}) \\ \vec{b}\,(\vec{k}) &= -\frac{i}{c}\,\vec{k}\times\vec{A}_\perp(\vec{k}) = \frac{1}{\omega}\,\vec{k}\times\vec{e}\,(\vec{k})\,. \end{aligned} \tag{6.40}$$

Thus, both $\vec{e}$ and $\vec{b}$ are perpendicular to the direction of the wave and they are perpendicular to each other as well.

6.6 Lorentz transformations of $\vec{e}$ and $\vec{b}$

The fields $\vec{e}$ and $\vec{b}$ are part of an antisymmetric $\binom{0}{2}$ tensor, and so their transformation properties are different than if they were each part of a covariant vector. In particular, the tensor in frame $\mathbf{S}'$ is

$$E_{\mu'\nu'} = \Lambda^\mu{}_{\mu'}\Lambda^\nu{}_{\nu'}E_{\mu\nu}\,. \tag{6.41}$$

If the Lorentz transformation is the boost in the x-direction as given in (3.10), then the various components of the fields are as follows:

$$\begin{aligned}
e'_1 &= E_{0'1'} = \Lambda^\mu{}_{0'}\Lambda^\nu{}_{1'}E_{\mu\nu} = \Lambda^0{}_{0'}\Lambda^1{}_{1'}E_{01} + \Lambda^1{}_{0'}\Lambda^0{}_{1'}E_{10} = \gamma^2 e_1 + \left(\frac{v}{c}\gamma\right)^2(-e_1) = e_1 \\
e'_2 &= E_{0'2'} = \Lambda^\mu{}_{0'}\Lambda^\nu{}_{2'}E_{\mu\nu} = \Lambda^0{}_{0'}\Lambda^2{}_{2'}E_{02} + \Lambda^1{}_{0'}\Lambda^2{}_{2'}E_{12} = \gamma e_2 + \frac{v}{c}\gamma(-cb_3) = \gamma(e_2 - vb_3) \\
e'_3 &= E_{0'3'} = \Lambda^\mu{}_{0'}\Lambda^\nu{}_{3'}E_{\mu\nu} = \Lambda^0{}_{0'}\Lambda^3{}_{3'}E_{03} + \Lambda^1{}_{0'}\Lambda^3{}_{3'}E_{13} = \gamma e_3 + \frac{v}{c}\gamma(+cb_2) = \gamma(e_3 + vb_2) \\
b'_1 &= -\frac{1}{c}E_{2'3'} = -\frac{1}{c}\Lambda^\mu{}_{2'}\Lambda^\nu{}_{3'}E_{\mu\nu} = -\frac{1}{c}\Lambda^2{}_{2'}\Lambda^3{}_{3'}E_{23} = b_1 \\
b'_2 &= -\frac{1}{c}E_{3'1'} = -\frac{1}{c}\Lambda^\mu{}_{3'}\Lambda^\nu{}_{1'}E_{\mu\nu} = -\frac{1}{c}\Lambda^3{}_{3'}\Lambda^1{}_{1'}E_{31} - \frac{1}{c}\Lambda^3{}_{3'}\Lambda^0{}_{1'}E_{30} = \gamma(b_2 + \frac{v}{c^2}e_3) \\
b'_3 &= -\frac{1}{c}E_{1'2'} = -\frac{1}{c}\Lambda^\mu{}_{1'}\Lambda^\nu{}_{2'}E_{\mu\nu} = -\frac{1}{c}\Lambda^1{}_{1'}\Lambda^2{}_{2'}E_{12} - \frac{1}{c}\Lambda^0{}_{1'}\Lambda^2{}_{2'}E_{02} = \gamma(b_3 - \frac{v}{c^2}e_2)\,.
\end{aligned} \tag{6.42}$$

Notice that the components along the boost direction, e_1 and b_1 are unchanged under the boost, while those that are transverse end up mixing the $\vec{e}$ and $\vec{b}$ fields together.

6.7 Invariants

Let us now look for scalar invariants of the electromagnetic fields. Since $E_{\mu\nu}$ has two indices, one invariant we can write down is $E_{\mu\nu}\eta^{\mu\nu}$. However, since $E_{\mu\nu}$ is antisymmetric and $\eta^{\mu\nu}$ is symmetric, the invariant is zero. This is still an invariant, it just happens to be a trivial one.

Next, we consider a product of $E_{\mu\nu}$ fields. In this case we can construct the invariant

$$\begin{aligned}
E_{\mu\nu}E^{\mu\nu} &= E_{\mu\nu}E_{\lambda\rho}\eta^{\mu\lambda}\eta^{\nu\rho} \\
&= 2\sum_{i=1}^{3}(E_{0i})^2\eta^{00}\eta^{ii} + 2\sum_{1\le i<j}^{3}(E_{ij})^2\eta^{ii}\eta^{jj} = -2\sum_{i=1}^{3}(E_{0i})^2 + 2\sum_{1\le i<j}^{3}(E_{ij})^2 \\
&= -2\,(\vec{e}\cdot\vec{e} - c^2\,\vec{b}\cdot\vec{b}\,)\,,
\end{aligned} \tag{6.43}$$

where we have used the symmetry that $E_{\mu\nu}E_{\mu\nu} = E_{\nu\mu}E_{\nu\mu}$ and we have only included those terms in the sum that are nonzero. From this invariant we can infer a few useful facts. For example, if in one frame we have $\vec{e} \neq 0$ and $\vec{b} = 0$, then in no frame is the electric field zero. To see this, we note that $\vec{e}\cdot\vec{e} - c^2\,\vec{b}\cdot\vec{b} > 0$ in this frame. But since this is an invariant, it is true in all frames. Hence $|\vec{e}\,| > c\,|\vec{b}|$ in all frames. Hence $|\vec{e}\,| > 0$ in all frames. We can also say a similar thing if $\vec{e}\cdot\vec{e} - c^2\,\vec{b}\cdot\vec{b} < 0$ or $\vec{e}\cdot\vec{e} - c^2\,\vec{b}\cdot\vec{b} = 0$.

There is another invariant that is quadratic in $E_{\mu\nu}$. But to build this, we need to introduce another tensor, called the ϵ-tensor ("epsilon tensor"), $\epsilon_{\mu\nu\lambda\sigma}$. This tensor is defined to be completely antisymmetric

$$\epsilon_{\mu\nu\lambda\sigma} = -\epsilon_{\nu\mu\lambda\sigma} = -\epsilon_{\lambda\nu\mu\sigma} = -\epsilon_{\sigma\nu\lambda\mu} \tag{6.44}$$

and is normalized to

$$\epsilon_{0123} = 1\,. \tag{6.45}$$

Since it is antisymmetric, all indices have to be different for $\epsilon_{\mu\nu\lambda\sigma}$ to be nonzero. Let us now show that $\epsilon_{\mu\nu\lambda\sigma}$ is a $\binom{0}{4}$ tensor. Consider the determinant of any Lorentz transformation matrix $\Lambda^{\mu}{}_{\mu'}$. The determinant is always 1 and a little thought reveals that

$$\det\Lambda = 1 = \Lambda^{\mu}{}_{0'}\Lambda^{\nu}{}_{1'}\Lambda^{\lambda}{}_{2'}\Lambda^{\sigma}{}_{3'}\,\epsilon_{\mu\nu\lambda\sigma}\,. \tag{6.46}$$

One can also check that the interchange of any two primed indices changes the sign of the expression on the right hand side. Hence, we have that

$$\Lambda^{\mu}{}_{\mu'}\Lambda^{\nu}{}_{\nu'}\Lambda^{\lambda}{}_{\lambda'}\Lambda^{\sigma}{}_{\sigma'}\,\epsilon_{\mu\nu\lambda\sigma} = \epsilon_{\mu'\nu'\lambda'\sigma'}\,, \tag{6.47}$$

and so $\epsilon_{\mu\nu\lambda\sigma}$ indeed transforms as a tensor.

With the ϵ-tensor we can then construct the invariant

$$\begin{aligned} E^{\mu\nu}E^{\lambda\sigma}\epsilon_{\mu\nu\lambda\sigma} &= 8(E^{01}E^{23}+E^{02}E^{31}+E^{03}E^{12}) = 8c(e_1b_1+e_2b_2+e_3b_3) \\ &= 8\,c\,\vec{e}\cdot\vec{b}\,, \end{aligned} \tag{6.48}$$

where we used the antisymmetry of the tensors involved. Hence, $\vec{e}\cdot\vec{b}$ is a Lorentz invariant. Therefore if $\vec{e}\cdot\vec{b}=0$ in one frame, it is zero in all frames.

6.8 The Lorentz force

Let us return to the force equation in (6.1). If we substitute for $E_{\mu\nu}$ the components in (6.14), then for the i^{th} spatial component of F_μ we have

$$\begin{aligned} F_i = -\gamma f_i = \frac{q}{c}E_{i\lambda}u^{\lambda} &= \frac{q}{c}(E_{i0}u^0 + \sum_{j=1}^{3}E_{ij}u^j) = \frac{q}{c}\gamma\,(-ce_i - c\sum_{j,k=1}^{3}\epsilon_{ijk}v_jb_k) \\ &= -q\,\gamma\,(e_i + (\vec{v}\times\vec{b}\,)_i)\,. \end{aligned} \tag{6.49}$$

Hence, we arrive at the Lorentz force equation

$$\vec{f} = q\,(\vec{e}+\vec{v}\times\vec{b}\,)\,. \tag{6.50}$$

Let's use the Lorentz force equation to compute how strong the magnetic fields will be at the LHC when the proton energies are 7 TeV. The protons are traveling in a circle, so while $|\vec{p}|$ is not changing in time, the direction of $\vec{p}$ is. In particular, we have that

$$\left|\frac{d\vec{p}}{dt}\right| = \frac{|\vec{v}\,|}{R}|\vec{p}\,|\,, \tag{6.51}$$

where R is the radius of the ring. As we have seen, the protons are moving very close to the speed of light, so we can replace $|\vec{v}\,|$ with c, and then write

$$\left|\frac{d\vec{p}}{dt}\right| \approx \frac{m_0c^2\gamma}{R} = \frac{7\text{ TeV}}{27/(2\pi)\text{ km}} = \frac{7\times 1.6\times 10^{-7}\text{ joules}}{4.29\times 10^3\text{ m}} = 2.6\times 10^{-10}\text{ joules/m}\,. \tag{6.52}$$

The charge of the proton in SI units is $q = +1.6 \times 10^{-19}$ coulombs. Hence the magnitude of the $\vec{b}$ field is

$$|\vec{b}\,| \approx \frac{1}{qc}\left|\frac{d\vec{p}}{dt}\right| = \frac{1}{(1.6\times 10^{-19}\text{ coulombs})(3\times 10^{8}\text{ m/s})}\,2.6\times 10^{-10}\text{ joules/m} = 5.4\,\text{tesla}\,, \tag{6.53}$$

which is 50,000 times stronger than the earth's magnetic field. Actually, this field is the average field needed around the ring. In reality, the LHC has 1232 magnets of length 14.3 m each. You can quickly check that the total length of the magnets is 17.6 km, covering about 65 % of the ring. Therefore, the field inside each magnet should be

$$(5.4\text{ tesla})\,\times\,(27\text{ km}/17.6\text{ km})\,= 8.3\text{ tesla}\,. \tag{6.54}$$

For the time component of F_μ we have

$$F_0 = \frac{\gamma}{c}\frac{dE}{dt} = \frac{q}{c}E_{0\lambda}u^\lambda = \frac{q}{c}\sum_{j=1}^{3}E_{0j}u^j) = \frac{q\gamma}{c}\vec{e}\cdot\vec{v}\,. \tag{6.55}$$

Hence,

$$\frac{dE}{dt} = q\,\vec{e}\cdot\vec{v}\,. \tag{6.56}$$

Notice that the magnetic field does not change the energy. Since the electromagnetic force is a pure force, the change in energy is a change in kinetic energy only. In fact, we could have derived this from (5.57), where

$$\frac{dE}{dt} = \vec{v}\cdot\vec{f} = q\,\vec{v}\cdot(\vec{e}+\vec{v}\times\vec{b}\,) = q\,\vec{v}\cdot\vec{e}\,, \tag{6.57}$$

since $\vec{v}$ is perpendicular to $\vec{v}\times\vec{w}$, where $\vec{w}$ is any vector.

7 Electromagnetism (advanced topics)

7.1 Potentials from moving charges (Liénard-Wiechert potentials)

In this section we find the gauge potential and the electric and magnetic fields for a moving point charge. First, let us suppose that the charge is static (it is at rest). We write the gauge potential as $A_\mu = (\varphi, -c\,\vec{w})$, in which case we have

$$\begin{aligned}\vec{e} &= -\vec{\nabla}\varphi - c\frac{\partial \vec{w}}{\partial t} \\ \vec{b} &= \vec{\nabla}\times\vec{w}\,.\end{aligned} \tag{7.1}$$

Since the charge is stationary, there is no current and hence no source for a $\vec{b}$ field, thus we set it to zero. We can satisfy this by setting $\vec{w} = 0$, which is consistent with the Lorenz gauge condition $\partial_\mu A^\mu = 0$ if $\frac{\partial\varphi}{\partial t} = 0$. With this choice, the electric field is

$$\vec{e} = -\vec{\nabla}\varphi\,, \qquad \frac{\partial \vec{e}}{\partial t} = 0\,. \tag{7.2}$$

If the charge is q, then

$$\vec{e} = \frac{q}{4\pi\epsilon_0 r^2}\hat{r} = \frac{q}{4\pi\epsilon_0 r^3}\vec{r}, \qquad \varphi = \frac{q}{4\pi\epsilon_0 r}\,, \tag{7.3}$$

where r is the distance to the point charge and $\hat{r}$ is the unit vector in the radial direction. The second relation in (7.3) follows from integrating the first. Let us now show the first. To this end, consider a point charge at the center of a sphere with radius r, as shown in figure 21. By symmetry, the electric field points radially from the center of the sphere and its magnitude is independent of the angle,

$$\vec{e} = e_r(r)\hat{r}\,. \tag{7.4}$$

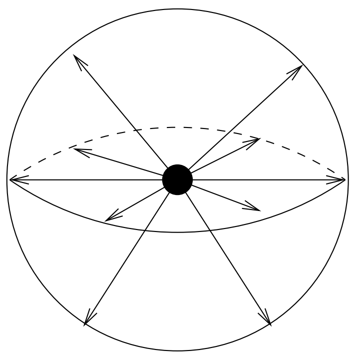


Figure 21: A charge q at the center of a sphere with radius r. The electric field is symmetric and points in the radial direction.

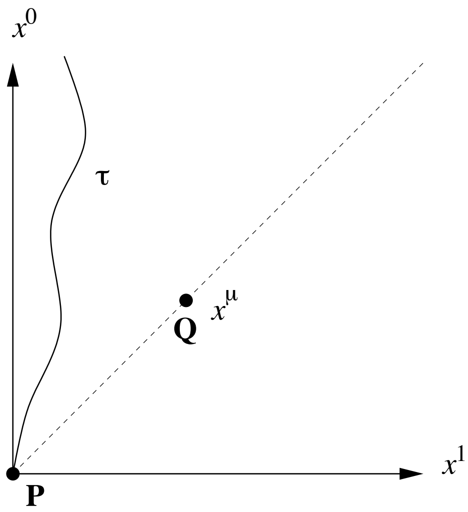


Figure 22: The world-line for a moving charge q. Given the coordinate x^μ at space-time point **Q**, τ is chosen so that $y^\mu(\tau)$ is the coordinate at **P**.

Integrating Gauss' law in eq. (6.18) over the volume of the sphere, gives

$$\begin{aligned} \int_V d^3x\,\vec{\nabla}\cdot\vec{e} &= \int_V d^3x\,\frac{\rho}{\epsilon_0} \\ \boxed{\begin{array}{c}\text{Gauss'Law}\\ \text{(of calculus)}\end{array}} \quad \Rightarrow \quad \int_{\partial V} d\vec{S}\cdot\vec{e} &= \frac{q}{\epsilon_0} \\ 4\pi r^2 e_r &= \frac{q}{\epsilon_0}\,, \end{aligned} \tag{7.5}$$

where we used that the surface area of the sphere is $4\pi r^2$. The first equation in (7.3) then follows.

Now assume that the charge is moving, and that its position as a function of its proper time is $y^\mu(\tau)$. From this, we want to find the fields at the space-time position **Q** with coordinate x^μ. Let us define $\Delta x^\mu = x^\mu - y^\mu(\tau)$, where τ is chosen such that Δx^μ is light-like, that is, $\Delta x^\mu \Delta x_\mu = 0$. This is shown in the space-time diagram in figure 22. In the figure, the event **Q** is on a light-like trajectory originating out of the event **P**. Hence τ in Δx^μ is chosen so that $y^\mu(\tau)$ is at **P**. Since **P** is in the past of **Q**, **P** is called the *retarded event*.

Let us now look for a gauge potential A^μ which is consistent with (7.3) when the particle is at rest. To this end we let

$$\Delta x^\mu = (ct, \vec{r})\,, \tag{7.6}$$

where $ct = r$ so that Δx^μ is light-like. We also define $D = \Delta x_\mu u^\mu$, where u^μ is the 4-velocity of the charged particle at the retarded event **P**. We then let

$$A^\mu = \frac{q\,u^\mu}{4\pi\epsilon_0 D}\,, \tag{7.7}$$

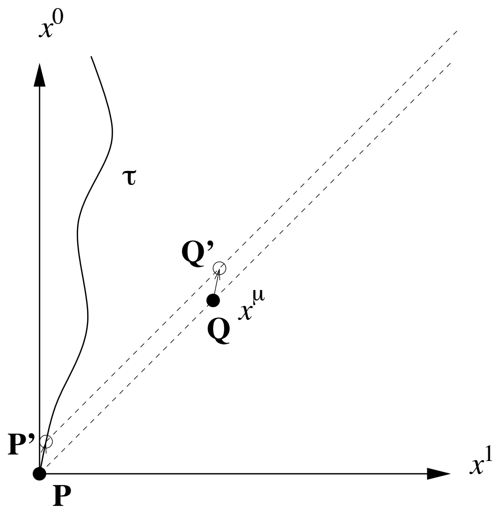


Figure 23: A small displacement of **Q** to **Q'** leads to a displacement of the retarded event from **P** to **P'**.

which we now show is consistent with the particle at rest. In particular, in the rest frame of the particle at **P**, we have $u^\mu = (c, 0, 0, 0)$ and so $D = c^2 t = cr$. Therefore,

$$\varphi = A^0 = \frac{q\,c}{4\pi\epsilon_0 cr} = \frac{q}{4\pi\epsilon_0 r} \qquad w^i = A^i = 0\,, \tag{7.8}$$

matching our previous result for a static charge.

If we now assume that the charge is moving, then $u^\mu = (c\gamma, \vec{v}\gamma)$ and

$$D = \gamma(c^2 t - \vec{v}\cdot\vec{r}\,) = \gamma(cr - \vec{v}\cdot\vec{r}\,)\,, \tag{7.9}$$

and so the potential is

$$\varphi = A^0 = \frac{q\,c\,\gamma}{4\pi\epsilon_0\gamma[cr - \vec{v}\cdot\vec{r}\,]} = \frac{q}{4\pi\epsilon_0[r - \vec{v}\cdot\vec{r}/c]}\,, \tag{7.10}$$

where the brackets $[\dots]$ symbolizes that the velocity and distance to be used are the retarded values, that is the values when the particle is at **P**. The potential is known as the *Liénard-Wiechert* potential for a moving charge.

Next let us find the fields $E_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. This is rather tricky because in order to take a derivative on the gauge potential we displace **Q** by a small amount. But this means we must also displace the retarded event **P** so that Δx^μ remains light-like. This is shown in figure 23.

To find $E_{\mu\nu}$ we will need the derivatives of several types of terms. First, the derivative of the displacement Δx^ν is

$$\partial_\mu \Delta x^\nu = \partial_\mu(x^\nu - y^\nu(\tau)) = \delta^\nu_\mu - \frac{\partial\tau}{\partial x^\mu}\frac{\partial y^\nu(\tau)}{\partial\tau} = \delta^\nu_\mu - \frac{\partial\tau}{\partial x^\mu}u^\nu\,. \tag{7.11}$$

We can then do the following trick:

$$
\begin{aligned}
0 = \partial_\mu(\Delta x_\nu \Delta x^\nu) &= 2\Delta x_\nu \partial_\mu \Delta x^\nu = 2\Delta x_\nu \left(\delta_\mu^\nu - \frac{\partial \tau}{\partial x^\mu} u^\nu\right) \\
&= 2\left(\Delta x_\mu - \frac{\partial \tau}{\partial x^\mu}\Delta x_\nu u^\nu\right) = 2\left(\Delta x_\mu - \frac{\partial \tau}{\partial x^\mu} D\right) . \qquad (7.12)
\end{aligned}
$$

Hence, we have that

$$
\frac{\partial \tau}{\partial x^\mu} = \frac{\Delta x_\mu}{D} \qquad (7.13)
$$

and so

$$
\partial_\mu \Delta x^\nu = \delta_\mu^\nu - \frac{\Delta x_\mu u^\nu}{D} . \qquad (7.14)
$$

Derivatives on u^ν are

$$
\partial_\mu u^\nu = \frac{\partial \tau}{\partial x^\mu}\frac{\partial u^\nu}{\partial \tau} = \frac{\Delta x_\mu}{D}\alpha^\nu , \qquad (7.15)
$$

where α^ν is the 4-acceleration of the charged particle[12]. Combining the results in (7.14) and (7.15) we find for the derivative of D

$$
\partial_\mu D = \partial_\mu(\Delta x^\nu u_\nu) = \left(\delta_\mu^\nu - \frac{\Delta x_\mu u^\nu}{D}\right) u_\nu + \Delta x^\nu \frac{\Delta x_\mu}{D}\alpha_\nu = u_\mu - \frac{\Delta x_\mu c^2}{D} + \frac{\Delta x_\mu \Delta x^\lambda \alpha_\lambda}{D} . \qquad (7.16)
$$

Using (7.15) and (7.16), we find

$$
\begin{aligned}
E_{\mu\nu} &= \frac{\partial}{\partial x^\mu}\left(\frac{q}{4\pi\epsilon_0}\frac{u_\nu}{D}\right) - \frac{\partial}{\partial x^\nu}\left(\frac{q}{4\pi\epsilon_0}\frac{u_\mu}{D}\right) \\
&= \frac{q}{4\pi\epsilon_0}\left(\frac{\Delta x_\mu \alpha_\nu}{D^2} - \frac{u_\nu}{D^2}\left(u_\mu - \frac{\Delta x_\mu c^2}{D} + \frac{\Delta x_\mu \Delta x^\lambda \alpha_\lambda}{D}\right)\right) \\
&\quad - \frac{q}{4\pi\epsilon_0}\left(\frac{\Delta x_\nu \alpha_\mu}{D^2} - \frac{u_\mu}{D^2}\left(u_\nu - \frac{\Delta x_\nu c^2}{D} + \frac{\Delta x_\nu \Delta x^\lambda \alpha_\lambda}{D}\right)\right) \quad \Leftarrow \quad \boxed{u_\mu u_\nu \text{ terms cancel}} \\
&= \Delta x_\mu \left(\frac{q}{4\pi\epsilon_0 D^3}\right)\left(D\alpha_\nu - \Delta x^\lambda \alpha_\lambda u_\nu + c^2 u_\nu\right) \\
&\quad - \Delta x_\nu \left(\frac{q}{4\pi\epsilon_0 D^3}\right)\left(D\alpha_\mu - \Delta x^\lambda \alpha_\lambda u_\mu + c^2 u_\mu\right) . \qquad (7.17)
\end{aligned}
$$

This last line can be written as

$$
E_{\mu\nu} = \Delta x_\mu V_\nu - \Delta x_\nu V_\mu , \qquad (7.18)
$$

[12] In section 2 we called the 4-acceleration a^ν, but to avoid later confusion we have changed it to α^ν

where

$$V_\nu = \frac{q}{4\pi\epsilon_0 D^3}\left(D\alpha_\nu - \Delta x^\lambda \alpha_\lambda u_\nu + c^2 u_\nu\right) . \tag{7.19}$$

We now want to express the fields in terms of the 3-velocity and 3-acceleration. Recall that $u^\nu = (\gamma c, \gamma\vec{v}\,)$ and

$$\alpha^\nu = \frac{du^\nu}{d\tau} = \gamma\frac{du^\nu}{d\tau} = \gamma\left(\frac{d}{dt}(\gamma c), \frac{d}{dt}(\gamma\vec{v}\,)\right) = u^\mu\frac{d\gamma}{dt} + \gamma^2(0,\vec{a}\,)\,, \tag{7.20}$$

where $\vec{a} = \frac{d}{dt}\vec{v}$. Notice that $Du_\nu - \Delta x^\lambda u_\lambda u_\nu = 0$, thus the first term in (7.20) will not contribute to V_ν. Hence, from now on we drop it and simply set $\alpha_\nu = \gamma^2(0,-\vec{a}\,)$ (the minus sign appears because we lowered the index.) Thus, using (7.6) we have that

$$\Delta x^\lambda \alpha_\lambda = -\gamma^2\vec{a}\cdot\vec{r}\,. \tag{7.21}$$

Using this along with the expression for D in (7.9) and the components of u_ν and α_ν we find

$$\begin{aligned}
V_0 &= \frac{q}{4\pi\epsilon_0}\left[\frac{\gamma^3 c\vec{a}\cdot\vec{r} + \gamma c^3}{\gamma^3c^3(r-\vec{v}\cdot\vec{r}/c)^3}\right] \\
&= \frac{q}{4\pi\epsilon_0 c^2}\left[\frac{\vec{a}\cdot\vec{r} + \gamma^{-2}c^2}{(r-\vec{v}\cdot\vec{r}/c)^3}\right] \\
\vec{V} &= \frac{q}{4\pi\epsilon_0}\left[\frac{-\gamma^3c(r-\vec{v}\cdot\vec{r}/c)\vec{a} - \gamma^3(\vec{a}\cdot\vec{r}\,)\vec{v} - \gamma c^2\vec{v}}{\gamma^3c^3(r-\vec{v}\cdot\vec{r}/c)^3}\right] \\
&= -\frac{q}{4\pi\epsilon_0 c^2}\left[\frac{(r-\vec{v}\cdot\vec{r}/c)\vec{a} + (\vec{a}\cdot\vec{r}\,)\vec{v}/c + \gamma^{-2}c\,\vec{v}}{(r-\vec{v}\cdot\vec{r}/c)^3}\right] ,
\end{aligned} \tag{7.22}$$

where the components of V_μ are $V_\mu = (V_0, \vec{V}\,)$. The square brackets still mean that the retarded values are being used. The components of the electric field are

$$e_i = E_{0i} = \Delta x_0 V_i - \Delta x_i V_0 = rV_i + r_i V_0\,, \tag{7.23}$$

hence the electric field is

$$\begin{aligned}
\vec{e} &= r\vec{V} + V_0\vec{r} \\
&= \frac{q}{4\pi\epsilon_0 c^2}\left[\frac{(\vec{a}\cdot\vec{r}\,)(\vec{r}-r\vec{v}/c) - r(r-\vec{v}\cdot\vec{r}/c)\vec{a}}{(r-\vec{v}\cdot\vec{r}/c)^3}\right] + \frac{q}{4\pi\epsilon_0}\left[\frac{(\vec{r}-r\vec{v}/c)}{\gamma^2(r-\vec{v}\cdot\vec{r}/c)^3}\right] \\
&= \qquad\qquad \vec{e}_{\rm rad} \qquad\qquad + \qquad\qquad \vec{e}_{\rm nr}
\end{aligned} \tag{7.24}$$

We have broken up the $\vec{e}$ field into two pieces, $\vec{e}_{\rm rad}$ and $\vec{e}_{\rm nr}$, where the subscripts stand for *radiative* and *nonradiative* respectively. The first term, $\vec{e}_{\rm rad}$ falls off as $1/r$ while the second term, $\vec{e}_{\rm nr}$ falls off as $1/r^2$. Notice as well that $\vec{e}_{\rm rad}$ is linear in $\vec{a}$ while $\vec{e}_{\rm nr}$ has no $\vec{a}$ dependence.

The components of the $\vec{b}$ field are

$$b_i = -\frac{1}{2c}\epsilon_{ijk}E_{jk} = -\frac{1}{2c}\epsilon_{ijk}(\Delta x_j V_k - \Delta x_k V_j) = \frac{1}{c}\epsilon_{ijk}r_jV_k\,, \tag{7.25}$$

where the repeated j and k indices are summed over. Thus we have that $\vec{b}$ is

$$\vec{b} = \frac{1}{c}\vec{r} \times \vec{V} \,. \tag{7.26}$$

If we now compare to $\vec{e}$ in the first line of (7.24) and use that $\vec{r} \times \vec{r} = 0$, we find that

$$\vec{b} = \frac{1}{rc}(\vec{r} \times \vec{e}\,) \,. \tag{7.27}$$

Because of its form in (7.27), $\vec{b}$ satisfies $\vec{r}\cdot\vec{b} = 0$ (for any vectors $\vec{A}$ and $\vec{B}$, $\vec{A}\cdot(\vec{A}\times\vec{B}) = 0$). It is also clear that $\vec{b} \cdot \vec{e} = 0$. In general $\vec{r} \cdot \vec{e} \neq 0$, however it is straightforward to check that $\vec{r} \cdot \vec{e}_{\rm rad} = 0$.

7.2 Example 1: The uniformly moving point charge

We now apply our results derived in the last section. Our first example is a point charge q moving with constant velocity $\vec{v} = v\hat{x}$ along the x axis. Let us further suppose that there is an observer measuring the field from the point charge on the y axis. We let $\hat{\Delta}x^\nu$ be the simultaneous displacement (in the observer's frame) between the observer and the charge. Hence we can write

$$\hat{\Delta}x^\nu = (0, \vec{r}_0) = (0, r_0\cos\theta, r_0\sin\theta, 0)\,, \tag{7.28}$$

and so r_0 is the simultaneous distance between the observer and the charge. Note that $\hat{\Delta}x^\nu$ is *not* the displacement between the observer and the retarded event Δx^ν that was discussed at length in the previous section. This displacement is

$$\Delta x^\nu = (r, \vec{r}\,) = (r, r\cos\phi, r\sin\phi, 0)\,. \tag{7.29}$$

Figure 24 shows these different distances and angles between the charge and the observer. In the figure the observer is making the measurement at space-time point **Q**, **P** is the retarded event for the point charge and **R** is the simultaneous event. The displacement $\vec{d}$ shows the movement of the charge between the time of the retarded event and the time of event **Q**. This elapsed time is $\Delta t = r/c$ and so $\vec{d} = \vec{v}\Delta t = r\vec{v}/c$. From the figure we can see that $\vec{r}_0 = \vec{r} - \vec{d} = \vec{r} - r\vec{v}/c$.

Since the charge has no acceleration, the electric field is

$$\begin{aligned} \vec{e} = \vec{e}_{\rm nr} &= \frac{q}{4\pi\epsilon_0}\left[\frac{(\vec{r} - r\vec{v}/c)}{\gamma^2(r - \vec{v}\cdot\vec{r}/c)^3}\right] \\ &= \frac{q}{4\pi\epsilon_0\gamma^2}\frac{(\vec{r} - r\vec{v}/c)}{(\vec{r}\cdot(\vec{r} - r\vec{v}/c)/r)^3} \\ &= \frac{q}{4\pi\epsilon_0\gamma^2}\frac{\vec{r}_0}{(\vec{r}\cdot\vec{r}_0)^3/r^3}\,. \end{aligned} \tag{7.30}$$

Interestingly, the electric field points along $\vec{r}_0$, that is between the simultaneous position of the charge and the observer, as opposed to the retarded position.

We can do better than this and get rid of the r dependence altogether. First, we can relate r^2 to the other quantities with

$$\begin{aligned} r^2 &= r_0^2 + d^2 + 2dr_0\cos\theta = r_0^2 + r^2\frac{v^2}{c^2} + 2r\,r_0\frac{v}{c}\cos\theta\,, \\ &\Rightarrow \;\gamma^{-2}r^2 - 2r\,r_0\frac{v}{c}\cos\theta - r_0^2 = 0 \end{aligned} \tag{7.31}$$

which we can solve for r in terms of r_0 and v. This results in

$$\begin{aligned} r &= \gamma^2\,r_0\frac{v}{c}\cos\theta + \gamma^2\sqrt{r_0^2\frac{v^2}{c^2}\cos^2\theta + r_0^2/\gamma^2} \\ &= \gamma^2 r_0\left(\frac{v}{c}\cos\theta + \sqrt{1-\frac{v^2}{c^2}\sin^2\theta}\right) \end{aligned} \tag{7.32}$$

where we used that $\gamma^{-2} = 1 - v^2/c^2$ and $1-\cos^2\theta = \sin^2\theta$. We then find

$$\begin{aligned} \vec{r}\cdot\vec{r}_0 &= (\vec{r}_0 + \vec{d}\,)\cdot\vec{r}_0 = r_0^2 + r\,r_0\frac{v}{c}\cos\theta \\ &= r_0^2\left(1+\gamma^2\left(\frac{v}{c}\cos\theta + \sqrt{1-\frac{v^2}{c^2}\sin^2\theta}\right)\frac{v}{c}\cos\theta\right) \\ &= \gamma^2 r_0^2\left(1-\frac{v^2}{c^2}\sin^2\theta + \frac{v}{c}\cos\theta\sqrt{1-\frac{v^2}{c^2}\sin^2\theta}\right) \\ &= r\,r_0\sqrt{1-\frac{v^2}{c^2}\sin^2\theta}\,. \end{aligned} \tag{7.33}$$

Therefore, we reach our final answer for $\vec{e}$,

$$\vec{e} = \frac{q}{4\pi\epsilon_0\gamma^2}\frac{\vec{r}_0}{r_0^3\left(1-\frac{v^2}{c^2}\sin^2\theta\right)^{3/2}}\,. \tag{7.34}$$

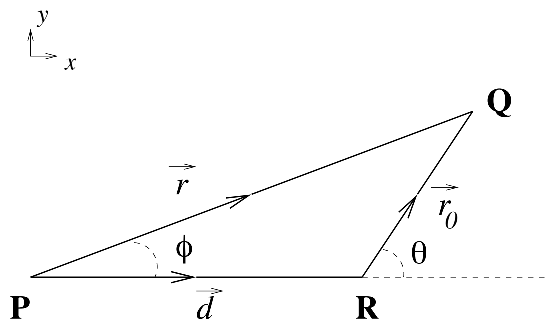


Figure 24: A uniformly moving charge in the x-y plane. The field is measured by the observer at space-time position **Q**. **P** is the retarded event for the charge and **R** is the simultaneous event.

As for the $\vec{b}$ field, we use (7.27) along with

$$\vec{r} \times \vec{r}_0 = \vec{d} \times \vec{r}_0 = r\frac{v}{c} r_0 \sin\theta\, \hat{z}\,, \tag{7.35}$$

to give

$$\vec{b} = \frac{q}{4\pi\epsilon_0\gamma^2 c^2} \frac{v\, \sin\theta\, \hat{z}}{r_0^2 \left(1 - \frac{v^2}{c^2}\sin^2\theta\right)^{3/2}}\,. \tag{7.36}$$

Notice that the $\vec{b}$ field is zero if v is zero, as is expected for a stationary charge. Notice further that $\vec{b}$ is pointing out of the x-y plane and so $\vec{e}\cdot\vec{b} = 0$, also as expected.

7.3 Example 2: The fields from an oscillating charge

Let us consider a single oscillating charged particle[13], with charge q, and assume that its motion is $\vec{x}(t) = d\cos\omega t\, \hat{x}$. Then the velocity and acceleration are

$$\vec{v} = -\omega\, d\, \sin\omega t\, \hat{x}\,, \qquad \vec{a} = -\omega^2 d\cos\omega t\, \hat{x}\,. \tag{7.37}$$

We will assume that the charge is moving at nonrelativistic speeds and so we can approximate $\gamma \approx 1$, $r - \vec{v}\cdot\vec{r}/c \approx r$ and $\vec{r} - r\vec{v}/c \approx \vec{r}$. We also assume that the distance of the observer to the charge is much greater than d. We can then approximate r to be the observer's distance to the origin. We assume that the observer is in the x-y plane at position $\vec{r} = (r\cos\theta, r\sin\theta, 0)$, as in figure 25.

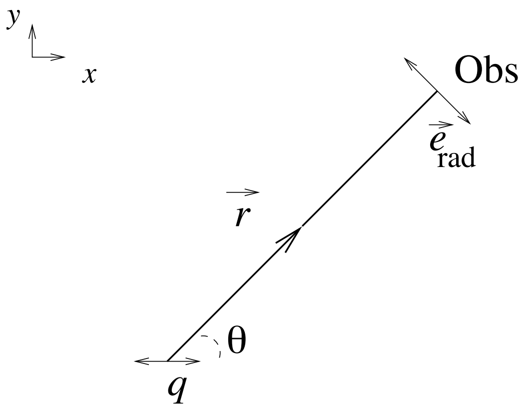


Figure 25: A charge q oscillating back and forth in the x direction. An observer in the x-y plane is a distance r from the origin. The radiative part of the electric field is perpendicular to $\vec{r}$. The magnetic field is perpendicular to both $\vec{r}$ and $\vec{e}$ and is directed out of the x-y plane.

With these assumptions we can approximate the radiative part of the electric field as

$$\vec{e}_{\rm rad} \;\approx\; \frac{q}{4\pi\epsilon_0 c^2}\left[\frac{(\vec{a}\cdot\vec{r})\vec{r} - r^2\vec{a}}{r^3}\right] \tag{7.38}$$

[13] An antenna is a collection of oscillating charges on a conductor.

We should use the retarded value for $\vec{a}$, which for the observer's time t is

$$\vec{a} = -\omega^2 d \cos\omega(t - r/c)\,\hat{x} = -\omega^2 d \cos(\omega t - \vec{k}\cdot\vec{r}\,)\,\hat{x}\,, \tag{7.39}$$

where $\vec{k}$ is directed along r and $|\vec{k}| = \omega/c$. Thus, we have

$$\vec{e}_{\rm rad} \;\approx\; \frac{q\omega^2 d}{4\pi\epsilon_0 c^2}\frac{\sin\theta}{r}\cos(\omega t - \vec{k}\cdot\vec{r}\,)\,(\sin\theta\,\hat{x} - \cos\theta\,\hat{y})\ . \tag{7.40}$$

The term in the last parentheses is a unit vector that is orthogonal to $\vec{r}$, so clearly $\vec{r}\cdot\vec{e}_{\rm rad} = 0$. Observe that there is also an overall factor of $\sin\theta$. Hence the field is zero along the direction of the acceleration and is largest when it is orthogonal to the acceleration. The magnetic field is

$$\vec{b}_{\rm rad} = \frac{1}{rc}\vec{r}\times\vec{e}_{\rm rad} = -\frac{q\omega^2 d}{4\pi\epsilon_0 c^3}\frac{\sin\theta}{r}\cos(\omega t - \vec{k}\cdot\vec{r}\,)\,\hat{z}\,. \tag{7.41}$$

Let's now compare $\vec{e}_{\rm rad}$ to the dominant part of $\vec{e}_{\rm nr}$, which for this nonrelativistic limit is the field we would find if the particle were at rest,

$$\vec{e}_{\rm nr} \approx \frac{q}{4\pi\epsilon_0}\frac{\vec{r}}{r^3} \tag{7.42}$$

If $\theta = \pi/2$, then $|\vec{e}_{\rm rad}| >> |\vec{e}_{\rm nr}|$ if

$$r >> \frac{c^2}{\omega^2 d} = \frac{\lambda^2}{4\pi^2 d}\,. \tag{7.43}$$

where λ is the wavelength of the radiation. Hence the radiation term dominates when r is much greater than λ^2/d, although this bound is itself much greater than λ in the nonrelativistic case where $\lambda >> d$.

7.4 The energy density of the electromagnetic field

Recall that energy is part of the momentum 4-vector $p^\mu = (E/c, \vec{p}\,)$. Recall further that charge density ρ is part of the current density 4-vector $j^\mu = (\rho\, c, \vec{j}\,)$. One can then ask what tensor the *energy density* should belong to. Since electric charge is a scalar and its current has one index, one would suspect that since p^μ already comes with an index, the corresponding current density should have two indices. We write this tensor $T^{\mu\nu}$ and call it the *energy-momentum tensor*. It is also called the *stress tensor*.

Let us make a list of the various components of $T^{\mu\nu}$.

- The energy density is the T^{00} component.
- $c\,T^{0i}$ is the energy current density in the $i^{\rm th}$ direction, in other words, it is the rate per unit area that energy is flowing in direction i. We will also call this the *energy flux*. Note that this has the dimensions of a power per unit area.

- T^{i0}/c is the density of the i^{th} component of momentum. However, it turns out that $T^{i0} = T^{0i}$. Roughly speaking, this is because if there is a momentum density, then energy must be flowing in that direction because whatever it is that has the momentum is moving and it is taking its energy with it.

- T^{ij} is the current density for the i^{th} component of momentum in direction j, which is the same as the current density for the j^{th} component of momentum in direction i and so $T^{ij} = T^{ji}$.

Hence, $T^{\mu\nu}$ is a symmetric tensor, and remains symmetric in all frames.

Let us now try to construct a symmetric $\binom{2}{0}$ tensor out of a combination of $E_{\mu\nu}$ tensors. We first claim that $T^{\mu\nu}$ should be quadratic in $E^{\mu\nu}$. To see this, we note that the fields are linear in the charges, as can be seen from Maxwell equations or the explicit expressions in (7.24) and (7.27). Therefore, the Lorentz force between two particles is proportional to the product of their charges, and so then is the work required to move one particle toward the other one. The work done on the system changes the energy density of the electromagnetic field. Since it is proportional to a product of charges, the energy density must be coming from a product of electromagnetic fields.

The most general symmetric tensor quadratic in $E_{\mu\nu}$ that we can write down is

$$T^{\mu\nu} = a\, E^{\mu\lambda} E_\lambda{}^\nu + b\, \eta^{\mu\nu} E^{\rho\lambda} E_{\lambda\rho}\,. \tag{7.44}$$

Lowering one index this becomes

$$T^\mu{}_\nu = a\, E^{\mu\lambda} E_{\lambda\nu} + b\, \eta^\mu{}_\nu E^{\rho\lambda} E_{\rho\lambda}\,. \tag{7.45}$$

Conservation of energy and momentum leads to a current conservation law for the energy momentum tensor

$$\partial_\mu T^\mu{}_\nu = 0\,. \tag{7.46}$$

Using the field equations in empty space, $\partial_\mu E^{\mu\nu} = 0$, we find

$$\begin{aligned}
\partial_\mu T^\mu{}_\nu &= a\, \partial_\mu (E^{\mu\lambda}) E_{\lambda\nu} + a\, E^{\mu\lambda} \partial_\mu E_{\lambda\nu} + b\, \partial_\nu (E^{\rho\lambda} E_{\rho\lambda}) \\
&= \quad 0 \quad + \ a\, E^{\mu\lambda} \partial_\mu E_{\lambda\nu} + 2b\, E^{\mu\lambda} \partial_\nu E_{\mu\lambda} \quad \Leftarrow \quad \boxed{\text{Relabel dummy index } \rho \to \mu} \\
&= E^{\mu\lambda} (a\, \partial_\mu E_{\lambda\nu} - 2b\, \partial_\mu E_{\lambda\nu} - 2b\, \partial_\lambda E_{\nu\mu}) \quad \Leftarrow \quad \boxed{\text{Bianchi identity (6.26)}} \\
&= E^{\mu\lambda} (a\, \partial_\mu E_{\lambda\nu} - 2b\, \partial_\mu E_{\lambda\nu} - 2b\, \partial_\mu E_{\lambda\nu}) \quad \Leftarrow \quad \boxed{\text{Relabel } \mu \leftrightarrow \lambda \text{ and switch order}} \\
&= E^{\mu\lambda} (a\, \partial_\mu E_{\lambda\nu} - 4b\, \partial_\mu E_{\lambda\nu})\,.
\end{aligned} \tag{7.47}$$

Hence, we find that current conservation in (7.46) requires that $b = a/4$. It turns out that $a = \epsilon_0$, which we will show shortly.

For now, let us see what the various components are in terms of $\vec{e}$ and $\vec{b}$. Using (6.43) for the invariant, we already see that

$$T^{\mu\nu} = \epsilon_0 \left(E^{\mu\lambda} E_\lambda{}^\nu - \tfrac{1}{2} \eta^{\mu\nu} (\vec{e}^{\,2} - c^2 \vec{b}^2) \right)\,. \tag{7.48}$$

The energy density T^{00} is then

$$
\begin{aligned}
T^{00} &= \epsilon_0 \left(E^{0i} E_i{}^0 - \tfrac{1}{2}\eta^{00}(\vec{e}^{\,2} - c^2\vec{b}^2) \right) \\
&= \epsilon_0 \left(\sum_{i=1}^{3} (-e_i)(-e_i) - \tfrac{1}{2}(\vec{e}^{\,2} - c^2\vec{b}^2) \right) \\
&= \frac{\epsilon_0}{2}(\vec{e}^{\,2} + c^2\vec{b}^2)\,. \qquad (7.49)
\end{aligned}
$$

The energy flux components $c\,T^{0i}$ are

$$
\begin{aligned}
c\,T^{0i} &= \epsilon_0\, c \left(E^{0j} E_j{}^i - \tfrac{1}{2}\eta^{0i}(\vec{e}^{\,2} - c^2\vec{b}^2) \right) \\
&= \epsilon_0\, c \left(\sum_{j,k=1}^{3} (-e_j)(\epsilon_{jik} c\, b_k) - \ \ 0 \ \ \right) \\
&= \epsilon_0\, c^2 \sum_{j,k=1}^{3} (\epsilon_{ijk} e_j\, b_k \\
&= \epsilon_0\, c^2 (\vec{e} \times \vec{b}\,)_i \equiv S_i\,. \qquad (7.50)
\end{aligned}
$$

The vector $\vec{S}$ is called the *Poynting vector* and it "points" in the direction of the energy flux. Since the flux comes from a cross product of the $\vec{e}$ and $\vec{b}$ fields, $\vec{S}$ is orthogonal to both.

We now justify the choice $a = \epsilon_0$. Suppose we have a large charge Q that is evenly distributed on a shell of radius R. In this case, the electric field satisfies

$$
\begin{aligned}
\vec{e} &= \frac{Q}{4\pi\epsilon_0}\frac{\vec{r}}{r^3} \qquad r > R \\
\vec{e} &= \ \ 0 \qquad\qquad r \leq R\,. \qquad (7.51)
\end{aligned}
$$

The energy density coming from the electric field according to (7.49) is

$$
\begin{aligned}
T^{00} &= \frac{Q^2 a}{32\pi^2\epsilon_0^2 r^4} \qquad r > R \\
T^{00} &= \ \ 0 \qquad\qquad r \leq R\,, \qquad (7.52)
\end{aligned}
$$

where we have left a unevaluated. Therefore, the energy is

$$
\begin{aligned}
E &= \int d^3x\, T^{00} = \int_R^\infty r^2 dr \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi\, \frac{Q^2 a}{32\pi^2\epsilon_0^2 r^4} \\
&= \frac{4\pi Q^2 a}{32\pi^2\epsilon_0^2} \int_R^\infty \frac{dr}{r^2} = \frac{Q^2 a}{8\pi\epsilon_0^2 R}\,. \qquad (7.53)
\end{aligned}
$$

If we add a small bit of charge δQ then the energy becomes

$$
E = \frac{(Q+\delta Q)^2 a}{8\pi\epsilon_0^2 R} \approx \frac{Q^2 a}{8\pi\epsilon_0^2 R} + \frac{Q\delta Q\, a}{4\pi\epsilon_0^2 R}\ \ . \qquad (7.54)
$$

Now let us think a little differently about adding δQ to Q. Let us start with the charge δQ out at infinity. As we bring it toward the shell there is a Lorentz force on the charge δQ from the $\vec{e}$ field in (7.51). Using (6.50) we have

$$\vec{f} = \frac{Q\delta Q}{4\pi\epsilon_0}\frac{\vec{r}}{r^3} \qquad r > R\,. \tag{7.55}$$

Therefore, the work done on bringing the charge δQ in from infinity to the surface of the shell at $r = R$ is

$$W = \int_R^\infty dr\,\frac{Q\delta Q}{4\pi\epsilon_0 r^2} = \frac{Q\delta Q}{4\pi\epsilon_0 R}\,. \tag{7.56}$$

The work done should equal the increase in the energy of the electromagnetic field. Comparing the work done with (7.54), we see that $a = \epsilon_0$.

7.5 Energy flux from a moving charge

Let us now find the energy flux (Poynting vector) $\vec{S}$ for a moving charge. In this case we have

$$\vec{S} = \epsilon_0\, c^2\, \vec{e}\times\vec{b} = \epsilon_0\, c^2\, \vec{e}\times\left(\frac{1}{rc}(\vec{r}\times\vec{e})\right)\,. \tag{7.57}$$

Using the identity $\vec{A}\times(\vec{B}\times\vec{C}\,) = (\vec{A}\cdot\vec{C}\,)\vec{B} - (\vec{A}\cdot\vec{B}\,)\vec{C}$ we have

$$\vec{S} = \epsilon_0\, c\,(\vec{e}\cdot\vec{e}\,)\,\hat{r} \;-\; (\vec{e}\cdot\hat{r})\,\vec{e}\,. \tag{7.58}$$

In particular, for large enough r where the radiative piece will dominate, $\vec{S}$ reduces to

$$\vec{S} = \epsilon_0\, c\,(\vec{e}_{\rm rad}\cdot\vec{e}_{\rm rad}\,)\,\hat{r} \tag{7.59}$$

since $\hat{r}\cdot\vec{e}_{\rm rad} = 0$. Notice that $\vec{S}$ is directed outward along $\vec{r}$ (the coefficient in front of $\hat{r}$ is nonnegative.)

Let us now apply this to the case of the oscillating charge. Again we assume that the charge's motion is nonrelativisitic so we can use the results in (7.40) and (7.41) for the fields. In this case the energy flux $\vec{S}$ is

$$\begin{aligned}\vec{S} &= \epsilon_0\, c\,(\vec{e}_{\rm rad}\cdot\vec{e}_{\rm rad}\,)\,\hat{r}\\ &= \frac{q^2\omega^4 d^2}{16\pi^2\epsilon_0 c^3}\frac{\sin^2\theta}{r^2}\cos^2(\omega t - \vec{k}\cdot\vec{r}\,)\,\hat{r}\,.\end{aligned} \tag{7.60}$$

To find the total power radiated, we integrate $\vec{S}$ over the surface of a sphere of radius r. Doing this we get

$$\begin{aligned}P &= \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi\, r^2\,\hat{r}\cdot\vec{S} = \int_0^\pi \sin^3\theta d\theta\,\frac{q^2\omega^4 d^2}{8\pi\epsilon_0 c^3}\cos^2(\omega t - \vec{k}\cdot\vec{r}\,)\\ &= \frac{q^2\omega^4 d^2}{6\pi\epsilon_0 c^3}\cos^2(\omega t - \vec{k}\cdot\vec{r}\,)\end{aligned} \tag{7.61}$$

One can consider the averaged power $\bar{P}$ over one oscillation cycle, in which case we replace $\cos^2(\omega t - \vec{k}\cdot\vec{r})$ with $\frac{1}{2}$. Hence, this gives

$$\bar{P} = \frac{q^2\omega^4 d^2}{12\pi\epsilon_0 c^3}\,. \tag{7.62}$$

We can also translate this into an energy per oscillation, given by

$$T\bar{P} = \frac{2\pi}{\omega}\bar{P} = \frac{q^2\omega^3 d^2}{6\epsilon_0 c^3} \tag{7.63}$$

where T is the time for one oscillation. The power is emitted in the form of electromagnetic radiation and this energy is sourced by the oscillating charge. Hence, if no further energy is put into the system the oscillating charge will lose energy.

Notice that the nonradiating piece does not lead to energy loss, since the nonrelativistic $\vec{e}_{\rm nr}$ is directed along $\vec{r}$ and so its contribution to $\vec{S}$ in (7.60) is orthogonal to $\hat{r}$. Therefore, it cannot contribute to $\hat{r}\cdot\vec{S}$.

One can also derive a compact expression for the rate that power is radiated for a general accelerating nonrelativistic charged particle. In this nonrelativistic limit, the radiative part in (7.24) is

$$\vec{e}_{\rm rad} \;=\; \frac{q}{4\pi\epsilon_0 c^2}\left[\frac{(\vec{a}\cdot\vec{r})\vec{r} - r^2\vec{a}}{r^3}\right] \tag{7.64}$$

In this case, the Poynting vector is

$$\vec{S} = \frac{q^2}{16\pi^2\epsilon_0 c^3}\left[\frac{r^2(\vec{a}\cdot\vec{a}\,) - (\vec{a}\cdot\vec{r}\,)^2}{r^4}\right]\hat{r}\,. \tag{7.65}$$

The power radiated, P, is then found by integrating $\vec{S}$ over a sphere of radius r. Without any loss of generality, we can choose the polar angle on the sphere to be the angle away from the direction of $\vec{a}$. Therefore, P is given by

$$\begin{aligned} P \;&=\; \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi\, r^2\,\hat{r}\cdot\vec{S} = \frac{q^2}{16\pi^2\epsilon_0 c^3}\,(1-\cos^2\theta)\,[\vec{a}\cdot\vec{a}\,] \\ &=\; \frac{q^2}{8\pi\epsilon_0 c^3}\left(2-\frac{2}{3}\right)[\vec{a}\cdot\vec{a}\,] \\ &=\; \frac{q^2}{6\pi\epsilon_0 c^3}\,[\vec{a}\cdot\vec{a}\,]\,. \end{aligned} \tag{7.66}$$

One can readily check that substituting the acceleration of an oscillating charge in (7.66) gives (7.62).

7.6 Radiation from a highly relativistic particle on a ring

As we have previously mentioned, the LHC is a particle collider with two opposing beams of 3.5 TeV protons (eventually it will be 7 TeV protons.). The protons themselves are made up of quarks and gluons, and it is actually the collisions between these constituents that will produce new particles, perhaps with masses up to a 1 TeV. However, it will not be easy to observe these new particles. The problem is that all the other constituents of the protons are still there, so the signature of the collision is a big mess and it will take a tremendous effort to extract the interesting information.

In fact, the 27 km LHC tunnel was built for a previous experiment, the Large Electron Positron collider, otherwise known as LEP. In the final version of this experiment the electrons and positrons each had over 80 GeV of energy. The nice thing about an electron-positron collider is that these particles are point particles which means they are not made up of anything else (or at least not at the scales that we can see). Thus when they collide, their signal is very clean. Furthermore, all of the energy is available to make particles. So you might be asking, why did they not just increase the energies of the electron-positron collider so that each particle had 500 GeV in order to create particles with 1 TeV of rest energy? We now show that it is because of energy loss due to radiation. This radiation is known as *synchrotron radiation*.

We are just interested in the radiation term, so we can use equation (7.59) for $\vec{S}$. We assume that the particle is moving at constant speed around a ring of radius R. The velocity vector $\vec{v}$ points along the ring and the acceleration $\vec{a}$ points toward the center of the ring. Taking the result from (7.24), we find

$$\begin{aligned}
\vec{S} &= \epsilon_0\, c\,(\vec{e}_{\rm rad}\cdot\vec{e}_{\rm rad}\,)\,\hat{r} \\
&= \frac{q^2}{16\pi^2\epsilon_0 c^3}\left[\frac{((\vec{a}\cdot\vec{r}\,)(\vec{r}-r\vec{v}/c)-r(r-\vec{v}\cdot\vec{r}/c)\vec{a})\cdot((\vec{a}\cdot\vec{r}\,)(\vec{r}-r\vec{v}/c)-r(r-\vec{v}\cdot\vec{r}/c)\vec{a})}{(r-\vec{v}\cdot\vec{r}/c)^6}\right]\hat{r} \\
&= \frac{q^2}{16\pi^2\epsilon_0 c^3}\left[\frac{r^2(r-\vec{v}\cdot\vec{r}/c)^2\vec{a}\cdot\vec{a}-2r(r-\vec{v}\cdot\vec{r}/c)(\vec{a}\cdot\vec{r}\,)^2+(r^2-2r\vec{v}\cdot\vec{r}+r^2v^2/c^2)(\vec{a}\cdot\vec{r}\,)^2}{(r-\vec{v}\cdot\vec{r}/c)^6}\right]\hat{r} \\
&= \frac{q^2}{16\pi^2\epsilon_0 c^3}\left[\frac{r^2(r-\vec{v}\cdot\vec{r}/c)^2\vec{a}\cdot\vec{a}-r^2(1-v^2/c^2)(\vec{a}\cdot\vec{r}\,)^2}{(r-\vec{v}\cdot\vec{r}/c)^6}\right]\hat{r} \\
&= \frac{q^2}{16\pi^2\epsilon_0 c^3}\left[\frac{r^2\vec{a}\cdot\vec{a}}{(r-\vec{v}\cdot\vec{r}/c)^4}-\frac{r^2(\vec{a}\cdot\vec{r}\,)^2/\gamma^2}{(r-\vec{v}\cdot\vec{r}/c)^6}\right]\hat{r} \qquad (7.67)
\end{aligned}$$

where we used that $\vec{v}\cdot\vec{a}=0$. Observe that the flux is greatest along the direction of $\vec{v}$ because of the denominators. In fact if $\vec{r}$ is parallel to $\vec{v}$ then $\vec{a}\cdot\vec{r}=0$ so only the first term contributes.

Suppose we compute the energy emitted by the charged particle as it travels a length d around the ring, where $d << R$. The space-time point at the beginning of this trajectory is labeled **P** and the point at the end is **Q**. Over such a short distance, the particle's trajectory is well approximated by a straight line. To find the energy emitted, we assume that we have a large sphere of radius r around the center of the particle's trajectory and we integrate the power through this sphere over the time from when the radiation emitted at **P** arrives to when the radiation emitted at **Q** arrives. There is a subtlety that needs to be taken into account and which is described in figure 26 (a). The figure shows the simultaneous circular wave-fronts from the radiation emitted at **P** and **Q**. Since the radiation is emitted at **P** before **Q**, the wave-front from **P** is larger. As is clear from the diagram, the time between the arrival of the **P** wave-front and the **Q**

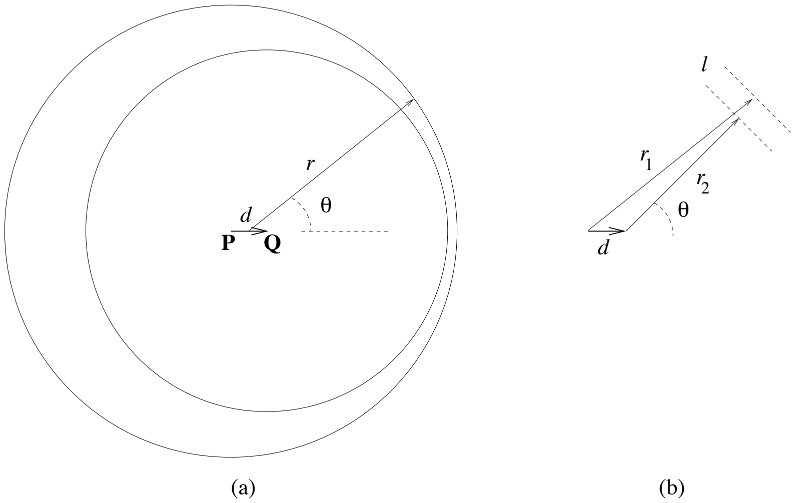


Figure 26: (a) Wave-fronts from a particle emitted at **P** and **Q**. The arrival between the first and second wave-front is shortest along the particle's trajectory. (b) Distance l between wave-fronts at an angle θ.

wave-front is smallest when the polar angle θ is $\theta = 0$, and it is largest when $\theta = \pi$. We can calculate this time for an arbitrary angle. Let r_1 be the radius of the first wave-front and r_2 is the radius of the second wave-front. This is shown in figure 26(b). The wave-front from **Q** is emitted a time d/v after the wave-front from **P** is emitted, therefore, $r_2 = r_1 - dc/v$. Assuming that $d << r_2$ so that the angles of r_2 and r_1 are the same, we then see that the distance between the wavefronts l is

$$l = r_1 - r_2 - d\cos\theta = \frac{dc}{v}\left(1 - \frac{v}{c}\cos\theta\right) , \tag{7.68}$$

and so the time, assuming a highly relativistic particle ($v \approx c$), is

$$t = l/c = \frac{d}{v}\left(1 - \frac{v}{c}\cos\theta\right) \approx \frac{d}{c}\left(1 - \frac{v}{c}\cos\theta\right) . \tag{7.69}$$

Now assume that the particle ring is in the z-x plane and that the trajectory between **P** and **Q** is pointing in the z direction. Therefore,

$$\vec{v} = v\hat{z} , \qquad \vec{a} = a\,\hat{x} , \qquad \vec{r} = \sin\theta\cos\phi\,\hat{x} + \sin\theta\sin\phi\,\hat{y} + \cos\theta\,\hat{z} . \tag{7.70}$$

The magnitude of the acceleration is $a = v^2/R \approx c^2/R$ for highly relativistic particles. Hence, using (7.67), the energy emitted from the particle on the length d trajectory is

$$\begin{aligned} E_{\text{emit,d}} &= \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi\, r^2 \vec{S}\cdot\hat{r}\,\frac{d}{c}\left(1 - \frac{v}{c}\cos\theta\right) \\ &= \frac{q^2 d}{16\pi^2\epsilon_0 R^2}\int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi \left(\frac{1}{(1-\frac{v}{c}\cos\theta)^3} - \frac{\cos^2\phi\sin^2\theta}{\gamma^2(1-\frac{v}{c}\cos\theta)^5}\right) \\ &= \frac{q^2 d}{8\pi\epsilon_0 R^2}\int_{-1}^1 dx \left(\frac{1}{(1-\frac{v}{c}x)^3} - \frac{1}{2}\frac{1-x^2}{\gamma^2(1-\frac{v}{c}x)^5}\right) \quad \Leftarrow \quad \boxed{x = \cos\theta} \end{aligned} \tag{7.71}$$

This integral is dominated near $x = 1$, but its evaluation is slightly tricky. If we write

$$1 - x^2 = \left(1 - \frac{c^2}{v^2}\right) + 2\frac{c^2}{v^2}\left(1 - \frac{v}{c}x\right) - \frac{c^2}{v^2}\left(1 - \frac{v}{c}x\right)^2 , \tag{7.72}$$

and keep only the contribution from the $x = 1$ boundary, we get

$$\begin{aligned} E_{\text{emit,d}} &\approx \frac{q^2 d}{8\pi\epsilon_0 R^2}\left(\frac{c}{2v}\frac{1}{(1-v/c)^2} - \frac{1}{2\gamma^2}\left(\frac{c}{4v}\frac{1-c^2/v^2}{(1-v/c)^4} + \frac{2c^3}{3v^3}\frac{1}{(1-v/c)^3} + \ldots\right)\right) \\ &\approx \frac{q^2 d}{8\pi\epsilon_0 R^2}\left(\frac{1}{2}\frac{1}{(1-v/c)^2} - \frac{1}{2\gamma^2}\left(\frac{1}{4}\frac{(-2)}{(1-v/c)^3} + \frac{2}{3}\frac{1}{(1-v/c)^3}\right)\right) , \end{aligned} \tag{7.73}$$

where the dots indicate a term that is subdominant so it is dropped. In the last step, we substitute v for c wherever it does not lead to a singularity. We can improve this some more but approximating γ^2 as

$$\gamma^2 = \frac{1}{1 - v^2/c^2} = \frac{1}{(1+v/c)(1-v/c)} \approx \frac{1}{2(1-v/c)} . \tag{7.74}$$

This leads to

$$\begin{aligned} E_{\text{emit,d}} &\approx \frac{q^2 d}{8\pi\epsilon_0 R^2}\left(\frac{1}{2} + \frac{1}{2} - \frac{2}{3}\right)\frac{1}{(1-v/c)^2} \\ &\approx \frac{q^2 d\,\gamma^4}{6\pi\epsilon_0 R^2} . \end{aligned} \tag{7.75}$$

To find the power emitted, P, we simply divide $E_{\text{emit,d}}$ by the time it takes the particle to move a distance d, d/c. This gives

$$P = \frac{q^2 c\,\gamma^4}{6\pi\epsilon_0 R^2} . \tag{7.76}$$

To instead get the energy emitted for one complete revolution, we only need to replace d with $2\pi R$, the circumference of the ring. This then gives our final result

$$E_{\text{emit}}/\text{revolution} \approx \frac{q^2\,\gamma^4}{3\,\epsilon_0\,R} . \tag{7.77}$$

Note in particular the γ^4 term. This means that lighter particles will radiate more than heavier ones when their kinetic energies are the same.

Let us now plug in some numbers to see the size of the effect. The radius of the LHC tunnel is 4.3 km. If we consider 7 TeV protons, then $\gamma \approx 7500$. Hence for every revolution, a single proton will radiate

$$\begin{aligned} E_{\text{proton}} &\approx \frac{(1.6\times10^{-19}\ \text{coul})^2(7500)^4}{3(8.85\times10^{-12}\ \text{F/m})(4.3\times10^3\ \text{m})} \\ &= 7.6\times10^{-16}\ \text{joules} = 4800\ \text{eV} . \end{aligned} \tag{7.78}$$

Hence, the proton gives up a very small percentage of its energy to radiation. Next consider an 80 GeV electron at LEP. In this case $\gamma \approx 1.6 \times 10^5$ and so the radiation loss per turn is

$$\begin{aligned} E_{\text{electron}} &\approx \frac{(1.6 \times 10^{-19}\text{ coul})^2 (1.6 \times 10^5)^4}{3(8.85 \times 10^{-12}\text{ F/m})(4.3 \times 10^3\text{ m})} \\ &= 1.5 \times 10^{-10}\text{ joules} = 0.9\text{ GeV}\,. \end{aligned} \tag{7.79}$$

Hence, the electron is radiating one per cent of its energy per turn. Energy has to be supplied to the electrons for each revolution in order to compensate for their radiation loss. This can lead to a huge power consumption.

We now see what we are in store for if we try to have 500 GeV electrons in the LHC tunnel. In this case $\gamma \approx 10^6$ and the radiation loss would be 1.5 TeV per revolution, three times the kinetic energy of the electron! As you might imagine, it is not practical to compensate for such a huge loss. Hence, it is better to go with protons.

7.7 Radiation from a highly relativistic particle on a ring, done differently (and quickly)

In this section we give a quick and dirty way of obtaining the result in the previous section. The idea is the following. In the instantaneous rest frame of the accelerating particle the power radiated can be computed using the nonrelativistic result in (7.66). The acceleration in the instantaneous rest frame is the proper acceleration $\vec{\alpha}$ which satisfies $\alpha^\nu \alpha_\nu = -\vec{\alpha} \cdot \vec{\alpha}$. For the particle going around the circle, γ is not changing, and so from (7.20) we have that $\vec{\alpha} \cdot \vec{\alpha} = \gamma^4\, \vec{a} \cdot \vec{a}$. Thus, the power radiated in the charged particle's frame, P', is

$$P' = \frac{q^2 c \gamma^4}{6\pi\epsilon_0 R^2}\,. \tag{7.80}$$

Going back to the lab frame we have

$$P = \frac{dE}{dt} = \frac{\gamma\, dE'}{\gamma\, dt'} = P'\,, \tag{7.81}$$

matching (7.76). One factor of γ in (7.81) comes from time dilation, $dt = \gamma\, dt'$. The other factor comes from Lorentz transforming the radiation's energy. The differential dE' is part of a 4-vector $dp^{\mu'}$. In the charged particle's rest frame we can see from (7.65) that the radiation is symmetric about the forward and backward direction since $\vec{\alpha}$ is orthogonal to the boost direction. Hence the 4-vector has the form $dp^{\mu'} = (dE'/c, d\vec{p}_\perp^{\,\prime})$, where $d\vec{p}_\perp^{\,\prime} \cdot \vec{v} = 0$. Therefore the boost direction is orthogonal to $d\vec{p}_\perp^{\,\prime}$ and so we have $dp^\mu = (\gamma dE'/c, d\vec{p}_\perp^{\,\prime})$, from which it follows that $dE = \gamma\, dE'$.

8 Special relativity and analytic mechanics

In this section we combine special relativity with what you have learned in the analytic mechanics part of the course.

8.1 The free relativistic particle

Suppose we have a free nonrelativistic particle with mass m. The Lagrangian for this system is given by

$$L = \tfrac{1}{2} m \dot{\vec{x}} \cdot \dot{\vec{x}}\,, \tag{8.1}$$

and the action is

$$S = \int L\,dt = \int \tfrac{1}{2} m \dot{\vec{x}} \cdot \dot{\vec{x}}\,dt\,. \tag{8.2}$$

Let us now consider the relativistic generalization of (8.1). We want to have translation invariance in $3+1$ space-time dimensions as well as Lorentz invariance. This last statement means that the Lagrangian should be a Lorentz invariant. The obvious generalization appears to be

$$L = C\ m \frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\eta_{\mu\nu} = -\,m \frac{dx^\mu}{d\tau}\frac{dx_\mu}{d\tau} = -\,m\,\dot{x}^\mu \dot{x}_\mu\,, \tag{8.3}$$

where C is a constant which we set to $C = -1$ (to be justified later). The action is then

$$S = -\int m\,\dot{x}^\mu \dot{x}_\mu\,d\tau\,. \tag{8.4}$$

But recall from earlier sections that the differentials are constrained by $dx^\mu dx_\mu = c^2 d\tau^2$ and so the Lagrangian in (8.3) is actually a constant,

$$L = -mc^2\,, \tag{8.5}$$

and the action is

$$S = -\int mc^2 d\tau\,. \tag{8.6}$$

Hence, the action is the negative rest energy multiplied by the elapsed proper time along the particle's world-line. By minimizing the action one finds the path that maximizes the proper time. Recall that when we discussed the twin paradox, the twin which took the straight path had a larger proper time than the twin who took a path that had a change in velocity along its world-line, so a maximization of the proper time seems reasonable from this perspective.

Substituting $c^2 d\tau^2 = dx^\mu dx_\mu$, the action can be rewritten as

$$S = -\int mc\sqrt{dx^\mu dx_\mu} = -\int mc\sqrt{\frac{\partial x^\mu}{\partial\lambda}\frac{\partial x_\mu}{\partial\lambda}}\,d\lambda\,, \tag{8.7}$$

where λ can be *any* variable that parameterizes the path of the particle. The Lagrangian is now written as

$$L = -mc\sqrt{\frac{\partial x^\mu}{\partial\lambda}\frac{\partial x_\mu}{\partial\lambda}}\,, \tag{8.8}$$

and the equations of motion are given by

$$\begin{aligned} \frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{x}^\mu}\right) - \frac{\partial L}{\partial x^\mu} &= 0 \\ -\frac{d}{d\lambda}\left(mc\frac{\dot{x}_\mu}{\sqrt{\dot{x}^\nu \dot{x}_\nu}}\right) &= 0\,, \end{aligned} \tag{8.9}$$

where $\dot{x}^\mu = \frac{\partial x^\mu}{\partial\lambda}$. If we let $\lambda = \tau$ then $\sqrt{\dot{x}^\nu \dot{x}_\nu} = c$ and the equations of motion become

$$-\frac{d}{d\tau}\left(m\dot{x}_\mu\right) = 0\,. \tag{8.10}$$

$m\dot{x}_\mu = p_\mu$ is the 4-momentum we discussed in previous sections, and so the equations of motion indicate that it is constant in proper time.

The freedom to choose any variable for λ is known as reparameterization invariance and will lead us to a surprising result (perhaps) when we construct a Hamiltonian. Recall how one constructs the canonical momentum and Hamiltonian from the Lagrangian. The canonical momentum conjugate to a coordinate is found by varying the Lagrangian with respect to its time derivative. In this case λ is playing the role of time, hence we have

$$r_\mu = \frac{\partial L}{\partial \dot{x}^\mu} = -mc\,\frac{\dot{x}_\mu}{\sqrt{\dot{x}^\nu \dot{x}_\nu}}\,. \tag{8.11}$$

and we use r_μ to distinguish the canonical momentum associated with λ from the 4-momentum p_μ above. If we were to let $\lambda = \tau$ then we would have $r_\mu = -p_\mu$. The Hamiltonian is given by

$$H = r_\mu \dot{x}^\mu - L \tag{8.12}$$

and the equations of motion for the coordinate x^μ and the canonical momentum r_μ are found from the Poisson brackets,

$$\begin{aligned} \dot{x}^\mu &= \{x^\mu, H\} \equiv \left(\frac{\partial x^\mu}{\partial x^\nu}\frac{\partial H}{\partial r_\nu} - \frac{\partial x^\mu}{\partial r_\nu}\frac{\partial H}{\partial x^\nu}\right) \\ \dot{r}_\mu &= \{p_\mu, H\} \equiv \left(\frac{\partial r_\mu}{\partial x^\nu}\frac{\partial H}{\partial r_\nu} - \frac{\partial r_\mu}{\partial r_\nu}\frac{\partial H}{\partial x^\nu}\right)\,. \end{aligned} \tag{8.13}$$

However, in our case the Hamiltonian is given by

$$H = \dot{x}^\mu r_\mu - L = -mc\,\frac{\dot{x}^\mu \dot{x}_\mu}{\sqrt{\dot{x}^\nu \dot{x}_\nu}} + mc\,\sqrt{\dot{x}^\nu \dot{x}_\nu} = 0\,, \tag{8.14}$$

in other words, H is constrained to be 0. Why do we get this contraint? Remember that in the usual nonrelativistic case the Hamiltonian is used to generate the time evolution of the coordinates and momenta through the Poisson brackets. In the case here, H is generating a reparameterization of λ. Since the theory is reparameterization invariant, we can change to a new value of λ and leave the values for x^μ and p_μ unchanged. Having $H = 0$ means that x^μ and r_μ stay fixed under the reparameterization.

Even though we have a trivial H, we have an extra conjugate momentum, r_0, as compared to the nonrelativistic case. However, it is possible to get a nontrivial Hamiltonian by choosing a parameterization for λ that evolves in time. To this end we let $\lambda = t$, the time coordinate in frame **S**. For this choice the action becomes

$$S = -\int mc\sqrt{c^2 - \frac{\partial\vec{x}}{\partial t}\cdot\frac{\partial\vec{x}}{\partial t}}\, dt\,, \tag{8.15}$$

and we see that x^0 is no longer a coordinate to be varied in the action and so there is one less conjugate momentum. The canonical momenta for the spatial coordinates with respect to t is

$$r_i = \frac{\partial L}{\partial\frac{\partial x^i}{\partial t}} = mc\frac{\frac{\partial x^i}{\partial t}}{\sqrt{c^2 - \frac{\partial\vec{x}}{\partial t}\cdot\frac{\partial\vec{x}}{\partial t}}} = m\frac{\frac{\partial x^i}{\partial t}}{\sqrt{1 - \frac{1}{c^2}\frac{\partial\vec{x}}{\partial t}\cdot\frac{\partial\vec{x}}{\partial t}}}\,. \tag{8.16}$$

This can be inverted using

$$\vec{r}\cdot\vec{r} = m^2\frac{\frac{\partial\vec{x}}{\partial t}\cdot\frac{\partial\vec{x}}{\partial t}}{1 - \frac{1}{c^2}\frac{\partial\vec{x}}{\partial t}\cdot\frac{\partial\vec{x}}{\partial t}} \qquad \Rightarrow \qquad \frac{1}{c^2}\frac{\partial\vec{x}}{\partial t}\cdot\frac{\partial\vec{x}}{\partial t} = \frac{\vec{r}\cdot\vec{r}}{m^2c^2 + \vec{r}\cdot\vec{r}} \tag{8.17}$$

which then gives

$$\frac{\partial\vec{x}}{\partial t} = \frac{\vec{r}}{\sqrt{m^2 + \vec{r}\cdot\vec{r}/c^2}}\,. \tag{8.18}$$

Note also that using $\vec{v} = \frac{\partial\vec{x}}{\partial t}$ the momentum is

$$\vec{r} = m\,\gamma\,\vec{v}\,, \tag{8.19}$$

which is the anticipated form if $\vec{r}$ is the spatial part of the 4-momentum.

The Hamiltonian is then

$$H = \frac{\partial x^i}{\partial t}r_i - L = \frac{mc^2}{\sqrt{1 - \frac{1}{c^2}\frac{\partial\vec{x}}{\partial t}\cdot\frac{\partial\vec{x}}{\partial t}}} = \sqrt{m^2c^4 + \vec{r}\cdot\vec{r}\,c^2}\,. \tag{8.20}$$

H now has the expected form of the energy in frame **S**, since $\vec{r}$ is the spatial part of the 4-momentum. In terms of $\vec{r}$ and H we have that $p_\mu = (H, -\vec{r})$.

The Hamiltonian now evolves x^i in time t, hence we have

$$\begin{aligned} \frac{\partial x^i}{\partial t} &= \{x^i, H\} = \frac{\vec{r}}{\sqrt{m^2c^2 + \vec{r}\cdot\vec{r}}} = \vec{v} \\ \frac{\partial r_i}{\partial t} &= \{r_i, H\} = 0\,. \end{aligned} \tag{8.21}$$

These are the expected equations of motion. Of course, we are free to choose λ to be the time coordinate for any inertial frame, and since $dx^\mu dx_\mu$ is a Lorentz invariant, we would find similar equations for any inertial frame we choose.

To close this subsection, we note that if $\frac{\partial\vec{x}}{\partial t}\cdot\frac{\partial\vec{x}}{\partial t} << c^2$, then the action in (8.15) is approximately

$$S \approx \int \left(-mc^2 + \tfrac{1}{2}m\frac{\partial\vec{x}}{\partial t}\cdot\frac{\partial\vec{x}}{\partial t}\right) dt\,. \tag{8.22}$$

The first term in the parentheses is a constant and does not affect the equations of motion. The second term is the usual nonrelativistic Lagrangian. Hence, this justifies the above choice $C = -1$.

8.2 Coupling to an external field

We now consider the case where the particle has charge q and is in the presence of an electromagnetic field. The action should be Lorentz invariant and also invariant under gauge transformations. The appropriate action is

$$S = -\int m\,c^2\,d\tau - \frac{1}{c}\int q\,A_\mu dx^\mu\,. \tag{8.23}$$

which is clearly Lorentz invariant. The second integral is a line integral along the particle's world-line. Under a gauge transformation it transforms to

$$-\frac{1}{c}\int q\,A_\mu dx^\mu - \frac{1}{c}\int q\,\partial_\mu\Phi dx^\mu\,. \tag{8.24}$$

The second term in (8.24) is a total derivative and hence cannot affect the equations of motion. Therefore, the equations of motion are invariant under gauge transformations.

Expressing A_μ in terms of the potentials, $A_\mu = (\varphi, -c\,\vec{w}\,)$, the second integral can be written as

$$-\frac{1}{c}\int q\,A_\mu dx^\mu = -\frac{1}{c}\int q\,A_\mu\frac{\partial x^\mu}{\partial t}dt = -\int q\left(\varphi - \vec{w}\cdot\frac{\partial\vec{x}}{\partial t}\right) dt\,. \tag{8.25}$$

The equations of motion are then found to be

$$\begin{aligned} 0 &= \frac{d}{dt}\left(\frac{\partial L}{\partial\frac{\partial x^i}{\partial t}}\right) - \frac{\partial L}{\partial x^i} \\ &= \frac{d}{dt}\left(m\frac{\frac{\partial x^i}{\partial t}}{\sqrt{1-\frac{1}{c^2}\frac{\partial\vec{x}}{\partial t}\cdot\frac{\partial\vec{x}}{\partial t}}} + q w_i\right) + q\frac{\partial\varphi}{\partial x^i} - q\frac{\partial w_j}{\partial x^i}\frac{\partial x^j}{\partial t} \end{aligned} \tag{8.26}$$

Hence, using the expressions for $\vec{e}$ and $\vec{b}$ in (7.1), this becomes

$$\begin{aligned} \frac{dp^i}{dt} = f^i &= -q\left(\frac{\partial\varphi}{\partial x^i} + \frac{\partial w_i}{\partial t}\right) - q\frac{\partial x^j}{\partial t}\frac{\partial w_i}{\partial x^j} + q\frac{\partial w^j}{\partial x^i}\frac{\partial x^j}{\partial t} \\ &= q\,e_i + q\,\varepsilon_{ijk}\,v^j\,b_k \\ \vec{f} &= q(\vec{e} + \vec{v}\times\vec{b})\,, \end{aligned} \tag{8.27}$$

which is the Lorentz force equation.

Note that the $\vec{p}$ that appears here is *not* the canonical momentum. Instead, this is given by

$$r_i = \frac{\partial L}{\partial \frac{\partial x^i}{\partial t}} = p^i + q\, w_i\,. \tag{8.28}$$

Therefore, one finds the Hamiltonian

$$\begin{aligned} H &= r_i \frac{\partial x^i}{\partial t} - L = p_i \frac{\partial x^i}{\partial t} + q\,\phi \\ &= \sqrt{m^2c^4 + (\vec{r} - q\,\vec{w})\cdot(\vec{r} - q\,\vec{w})c^2} + q\,\phi \end{aligned} \tag{8.29}$$

where we essentially replaced $\vec{r}$ with $\vec{r} - q\vec{w}$ in (8.20) and added the contribution from the electric potential.

8.3 Symmetries and constants of the motion for the free particle system

Let us return to the case of the nonrelativistic free particle. The Lagrangian and the action in (8.1) and (8.2) have some symmetries. First, they are invariant under constant shifts in $\vec{x}$,

$$\vec{x}(t) \to \vec{x}(t) + \vec{x}_0\,. \tag{8.30}$$

Notice that if the Lagrangian had had a potential term,

$$L = \tfrac{1}{2}\, m\, \dot{\vec{x}}\cdot\dot{\vec{x}} - V(\vec{x})\,, \tag{8.31}$$

then in general L would not be invariant under the shift in (8.30).

We now show that this invariance leads to a conservation law. Suppose we consider an infinitesimal shift, $\vec{x} \to \vec{x} + \vec{\epsilon}$. Then the change in the Lagrangian is

$$\begin{aligned} \delta L = 0 &= \frac{\partial L}{\partial \dot{\vec{x}}}\cdot\frac{d}{dt}\delta\vec{x} + \frac{\partial L}{\partial \vec{x}}\cdot\delta\vec{x} \\ &= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\vec{x}}}\cdot\vec{\epsilon}\right) - \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\vec{x}}}\right) - \frac{\partial L}{\partial \vec{x}}\right)\cdot\vec{\epsilon}, \end{aligned} \tag{8.32}$$

where we assumed that the change of L is zero because of the invariance. The second term on the second line of (8.32) is zero by the equations of motion, hence if we use that $\vec{\epsilon}$ is independent of t then we find the conservation law

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\vec{x}}} = \frac{d}{dt} m\dot{\vec{x}} = \frac{d}{dt}\vec{p} = 0\,. \tag{8.33}$$

Hence, the translation invariance leads to a conservation equation, in this case conservation of the three-momentum.

The Lagrangian in (8.1) is also invariant under the three-dimensional rotations, $x^i \to R^i{}_j\, x^j$, where $R^i{}_j$ is a three-dimensional rotation matrix and where the components i, j, run from 1 to 3. If we assume that the rotation is very small then we can write $R^i{}_j$ as

$$R^i{}_j = \delta^i{}_j + i\,\epsilon^a\, T_a{}^i{}_j\,, \qquad a = 1\ldots 3\,, \tag{8.34}$$

where the ϵ^a are three independent infinitesimal angles that define the rotation, and the T_a are matrices that "generate" the rotation. The repeated a index in (8.34) is summed over. A useful choice for the independent rotations are rotations in the $y-z$, $z-x$ and $x-y$ planes, in which case we would have

$$T_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & i \\ 0 & -i & 0 \end{pmatrix}, \qquad T_2 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \qquad T_3 = \begin{pmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{8.35}$$

Notice that the components of the T_a can be written as

$$T_a{}^i{}_j = i\,\varepsilon_{aij}\,, \tag{8.36}$$

where ε_{aij} is the completely antisymmetric tensor in three spatial directions,

$$\varepsilon_{aij} = -\varepsilon_{iaj} = -\varepsilon_{jia}, \qquad \varepsilon_{123} = 1\,. \tag{8.37}$$

With the definition in (8.34), the transformation on the components of $\vec{x}$ are

$$x^i \to x^i + \delta x^i = x^i + i\,\epsilon^a T_a{}^i{}_j\, x^j\,. \tag{8.38}$$

Repeating the exercise in (8.32) we find

$$\begin{aligned} \delta L = 0 \quad &= \quad \frac{\partial L}{\partial \dot{\vec{x}}}\cdot\frac{d}{dt}\delta\vec{x} + \frac{\partial L}{\partial \vec{x}}\cdot\delta\vec{x} \\ &= \quad \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}(i\,T_a{}^i{}_j)x^j\right)\epsilon^a - \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\vec{x}}}\right) - \frac{\partial L}{\partial \vec{x}}\right)\cdot\vec{\delta}x\,, \end{aligned} \tag{8.39}$$

Again using the equations of motion and assuming that ϵ^a is constant in t, we find that

$$L_a \equiv \varepsilon_{aij} x^i p^j \tag{8.40}$$

are constants of the motion. These are simply the components of the angular momentum vector

$$\vec{L} = \vec{x}\times\vec{p}\,. \tag{8.41}$$

In both the cases of translations and rotations, we have a continuous symmetry which allows us to construct a set of conserved quantities. In fact for any continuous symmetry of the Lagrangian we could have constructed a conserved quantity. This observation is known as Noether's theorem.

Turning now to the relativistic case, the translation symmetry of the action in (8.7) is very similar to the translation symmetry in (8.2) and leads to the conserved quantities

$$\frac{d}{d\lambda}\frac{\partial L}{\partial \dot{x}^\mu} = \frac{d}{d\lambda} r_\mu = 0\,. \tag{8.42}$$

If we let $\lambda = \tau$ then $r_\mu = -p_\mu$ and the four dimensional space-time translation symmetry leads to conservation of the 4-momentum, $\frac{d}{d\tau}p_\mu = 0$, which includes a component for conservation of energy.

The invariance under Lorentz translations leads to a new set of conserved quantities. Recall that a Lorentz transformation relates the coordinates in frame $\mathbf{S}$ of a space-time point to the coordinates in $\mathbf{S}'$ for the *same* space-time point by

$$(x^{\mu'} - \bar{x}^{\mu'}) = \Lambda^{\mu'}{}_\nu (x^\nu - \bar{x}^\nu)\,, \tag{8.43}$$

where $\bar{x}^\mu$ is a reference point so that the relation is between displacements. From now on we choose the reference point to be at the origin where $\bar{x}^\mu = 0$. Instead of changing the reference frame we want to change the the space-time point so that the coordinates of the new point in $\mathbf{S}$ are the same as if we had kept the point the same and boosted to $\mathbf{S}'$. Hence the transformation of the space-time point is

$$x^\mu \to \Lambda^\mu{}_\nu x^\nu\,. \tag{8.44}$$

Notice that both indices on Λ are unprimed. For an infinitesimal transformation we would have

$$x^\mu \to x^\mu + i\,\tfrac{1}{2}\,\epsilon^{\sigma\rho} M_{\sigma\rho}{}^\mu{}_\nu x^\nu\,, \tag{8.45}$$

where $\epsilon^{\sigma\rho}$ is an infinitesimal angle or rapidity in the $3+1$ dimensional space-time, which satisfies $\epsilon^{\sigma\rho} = -\epsilon^{\rho\sigma}$. For example, the $\{01\}$ component would correspond to an infinitesimal rapidity from boosting in the $\hat{x}$ direction. The generators $M_{\sigma\rho}$ are given by

$$M_{01} = \begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad M_{02} = \begin{pmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad M_{03} = \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}$$
$$M_{12} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad M_{23} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{pmatrix} \qquad M_{31} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix} \tag{8.46}$$

In terms of the components we can write each of these as

$$M_{\sigma\rho}{}^\mu{}_\nu = -M_{\rho\sigma}{}^\mu{}_\nu = i(\eta_{\sigma\nu}\delta^\mu_\rho - \eta_{\rho\nu}\delta^\mu_\sigma)\,. \tag{8.47}$$

If we lower the μ index this takes the more symmetric form

$$M_{\sigma\rho\,\mu\nu} = i(\eta_{\sigma\nu}\eta_{\rho\mu} - \eta_{\sigma\mu}\eta_{\rho\nu})\,. \tag{8.48}$$

The Lagrangian is invariant under the transformation in (8.45), hence we are led to

$$\begin{aligned} \delta L = 0 &= \frac{\partial L}{\partial \dot{x}^\mu}\frac{d}{d\lambda}\delta x^\mu + \frac{\partial L}{\partial x^\mu}\delta x^\mu \\ &= \frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{x}^\mu}(\tfrac{i}{2}\, M_{\sigma\rho}{}^{\mu}{}_{\nu})x^\nu \epsilon^{\sigma\rho}\right) - \left(\frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{x}^\mu}\right) - \frac{\partial L}{\partial x^\mu}\right)\delta x^\mu \,. \end{aligned} \tag{8.49}$$

Therefore, $\Theta_{\sigma\rho}$ is a constant of the motion, where

$$\Theta_{\sigma\rho} = i\, M_{\sigma\rho\mu\nu} r^\mu x^\nu = -x_\sigma r_\rho + x_\rho r_\sigma \,. \tag{8.50}$$

If we now let $\lambda = \tau$ then this becomes

$$\Theta_{\sigma\rho} = x_\sigma p_\rho - x_\rho p_\sigma \,. \tag{8.51}$$

$\Theta_{\sigma\rho}$ is an antisymmetric $\left(\begin{smallmatrix} 0 \\ 2 \end{smallmatrix}\right)$ tensor and is the generalization of the angular-momentum vector. In fact, we can see that $\Theta_{ij} = -\varepsilon_{ijk} L_k$.