“Relativistic Particles and Fields”
Lecture notes, PHS 3131
David Paganin, 2012

CHAPTER 1: DEPARTURES FROM NEWTONIAN MECHANICS

1. The ultimate speed[1]

In this section, we come up against an observation which is contrary to the predictions of Isaac Newton’s mechanics. Such a departure from Newtonian mechanics - one of many! - will be accounted for, in due course, by our exposition of the theory of Relativity. The foundation of this theory, which includes Newtonian mechanics as a limiting case, was largely due to the work of Albert Einstein (see Figure 1).

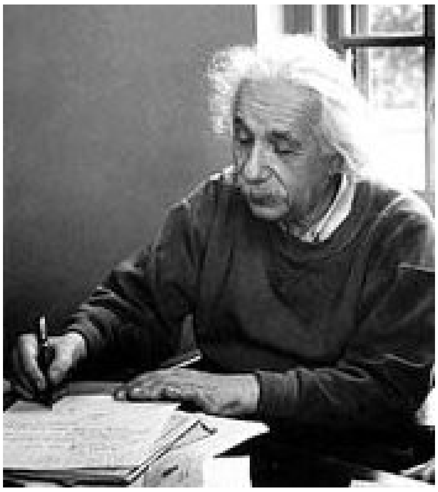


Figure 1: Albert Einstein

The electron is readily accelerated to very high speeds which are many orders of magnitude larger than anything we encounter in our everyday experience. Thus, for example, an electron travelling from cathode to anode of a vacuum tube, with 100 volts of potential difference between the electrodes, would (if it started from rest) arrive at the anode with a speed of about 6000 km/sec.

★ *Exercise 1: Use a calculation based on Newtonian mechanics to predict that an electron, initially at rest at the cathode of a vacuum tube, will arrive at the anode with a speed of about 6000 km/sec, given that the potential difference between anode and cathode is 100 volts.*

Even under these conditions, the predictions of Newtonian mechanics agree well with the outcomes of experiment. However, suppose now that the electron is accelerated through much larger voltages, say, millions of volts rather than hundreds of volts. How does Newtonian mechanics fare in this situation? Let’s turn to nature to give us the answer, by considering the acceleration of electrons to very high energies.

Because time is short, we will not give a detailed description of this experiment[2]. Suffice it to say that the squared speed v^2 of electrons was measured as a function of their kinetic energy K, with the results as shown in Figure 2 below.

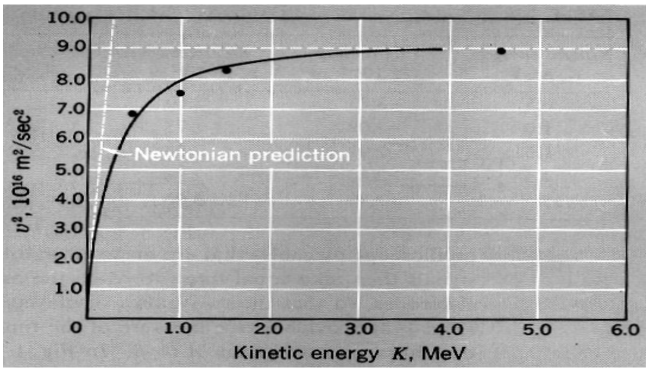


Figure 2: Results of the “ultimate speed” experiment, which plots the squared speed v^2 of the electrons as a function of their kinetic energy K. The departure of these data from the Newtonian prediction is rather dramatic.

Now, what would Newtonian mechanics predict for this experiment? Well, let us write down the Newtonian expression which relates the kinetic energy K of a particle of mass m to its speed v:

$$(1) \quad K = \frac{1}{2} m v^2 .$$

A trivial re-arrangement of this formula leads to:

$$(2) \quad v^2 = \frac{2}{m} K ,$$

from which we predict that a plot of the squared speed v^2 versus kinetic energy K will yield a straight line of gradient $2/m$. This Newtonian prediction is marked in Figure 2, from which it is evident that nature departs from the Newtonian prediction for sufficiently-high kinetic energies of the electron.

---

[1] A. P. French, *Special Relativity*, Chapman and Hall, London (1968), pp. 6-11.

[2] If interested, see W. Bertozzi, *American Journal of Physics*, volume **32**, pages 551-555 (1964).

Apart from the fact that this experiment exhibits a departure from Newtonian mechanics, we also see an indication of the notion of an absolute speed. The data show that an increase in the kinetic energy of the electron leads to an increase in squared speed v^2, but this squared speed appears to asymptote at the value of c^2, where c is the speed of light. This notion of an "ultimate speed" for the electron is once again in contradiction to the predictions of Newtonian mechanics, which predict, via equation (2), that the squared speed will increase, without an upper limit, in direct proportion to the kinetic energy.

2. Einstein's box and the inertia of energy[3]

Einstein made great use of the so-called *gedanken experiment*, which is German for "thought experiment", i.e. an experiment which takes place in the mind of the thinker as opposed to an actual physical experiment which takes place in the laboratory.

Now, one of the most famous of Einstein's thought experiments is the *Einstein Box*, illustrated in Figure 3 below. Here, we see a box of mass M and length L which is isolated from its environment. The box is initially stationary relative to some observer. At some instant of time, Figure 3(a) shows that a burst of light with energy E is emitted from the left end of the box; this energy travels from left to right, towards the right end of the box.

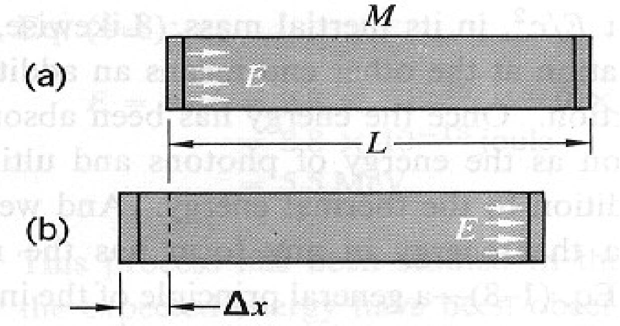


Figure 3: Diagram illustrating the thought experiment known as "Einstein's box".

Now, light is known to carry momentum, as can be demonstrated by measuring the force exerted by light on a reflective foil which is attached to a thin torsion fibre.

★ *Exercise 2: This exercise builds on the previous comment that light carries momentum, by asking you to derive the energy-momentum relation for photons, namely:*

(3) $E = pc$,

where E *is the energy of the photon,* p *is its momentum and* c *is the speed of light. Please show that (3) is a consequence of the well-known equations* $\lambda = h/p$, $E = h\nu$ *and* $c = \nu\lambda$, *where* h *is the Planck constant,* ν *is the photon frequency, and* λ *is its wavelength.*

Since momentum is carried by the burst of light emitted from the left side of the Einstein box, the cherished principle of the conservation of momentum - which we shall not abandon - leads us to conclude that the box recoils to the left with a given constant speed v while the burst of radiation is in flight from the left side of the box to the right side of the box. What is the recoil speed of the box? According to the results of exercise 2, the energy E in the burst of light carries a momentum p given by :

(4) $p = E/c$.

This momentum p of the burst of light is equal and opposite to the recoil momentum $-Mv$ of the box, from which we conclude that the recoil speed v of the box is:

$$(5) \quad v = -\frac{E}{Mc}.$$

The recoiling box travels for a time Δt which is well approximated by L/c; after this time the light burst hits the right side of the box, and is absorbed by the wall, bringing the box to rest once more (see Figure 3(b)). Since the box travelled with the constant speed in (5) for a time L/c, we see that the box has travelled through a distance Δx given by:

$$(6) \quad \Delta x = -\frac{E}{Mc} \times \frac{L}{c} = -\frac{EL}{Mc^2}.$$

Now, we mentioned earlier that the box plus its contents was isolated from its environment. How do we reconcile this to the fact that the box has moved? One way to escape this difficulty is to postulate that a mass m is carried by the burst of radiation, such that the centre of mass of the box + radiation remains fixed:

---

[3] A. P. French, *Special Relativity*, Chapman and Hall, London (1968), pp. 16-20.

$$(7)\quad \frac{mL + M\Delta x}{M + m} = 0$$

If we substitute equation (6) into (7) and solve for the energy E of the light, we arrive at the famous result:

$$(8)\quad E = mc^2 .$$

This equation asserts the inertia of energy, i.e. the fact that a photon of energy E has a mass mc^2. The equation also asserts the energy of inertia, a point which is perhaps made clearer by commenting that our thought experiment converts matter into light energy when the pulse is emitted ("the energy of inertia") and then converts light energy into matter when it is re-absorbed ("the inertia of energy"). We are quickly led to the idea that energy of any form has a mass equivalent given by (8). In this context, let us quote Einstein himself:

> The most important result of a general character to which the special theory has led is concerned with the conception of mass. Before the advent of relativity, physics recognized two conservation laws of fundamental importance, namely, the law of the conservation of energy and the law of the conservation of mass; these two fundamental laws appeared to be quite independent of each other. By means of the theory of relativity they have been united into one law[4].

Examples of the inertia of energy include the fact that a warm can of lemonade is heavier than a cold can, and the fact that a charged capacitor has more mass than an uncharged capacitor. When an electron meets an anti-electron, the two annihilate and create gamma rays of energy equal to $2m_e c^2$, where m_e is the mass of both an electron and an anti-electron at rest. An infamous example of the inertia of energy, as expressed in equation (8), is the conversion of matter energy into radiant energy in nuclear weapons. Less malignantly, this conversion of matter energy into radiant energy via thermonuclear reactions is what powers our sun (see Figure 4, and exercise 3).

★ *Exercise 3: The radiant energy reaching us from the sun has a value of about 1350 watts per square metre. Show that the mass of the sun is decreasing at the rate of about 4.5 million tons per second.*

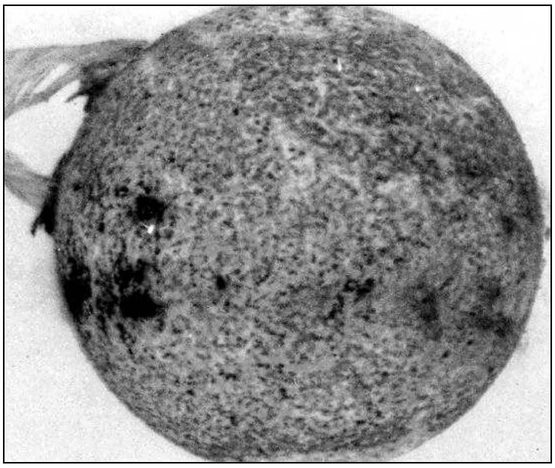


Figure 4: The nuclear energy produced by our sun comes about because of the conversion of rest-mass energy into radiant energy

3. Energy, momentum and mass[5]

Eliminate the energy E from equations (3) and (8), to see that:

$$(9)\quad p = mc$$

for photons. This appears to be a special case of the well-known result from Newtonian mechanics which relates the magnitude of the momentum to the mass and speed:

$$(10)\quad p = mv .$$

If we assume the universality of (8), we may eliminate m from both (8) and (10) to give:

$$(11)\quad E = \frac{pc^2}{v} .$$

Let us put this equation to one side for the moment. Suppose that a force F acts on a particle, with the said particle moving through an infinitesimal distance dx while acted upon by that force. The increment dE of the kinetic energy of the particle is equal to the work (i.e. force F multiplied by displacement dx), and so:

$$(12)\quad dE = F\,dx = \frac{dp}{dt}dx = dp\frac{dx}{dt} = v\,dp ,$$

where t is time, v is speed, and we have used the fact that force is the time rate of change of momentum. Next, solve (11) for v and substitute into (12), to give:

---

[4] A. Einstein, *Relativity*, Crown, New York (1961).

[5] A. P. French, *Special Relativity*, Chapman and Hall, London (1968), pp. 20-24.

$$(13)\quad dE = \frac{pc^2}{E}dp \quad \Rightarrow \quad E\,dE = pc^2 dp\,.$$

Integrate both sides of this equation to give:

$$(14)\quad \tfrac{1}{2}E^2 = \tfrac{1}{2}p^2c^2 + \alpha\,,$$

where the constant of integration α is to be determined.

We shall obtain the value of this constant of integration by demanding that the non-relativistic limit of our calculations should conform with Newtonian mechanics. To this end, solve equation (11) for pc, substitute the resulting equation into (14), and then solve for E, to give:

$$(15)\quad E = \sqrt{\frac{2\alpha}{1-\frac{v^2}{c^2}}}\,.$$

When the speed $v << c$, we have $v/c << 1$ and so we can use the binomial approximation:

$$(16)\quad (1+\theta)^a \approx 1 + a\theta, \quad \theta << 1$$

to approximate (15) by:

$$\begin{aligned} E &= \sqrt{2\alpha}\left(1-\tfrac{v^2}{c^2}\right)^{-1/2} \\ (17)\qquad &\approx \sqrt{2\alpha}\left(1+\tfrac{v^2}{2c^2}\right) \\ &= \sqrt{2\alpha} + \frac{1}{2}\left\{\frac{1}{c^2}\sqrt{2\alpha}\right\}v^2. \end{aligned}$$

This must agree with the Newtonian expression (1) for the kinetic energy. (We do not worry about the additive constant $\sqrt{2\alpha}$ as we only ever measure relative energies.) Therefore, the quantity in braces is equal to the classical mass of the particle, which we denote by m_0. Thus:

$$(18)\quad \frac{\sqrt{2\alpha}}{c^2} = m_0 \quad \Rightarrow \quad \alpha = \tfrac{1}{2}m_0^2c^4\,.$$

Therefore (14) becomes the following important relation between energy, mass and momentum:

$$(19)\quad E^2 = p^2c^2 + m_0^2c^4\,.$$

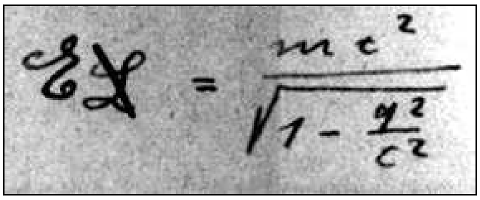
For zero momentum, this equation reduces to (8); we therefore call m_0 the "rest mass".

This phrase, which implicitly suggests that the mass changes with speed, is in harmony with our previous findings regarding the inertia of energy; indeed the inertia of energy leads us to expect the mass to increase with increasing speed. To investigate this point further, first note that substitution of (18) into (15) yields:

$$(20)\quad E = \frac{m_0c^2}{\sqrt{1-\frac{v^2}{c^2}}}\,,$$

an equation which may be combined with (8) to yield a formula for the increase of mass with v:

$$(21)\quad m(v) = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}\,.$$



$$\mathfrak{E} = \frac{mc^2}{\sqrt{1-\frac{q^2}{c^2}}}$$

Figure 5: Equation (20) of our notes, in a slightly different notation, as it appeared in Einstein's original manuscript.

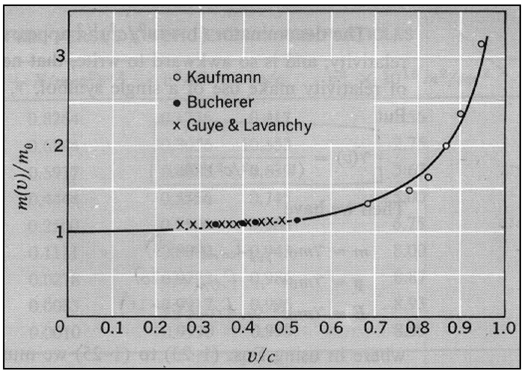


Figure 6: Variation of the inertial mass with speed of the electrons. The straight curve is the prediction of equation (21) for m(v)/m₀ versus v/c, while the data points are the results of experiments by three different groups of workers.

An experimental verification of equation (21) is shown in Fig. 6.

We also point out that (19) reduces to the photon relation (3) when the rest mass is taken to zero.

The kinetic energy K is equal to the difference between the particle energy E and the rest mass energy m_0c^2:

$$
\begin{aligned}
K &= E - m_0c^2 \\
(22) \qquad &= \frac{m_0c^2}{\sqrt{1-\frac{v^2}{c^2}}} - m_0c^2 \text{ ... we used } (20) \\
&= m_0c^2\left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right).
\end{aligned}
$$

This formula can be re-arranged to give the following relation between v^2 and K:

$$(23) \quad v^2 = c^2\left(1-\left(1+\frac{K}{m_0c^2}\right)^{-2}\right).$$

This formula was used to plot the curve of v^2 versus K in Figure 2. As can be seen, the prediction of (23) is in excellent agreement with experiment.

★ *Exercise 4: Suppose that a constant force* F *acts on a body of rest mass* m_0 *for a time* t. *Suppose that the body is initially at rest. Show that the speed* $v(t)$ *at time* t *is given by:*

$$(24) \quad v(t) = \frac{c}{\sqrt{1+\left(\frac{m_0c}{Ft}\right)^2}}.$$

Show that the limiting speed of the particle is equal to the speed of light.

★ *Exercise 5: The "Einstein box" argument we presented has a flaw. We assumed that the entire box was instantaneously set in motion when the burst of light was emitted from the left side (see Figure 3(a)). We therefore implicitly assumed that the influence of the burst was able to travel with infinite speed, from one end of the box to the other. The purpose of this exercise is to refine the "Einstein box" argument to take this criticism into account. To this end, consider the variant of Einstein's box depicted in Figure 7. We replace the box with two initially-stationary masses which are isolated from the environment, and separated by a distance* L. *At some instant the left mass emits a burst of light which travels towards the right mass; the left mass is* m_1 *before emitting the burst and* m_1' *after emitting the burst. The recoil speed of the first mass is* $-v_1$, *as shown. The right mass is* m_2 *before absorbing the light burst and* m_2' *after absorbing the light burst. Using this thought experiment, show that the conclusions of our original "Einstein box" argument are unchanged.*

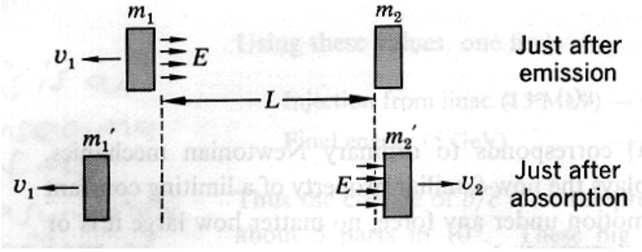


Figure 7: Refined version of the thought experiment known as "Einstein's box".

★ *Exercise 6: What fractional error does one make if one uses the Newtonian expression for the kinetic energy of a body of mass* m *moving at speed* v, *if the speed of that body is: (i) 3 m/s; (ii) 300 m/s (speed of sound in air); (iii) 10,000 m/s (speed needed to escape from the earth); (iv) 0.1c; (v) 0.9c?*

CHAPTER 2: PERPLEXITIES IN THE PROPAGATION OF LIGHT

4. The nature of light and the luminiferous ether[6]

What is the nature of light? One may picture light as a stream of particles emitted from a source, an idea which dates back at least as far as Pythagoras in the sixth century B.C. This theory accounts for the propagation of light in straight lines, the sharpness of shadows, and the ability of light to travel through vacuum.

A different theory emerged in the seventeenth century, as can be seen from the writings of Hooke and Huygens; this theory considered light to be a *wave* rather than a *particle* phenomenon. This wave theory was able to model the phenomena of interference (e.g. the brilliant colours seen on a soap film, together with Young's famous double-slit experiment) and diffraction (e.g. the fine oscillatory structure near the edge of a geometric shadow). The wave picture of radiation dominated nineteenth-century physics, culminating in Maxwell's theory of electrodynamics, which explained light as an electromagnetic wave phenomenon.

Despite the many successes of the wave theory of light, which led to its pre-eminence in the minds of most nineteenth century physicists, there was one serious deficiency of the wave theory when compared to the particle theory: the former theory was unable to explain how light is able to travel through vacuum.

---

[6] A. P. French, *Special Relativity*, Chapman and Hall, London (1968), pp. 38-40.

The waves of our everyday experience require a *medium* through which to travel: sound waves travel through air, water waves travel through water *etc*. It was therefore assumed that there must exist a medium for the light waves to travel through; such a medium was known as the *luminiferous ether.*

The ether was believed to permeate all of space, to be sufficiently tenuous so that the planets could travel through it for year after year with no detectable loss of speed, yet still be capable of very strong restoring forces when displaced from equilibrium (the speed of propagation of a wave in a medium increases with the strength of the said restoring forces.) This replacement of the vacuum with a medium known as the ether is well described by Maxwell himself:

> The vast interplanetary and interstellar regions will no longer be regarded as waste places in the universe ... We shall find them to be already full of this wonderful medium; so full that no human power can remove it from the smallest portion of space, or produce the slightest flaw in its infinite continuity[7].

Continuing in this historical vein, let me quote Sir Oliver Lodge, writing as late as 1925:

> The *air* pressure is a ton to each square foot: an ordinary barometer demonstrates this; it is an instrument for measuring the pressure of the atmosphere. But there is no instrument for measuring the pressure of the Ether, which is probably millions of times greater: its is altogether too uniform for direct apprehension. A deep-sea fish probably has no means of apprehending the existence of water, it is too uniformly immersed in it: and that is our condition in regard to the Ether. But we can feel its vibrations. Hold your hand in front of a blazing fire! ... What we feel is due to Ether vibrations: they excite the nerves of the skin ... It is really the temperature of the skin that we feel, but it is excited by the tremors in the Ether. ... All these skin sensations are directly due to the Ether and its vibrations. The vibrations originated in the Sun, and have travelled 92 million miles of cold empty space, taking eight minutes on the journey before they reach us[8].

Given the existence of the ether, it was quite clear what was meant by the "speed of light", namely "the speed of light relative to the ethereal medium in which it moves". The speed of light would be expected to be quite independent of the motion of the source (for example, the speed of sound in air is unchanged whether the sound is coming from a stationary or a moving aeroplane). However, the perceived speed of light would be expected to be dependent on the motion of the receiver with respect to the ether.

To clarify this point, an analogy might be useful. Replace the ether with the surface of a still lake, the moving source with a pebble thrown into the water, and the moving receiver with a physicist in a moving speedboat. The speed of water ripples, relative to a point (e.g. a cork) which is stationary with respect to the lake's surface, is independent of the speed with which the stone strikes the water. However, the speed of such waves relative to the physicist in the speedboat depends on the speed and direction of the speedboat.

5. The Michelson-Morley experiment[9]

The detection of the earth's motion through the ether was one of the most urgent physical problems of late nineteenth-century physics. The fact that the earth moves around the sun suggested that the speed of the "ether wind"[10] should be at least of the order of the speed of the earth relative to the sun, and should vary by a similar quantity with a period equal to one year.

A famous experiment, designed to detect this motion, is known as the Michelson-Morley experiment (see the "Michelson interferometer" in Figure 8). In this experiment we have a light source S, a pair of mirrors M_1 and M_2, a semi-transparent glass mirror P (which transmits half of the incident light and reflects the other half), and a telescope T. The distance from P to M_1 is l_1; the distance from P to M_2 is l_2. Suppose that the speed of the ether wind is v, moving from right to left, as shown in Figure 8 (we shall consider an arbitrary direction for the ether wind later in our discussion).

Consider light travelling from P to M_1 and back again. On its journey from P to M_1, the speed of light relative to the interferometer is $c - v$; for

---

[7] J. C. Maxwell, quoted in Sir O. Lodge, *Ether and Reality*, Hodder and Stoughton, London (1925), p. 25.

[8] Sir O. Lodge, *Ether and Reality*, Hodder and Stoughton, London (1925), pp. 28-29.

[9] A. P. French, *Special Relativity*, Chapman and Hall, London (1968), pp. 6-11.

[10] Since the earth was believed to be moving through the ether, an observer who is stationary on the surface of the earth will have the ether moving past them, i.e. an "ether wind".

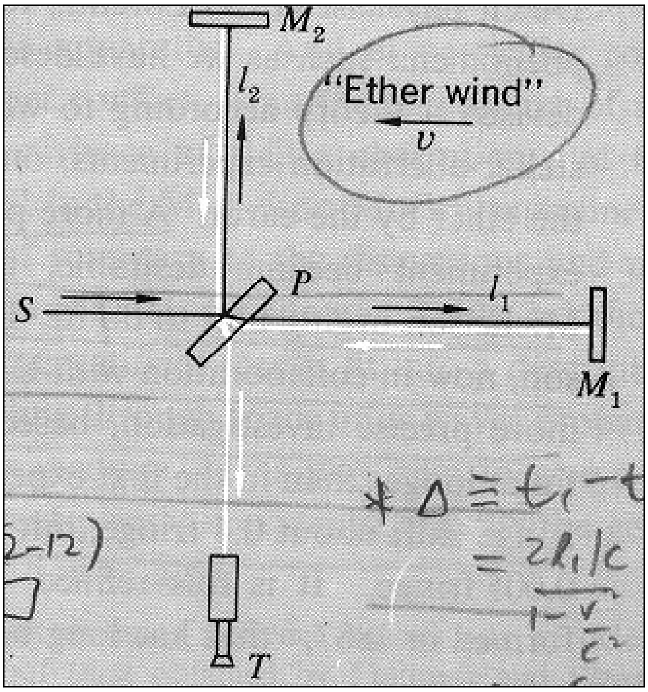


Figure 8: The Michelson-Morley experiment for detecting the effects of the motion of the Earth with respect to the ether.

the return trip, from M_1 to P, the speed of the light relative to the interferometer is $c + v$. Hence the total time t_1 taken for light to travel from P to M_1 and back again is:

$$
\begin{aligned}
t_1 &= \frac{l_1}{c-v} + \frac{l_1}{c+v} \\
&= \frac{l_1(c+v) + l_1(c-v)}{(c-v)(c+v)} \\
&= \frac{2l_1 c}{c^2 - v^2} && (25) \\
&= \frac{2l_1/c}{1 - \frac{v^2}{c^2}}.
\end{aligned}
$$

Similarly, it can be shown that the total time t_2 taken for light to travel from P to M_2 and back again is:

$$(26) \quad t_2 = \frac{2l_2/c}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

★ *Exercise 7: Derive equation (26).*

This defines a time difference Δ which is the difference between t_1 and t_2. For the case where $v << c$, we obtain the following approximate expression for Δ:

$$
\begin{aligned}
\Delta &\equiv t_1 - t_2 \\
&= \frac{2l_1/c}{1 - \frac{v^2}{c^2}} - \frac{2l_2/c}{\sqrt{1 - \frac{v^2}{c^2}}} \\
(27) \quad &= \frac{2}{c}\left(l_1\left(1 - \frac{v^2}{c^2}\right)^{-1} - l_2\left(1 - \frac{v^2}{c^2}\right)^{-1/2} \right) \\
&\approx \frac{2}{c}\left(l_1\left(1 + \frac{v^2}{c^2}\right) - l_2\left(1 + \frac{v^2}{2c^2}\right) \right) \text{... we used (16)} \\
&= \frac{2}{c}(l_1 - l_2) + \frac{v^2}{c^3}(2l_1 - l_2).
\end{aligned}
$$

If the whole apparatus is now rotated through 90^0, so that PM_2 now points along the direction of the ether wind, one can show that there is a new time difference Δ', given by:

$$(28) \quad \Delta' \approx \frac{2}{c}(l_1 - l_2) + \frac{v^2}{c^3}(l_1 - 2l_2).$$

★ *Exercise 8: Derive equation (28).*

This change in time difference would result in a shifting of the fringes of the interference pattern observed by the telescope. This shift would be by an amount of δ fringes, where:

$$
\begin{aligned}
\delta &= \text{number of wavelengths of light} \\
&\quad \text{travelled during the time difference} \\
&= \text{distance travelled by light during the} \\
&\quad \text{time difference, divided by wavelength} \\
&= \frac{c(\Delta - \Delta')}{\lambda} \text{... now use (27) \& (28)} \\
&= \frac{c}{\lambda}\begin{pmatrix} \frac{2}{c}(l_1 - l_2) + \frac{v^2}{c^3}(2l_1 - l_2) \\ -\frac{2}{c}(l_1 - l_2) - \frac{v^2}{c^3}(l_1 - 2l_2) \end{pmatrix} \\
(29) \quad &= \frac{v^2}{\lambda c^2}(l_1 + l_2)
\end{aligned}
$$

Michelson's original experiment of 1881 had $l_1 = l_2 \equiv l$, so let us rewrite (29) as:

$$(30) \quad \delta = \frac{2v^2 l}{\lambda c^2} = \frac{2(v/c)^2}{(\lambda/l)}.$$

In Michelson's experiment, $\lambda \approx 6 \times 10^{-7}$ m and $l \approx 1.2$ m, giving $\lambda/l \approx 5 \times 10^{-7}$. Given the fact that the earth moves around the sun at about 30 km/s = $10^{-4}c$, we expect v/c to be no less than 10^{-4} during at least some part of the year. Accordingly, we expect the fringe shift to obey:

$$(31)\quad \delta > \frac{2\left(10^{-4}\right)^2}{5\times 10^{-7}} = 0.04 .$$

The observation of this shift of 0.04 of a fringe was well within the capability of Michelson's original apparatus, together with the improved apparatus used in the 1887 experiment (which included Morley as a collaborator).

In one of the most famous null results in the history of physics, Michelson announced in 1881 that "there is no displacement of the interference bands", regardless of the season of the year during which the experiment was conducted. No such shift has ever been observed, even in modern experiments which are many orders of magnitude more sensitive than those of Michelson and Morley.

★ *Exercise 9: We have considered the ether-wind theory of the Michelson-Morley experiment, for the special case where the arms of the interferometer are parallel and perpendicular to the ether wind. Consider the general case for an angular setting θ as shown in Figure 9. Show that, for arms of equal length, the time difference for the two paths is, to a good approximation:*

$$(32)\quad \Delta t(\theta) \approx \frac{v^2 \ell}{c^3}\cos(2\theta).$$

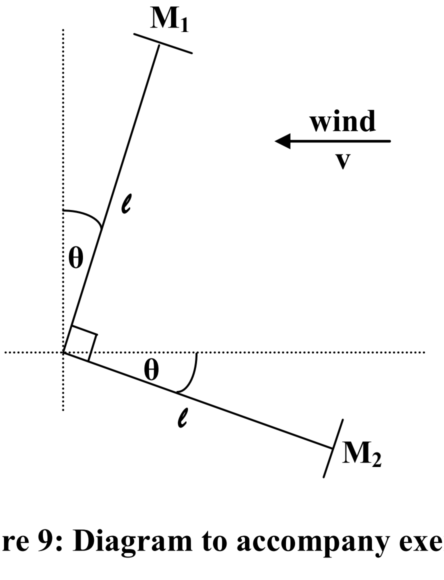


Figure 9: Diagram to accompany exercise 9.

CHAPTER 3: THE FOUNDATIONS OF SPECIAL RELATIVITY[11]

6. Inertial reference frames

A frame of reference is a conventional standard of rest relative to which measurements can be made and experiments described. For example, if we choose a frame rigidly attached to the earth, the various points of the earth remain at rest in this frame, while the stars trace out circles about this frame in the course of a day. If, on the other hand, we choose a rigid reference frame attached to the stars then these stars remain at rest while points on the earth (other than points on the earth's axis!) trace out approximate circles in the course of each day, with the earth itself tracing out an ellipse in the course of each year.

For much of this course, we will make use of a particular class of frames of reference which are known as "inertial reference frames" or, more succinctly, "inertial frames". Such frames of reference are defined as follows:

> An inertial frame is one in which spatial relations, as determined by rigid scales at rest in the frame, are Euclidean and in which there exists a universal time in terms of which free particles remain at rest or continue to move with constant speed along straight lines (i.e. in terms of which free particles obey Newton's first law).

Free particles which are placed without velocity at fixed points in an inertial reference frame will remain at those points, by definition. We can therefore picture an inertial reference frame as a regular lattice of free test particles which are mutually at rest; the distances between these "defining" particles satisfy the Euclidean axioms. We can further picture each of these defining particles (which we identify with points in a given coordinate system) to carry a clock that indicates the universal time throughout the frame.

7. Einstein's two axioms for Special Relativity

Einstein's first axiom for Special Relativity is intimately related to the concept of inertial references frames:

> **Einstein's first axiom:** The laws of physics are identical in all inertial frames, or, equivalently, the outcome of any physical experiment is the

---

[11] This chapter of the lecture course is closely based on W. Rindler, *Introduction to Special Relativity*, Oxford University Press, Oxford (1991), chapter 1.

> same when performed with identical initial conditions relative to any inertial reference frame.

This first axiom represents a generalization to the whole of physics of a relativity principle long known to be satisfied by Newtonian mechanics.

As an example of the relativity principle in *Newtonian* mechanics, consider a physicist who is inside a windowless spaceship which is travelling at uniform speed through space. This physicist has a set of billiard balls, springs, inclined planes etc. with which she can perform experiments. According to the laws of Newtonian mechanics, there is no way of using the outcome of experiments to determine the state of motion of the rocket.

This first axiom seems harmless, and even conservative, in light of the fact that such a relativity principle was long known to be obeyed by Newtonian mechanics.

Let us change tack for a moment. One of the greatest triumphs of Maxwell's electromagnetic theory was the explanation of light as an electromagnetic wave phenomenon. But waves in what? We have already considered the theory of the ether, and the Michelson-Morley experiment for detecting the effects of this putative absolute frame of reference for electromagnetic waves. Now, *the existence of such an absolute frame of reference is evidently in contradiction to Einstein's first axiom*. We have already seen that the null result of the Michelson-Morley experiment was perplexing from the point of view of the ether theory. Attempts were made to reconcile this null result by use of an appropriately-modified ether theory, but they could still not resolve the problem.

Rather than looking upon the null result of the Michelson-Morley experiment as an anomaly which needed to be explained within the context of a modified theory of the ether, *Einstein considered this null result as uncovering a hitherto unknown principle of nature*. Indeed, he elevated this null result into his "law of light propagation". This law is arrived at via Einstein's second axiom of Special Relativity:

> **Einstein's second axiom:** There exists an inertial frame in which light signals in vacuum always travel rectilinearly at constant speed c, in all directions, independently of the motion of the source.

Combine this with the first axiom, to arrive at:

> **Einstein's law of light propagation:** Light signals in vacuum are propagated rectilinearly, with the same speed c, at all times, in all directions, in all inertial frames.

A logical consequence of Einstein's two axioms was to eliminate the ether concept from physics. There are a plethora of further consequences, all of which can be logically derived from Einstein's two axioms, and the most important of which will occupy us for the remainder of this chapter.

8. Events

We need to give a precise meaning to the concept of an "event", for it plays a very important role in Special Relativity. We consider an "event" (more precisely, a "instantaneous point event") to be completely characterised by its position and time, e.g. (t, x, y, z), where t is time and (x, y, z) are the usual Cartesian coordinates.

How are these four coordinates ("space-time coordinates") to be assigned to particular events? Suppose there is a presiding observer at the origin of a given reference frame. The distance $r = \sqrt{x^2 + y^2 + z^2}$ to any particle occupying a position (x, y, z) can be determined by the radar method of bouncing a light echo off the particle, and then multiplying the elapsed time by $\frac{1}{2}c$.

★ *Exercise 10: Derive the just-mentioned result for measuring distance from the presiding observer to the particle at position (x,y,z).*

Angle measurements with an appropriate surveying instrument ("theodolite") then allow the polar angles (θ, ϕ) in spherical polar coordinates to be determined. Equipped with measurements of (r, θ, ϕ) the presiding observer can then determine the Cartesian coordinates of the particle using:

(33) $(x, y, z) = (r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)$.

The same radar signal can also be used to determine the time t of the reflection event, as the average of the time of emission and the time of reception of the radar pulse.

9. The Lorentz transformations

In the previous section we discussed a method for assigning space-time coordinates (t,x,y,z) to events in a given inertial reference frame. Such coordinates will be called *standard coordinates* for an inertial reference frame. We shall now consider the transformation of the standard coordinates of a given event from one inertial frame to another.

We now set up the standard coordinates in two inertial frames S and S' so as to be in a "standard configuration" with each other, as shown in Figure 10. We (i) choose the line of motion of the spatial origin of S' as the x-axis of S; (ii) choose the time settings in S and S' so

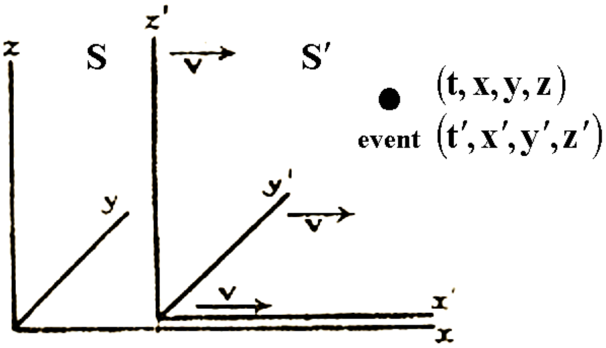


Figure 10: Two inertial reference frames, one of which moves with constant velocity with respect to the other

so that the two spatial origins coincide when both their clocks read zero; (iii) choose the directions of the x and y axes so that these axes, for both frames, coincide when both clocks read zero . The frame S' moves with some constant velocity v in the x-direction of S.

The "standard Galilean transformations" for transforming the coordinates (t, x, y, z) of the event relative to S, into the coordinates (t', x', y', z'), are as follows:

$$(34)\quad \begin{cases} t' = t \\ x' = x - vt \\ y' = y \\ z' = z \end{cases}$$

The form of these transformations corresponds to what might be called "pre-relativistic common sense". In particular, the time coordinate of the event is unaffected by the change of reference frame, in line with Newton's dictum that "Absolute, true and mathematical time, of itself, and by its own nature, flows uniformly on, without regard to anything external."

If we differentiate the Galilean transformations with respect to time $t' = t$, we obtain the "common-sense" classical velocity transformation laws, which relates the velocity components of an arbitrarily moving particle in S with those in S':

$$(35)\quad \begin{cases} \dfrac{dx'}{dt'} = \dfrac{dx}{dt} - v \\ \dfrac{dy'}{dt'} = \dfrac{dy}{dt} \\ \dfrac{dz'}{dt'} = \dfrac{dz}{dt} \end{cases}$$

Thus, for example, if a particle has velocity α along the x' axis then it has the velocity $\alpha + v$ in S; this is the "common sense" law for velocity addition.

Yet if we set $\alpha = c$, where c is the speed of light, then we come in conflict with Einstein's law of light propagation. Evidently, then, the two axioms of Special Relativity imply the need to modify the Galilean transformations. The modified form of the transformation which conform to Einstein's axioms are known as the Lorentz transformations, which we now derive[12].

Suppose that a light burst is travelling along the positive x axis; in S, therefore, the position x of the light burst will obey:

$$(36)\quad x - ct = 0.$$

In the frame S', by axioms 1 and 2, we have:

$$(37)\quad x' - ct' = 0.$$

Hence,

$$(38)\quad x' - ct' = \lambda(x - ct),$$

where λ is a constant. Using a similar argument applied to a burst of light travelling along the negative x axis, we get:

$$(39)\quad x' + ct' = \mu(x + ct),$$

where μ is another constant. The sum and difference of equations (38) and (39) leads to the respective equations:

---

[12] This derivation of the Lorentz transformation follows the presentation given by Einstein himself in Albert Einstein, *Relativity: The Special and General Theory*, Appendix I, available online at http://www.bartleby.com/173/a1.html

(40a) $x' = \frac{1}{2}(\lambda + \mu)x - \frac{1}{2}(\lambda - \mu)ct$,

(40b) $ct' = -\frac{1}{2}(\lambda - \mu)x + \frac{1}{2}(\lambda + \mu)ct$.

If we now let:

(41a) $a \equiv \frac{1}{2}(\lambda + \mu)$,

(41b) $b \equiv \frac{1}{2}(\lambda - \mu)$,

then equations (40) become:

(42a) $x' = ax - bct$,

(42b) $ct' = act - bx$.

Now suppose that we are sitting at the origin of the S' axis, where $x' = 0$. For this case, (42a) becomes:

(43) $ax = bct$, $x' = 0$,

which can be re-arranged to give:

(44) $\frac{x}{t} = \frac{bc}{a}, \quad x' = 0$.

But the left side of (44) is v, i.e. the relative velocity of the two reference frames; hence (44) becomes:

(45) $v = \frac{bc}{a}, \quad x' = 0$.

Even though equation (45) was derived under the restriction $x' = 0$, it is evidently a relation among constants; hence (45) is valid without restriction. We may therefore use (45) to eliminate b from equations (42), leaving:

(46a) $x' = a(x - vt)$,

(46b) $ct' = a\left(ct - \frac{vx}{c}\right)$.

Now suppose that a unit-length measuring rod is sitting at the origin of each of the coordinate systems S and S', the said rods being aligned with the x and x' axes, respectively. The principle of relativity teaches us that, as judged from S, the length of a unit measuring-rod which is at rest with respect to S' must be exactly the same as the length, as judged from S', of a unit measuring-rod which is at rest relative to S.

Let us see how the points of the x' axis appear as viewed from S, by taking a "snapshot" of S' from S; this means that we have to insert a particular value of t (time of S) into (46a); we choose $t = 0$. At time $t = 0$, equation (46a) becomes:

(47) $x' = ax, \quad t = 0$,

which leads to the following formula for coordinate differences:

(48) $\Delta x' = a\Delta x, \quad t = 0$.

Since we are looking at the unit measuring rod of S', let $\Delta x' = 1$; hence (48) becomes:

(49) $\Delta x = \frac{1}{a}, \quad t = 0$

for the apparent length of the S' rod with respect to the frame S.

Next, we eliminate t from equations (46a); this is done by solving (46a) for t, then solving (46b) for t, and letting the two expressions equal one another, leading to:

(50) $\frac{t'v}{a} + \frac{x'}{a} = x\left(1 - \frac{v^2}{c^2}\right)$.

Since the observer in S' is taking a snapshot at $t' = 0$, equation (50) becomes:

(51) $\frac{x'}{a} = x\left(1 - \frac{v^2}{c^2}\right), \quad t' = 0$,

the difference form of which is:

(52) $\frac{\Delta x'}{a} = \Delta x\left(1 - \frac{v^2}{c^2}\right), \quad t' = 0$.

If the observer in S' is looking at the unit rod attached to S, then $\Delta x = 1$ and (52) becomes:

(53) $\Delta x' = a\left(1 - \frac{v^2}{c^2}\right), \quad t' = 0$

for the apparent length $\Delta x'$ of the S unit rod with respect to the frame S.

By symmetry, the apparent length Δx of S''s unit rod as observed from S must be the same as the apparent length $\Delta x'$ of S's unit rod as observed from S'. Hence we may equate the right-hand sides of equations (49) and (53), to obtain:

(54) $a = \dfrac{1}{\sqrt{1-\dfrac{v^2}{c^2}}}$.

Thus equations (46a) and (46b) become:

(55a) $x' = \gamma(x - vt)$,

(55b) $t' = \gamma\left(t - \dfrac{vx}{c^2}\right)$,

where we have introduced the standard notation:

(55c) $\gamma \equiv \dfrac{1}{\sqrt{1-\dfrac{v^2}{c^2}}}$ ("Lorentz factor").

The extension of this result, to include events which take place outside the x axis, is obtained by retaining equations (55) and supplementing them by the relations:

(55d) $y' = y$, (55e) $z' = z$.

Equations (55) are the famous *Lorentz transformations* which play a pivotal role in the theory of Special Relativity.

For later reference, we collect these transformations, together with their inverses, in one place:

$$\text{(56)}\quad \begin{cases} t' = \gamma\left(t - \dfrac{vx}{c^2}\right) \\ x' = \gamma(x - vt) \\ y' = y \\ z' = z \end{cases}, \quad \text{(57)}\quad \begin{cases} t = \gamma\left(t' + \dfrac{vx'}{c^2}\right) \\ x = \gamma(x' + vt') \\ y = y' \\ z = z' \end{cases}.$$

★ *Exercise 11: Derive the inverse Lorentz transformations from the forward Lorentz transformations, using both a "brute force" approach and an approaches which makes use of a relativistic symmetry.*

★ *Exercise 12: Prove the following statement: "Any relativistic transformation formula which relates unprimed and primed quantities (from* S *and* S' *respectively) remains valid when* v *is replaced by* $-v$*, and primed quantities are interchanged with unprimed quantities." This symmetry is called a "v reversal", and we shall make use of a number of times during this course.*

★ *Exercise 13: Derive the following identities:*

(a) $\gamma v = c\sqrt{\gamma^2 - 1}$*;*

(b) $c^2 d\gamma = \gamma^3 v\, dv$*;*

(c) $d(\gamma v) = \gamma^3 dv$.

10. Properties of the Lorentz transformations

In this section we explore a number of properties of the Lorentz transformations, namely:

(a) Relativity of simultaneity;
(b) Newtonian limit;
(c) Difference and differential versions;
(d) Lorentz invariance of squared interval;
(e) Relativistic speed limit and causality;
(f) Group properties.

10.1 Relativity of simultaneity

An immediate consequence of the Lorentz transformations is that the concept of "simultaneous events" is relative, i.e. dependent on the state of motion of the observer. Events which are deemed simultaneous in one reference frame will not necessary appear simultaneous in another reference frame. This concept is known as the "relativity of simultaneity".

We work one space and one time dimension, for simplicity. Consider two events, which have coordinates (x_1, t_1) and (x_2, t_2) in frame S, with $x_1 \neq x_2$. Transforming each of these events to the frame S' using the Lorentz transformation, we see that:

$$\text{(57)}\quad \begin{cases} (x_1', t_1') = \gamma\left(x_1 - vt_1, t_1 - \dfrac{vx_1}{c^2}\right) \\ (x_2', t_2') = \gamma\left(x_2 - vt_2, t_2 - \dfrac{vx_2}{c^2}\right) \end{cases}.$$

If we set $t_1 = t_2 \equiv t$, meaning that the two events are simultaneous in frame S, we see that:

(58) $t_2' - t_1' = \dfrac{\gamma v}{c^2}(x_1 - x_2)$.

Thus the events are not simultaneous in S', unless $x_1 = x_2$. Therefore simultaneity is a relative (i.e. frame dependent) concept rather than an absolute concept. Note that simultaneity becomes absolute if we formally take the speed of light to infinity or take $v \ll c$, in which case the right-hand side of (58) vanishes.

This concludes our *mathematical* derivation of the relativity of simultaneity. The same result can be derived using the following *intuitive* argument, known as "Einstein's train".

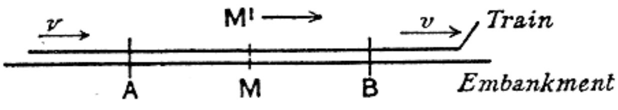


Figure 11: Einstein's train

Suppose we have a train which travels to the right, as shown in Figure 11. Frame S will be that of an observer standing on the embankment at point M; frame S' will be that of a passenger at point M'. Suppose now that the observer M on the embankment notes that lightning simultaneously strikes points A and B on the ground; the time of the strikes occurs when M and M' are at the same point. When we say that "the lightning strikes A and B are simultaneous with respect to the observer M on the embankment", we mean: the rays of light emitted at the places A and B, where the lightning occurs, meet each other at the mid-point M of the length AB of the embankment.

Question: Are the events "Lightning strikes A" and "Lightning strikes B", which are simultaneous events to the observer M on the embankment, also simultaneous events to the passenger at M'?

The passenger is moving towards the beam of light coming from B, whilst she is riding on ahead of the beam of light coming from A. Hence *the observer will see the beam of light emitted from B earlier than she will see that which is emitted from A.* The passenger using the railway carriage as their frame of reference will therefore conclude that the lightning flash B took place earlier than the lightning flash A. Since the two observers do not agree on the simultaneity or otherwise of the two lightning strikes, we conclude that the concept of simultaneity is relative rather than absolute.

10.2 Newtonian limit and the Galilean transformations

If we deduce the non-relativistic limit of the Lorentz transformations, by formally taking the speed of light to infinity, or by taking v<<c, we obtain the standard form of the Galilean transformations given in (34). Thus the new theory reduces to the old as a special case, in the domain of validity of the older theory.

10.3 Difference and differential versions

Since the Lorentz transformations are linear and homogeneous, coordinate differences $\Delta t = t_2 - t_1$ etc., as well as differentials, satisfy the same transformation equations as the coordinates themselves. For example:

$$(59)\quad \begin{cases} \Delta t' = \gamma\left(\Delta t - \dfrac{v\Delta x}{c^2}\right) \\ \Delta x' = \gamma(\Delta x - v\Delta t) \\ \Delta y' = \Delta y \\ \Delta z' = \Delta z \end{cases} \qquad (60)\quad \begin{cases} dt' = \gamma\left(dt - \dfrac{vdx}{c^2}\right) \\ dx' = \gamma(dx - vdt) \\ dy' = dy \\ dz' = dz \end{cases}$$

10.4 Lorentz invariance of squared interval

★ *Exercise 14: In the reference frame* S*, show that the equation of an expanding spherical burst of light is:*

(61a) $c^2t^2 - x^2 - y^2 - z^2 = 0$.

Now Lorentz transform to S'*, showing that the equation in the transformed frame is:*

(61b) $x'^2 + y'^2 + z'^2 = c^2t'^2$.

Therefore the left side of (61a) is invariant under Lorentz transformations (i.e. "Lorentz invariant"). Discuss this result in the context of the relativity of simultaneity.

★ *Exercise 15: Demonstrate the Lorentz invariance of the "squared interval"* $\mathbf{\Delta s^2}$*:*

(62) $\mathbf{\Delta s^2} \equiv c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2$.

This quantity can be positive, negative or zero; by the results of the previous exercise, the squared interval is clearly zero for events which can be connected by a light signal.

10.5 Relativistic speed limit and causality

As $v \to c^-$ (i.e. "v tends to c from below"), $\gamma \to \infty$. When $v > c$, γ becomes complex. Therefore the relative velocity of two inertial frames must be less than c, since (i) the γ factor appears in the Lorentz transformations; and (ii) a finite real coordinate in one inertial reference frame must correspond to a finite real coordinate in another inertial reference frame.

Indeed, we can show that the speed of particles, and, more generally, of all physical "signals", is limited by c, *if we insist on the Lorentz invariance of causality* (i.e. we demand that cause must precede effect in all inertial reference frames connected by a Lorentz transformation).

The proof of this statement is a proof by contradiction. Assume that an event $\mathsf{P} = (0,0)$ leads to an event $\mathsf{Q} = (x,t)$ with $x > ct$ (i.e. the influence of event P has travelled at super-luminal speed to event Q). Denote by η the time difference between these two events; in S, $\eta = t > 0$, which indicates that the cause P preceded the effect Q. Now, Lorentz transform each of these events:

$$\begin{aligned} &\mathsf{P} = (0,0) \xrightarrow{\text{Lorentz trans.}} \mathsf{P}' = (0,0) \\ (63)\quad &\mathsf{Q} = (x,t) \xrightarrow{\text{Lorentz trans.}} \mathsf{Q}' = \gamma\left(x - vt, t - \frac{vx}{c^2}\right) \end{aligned}$$

Here is the time difference η' in the frame S':

$$(64)\quad \eta' = \gamma\left(t - \frac{vx}{c^2}\right).$$

Since $x > ct$, let $x = ct + \Delta$, where $\Delta \in \Re^+$ (i.e. "Δ is a real positive number"). Thus:

$$(65)\quad \eta' = \gamma\left(t - \frac{v(ct + \Delta)}{c^2}\right) = \frac{\gamma}{c^2}\left(ct(c - v) - v\Delta\right).$$

Now, there exist inertial frames where $c - v$ can be made arbitrarily small. In such frames:

$$(66)\quad \eta' \to \frac{-v\Delta\gamma}{c^2}, \quad v \to c.$$

Thus, in such frames, $\eta' < 0$ and thus the order of cause and effect is reversed. This violates causality, and so the initial assumption of "Assume that an event $\mathsf{P} = (0,0)$ leads to an event $\mathsf{Q} = (x,t)$ with $x > ct$" must be incorrect. Therefore the speed of light forms a "relativistic speed limit" (cf. Figure 2). Corollary: Rigid bodies do not exist (cf. Exercise 5).

10.6 Group properties

The Lorentz transformations form a mathematical group. Specifically, a set G of operations $\{a, b, c, \cdots\}$, together with a binary operator that allows us to form the "product" of any two elements, forms a group if the following hold:

(a) If a and b are in G, then ab is also in G;

(b) There exists a "unit element" I in G such that, for any a in G, $Ia = aI = a$;

(c) If a is in G then there exists a^{-1} in G such that $aa^{-1} = a^{-1}a = I$;

(d) The product is associative, i.e. $(ab)c = a(bc)$.

★ *Exercise 16: Show that Lorentz transform-ations form a group (the "Lorentz group").*

★ *Exercise 17: Graph the Lorentz factor* γ *versus* v *for* $-c < v < c$.

★ *Exercise 18: If two events occur at the same point in some inertial reference frame* S*, prove that their temporal order is the same for all reference frames, and that the least time separation between them is assigned in* S.

★ *Exercise 19: If two events occur at the same time in some inertial frame* S*, prove that there is no limit on the time separation assigned to these events in other frames. Also, prove that their space separation varies from infinity to a minimum which is measured in* S.

★ *Exercise 20: In* S' *a straight rod parallel to the* x' *axis moves in the* y' *direction with constant velocity* u*. Show that in* S *the rod is inclined to the* x *axis at an angle:*

$$(67)\quad \theta = -\tan^{-1}(\gamma uv / c^2).$$

CHAPTER 4: RELATIVISTIC KINEMATICS[13]

11. On world pictures and world maps

At this stage in our discussions, we need to make a clear distinction between the concepts of "world picture" and "world map".

(i) What a given observer actually *sees* at a given instant of time will be denoted as her "world picture". This notion of a "world picture" is actually quite a complicated construct, for this picture is made up of *a composite of events which occurred progressively earlier as one looks at events which occurred further and further away*. As an example of this, when you look at the moon against a backdrop of stars, you see the moon as it was a little over a second ago[14], but you see most of the stars as they were several tens of thousands of years ago[15].

(ii) For our purposes, a more useful concept is that of the "world map", which is a map of events in an observer's instantaneous space $t = t_0$.

When we speak of such things as "the length of a measuring rod is frame S" *etc.*, we will always be referring to the "world map", unless otherwise stated.

12. Length contraction

Consider two inertial frames S and S' which are in standard configuration. Suppose we have a measuring rod of length $\Delta x'$ which is placed at rest along the x' axis of S'. To measure the length of the rod in any inertial reference frame in which the said rod is moving, we need to measure the ends points of the rod *simultaneously*; since simultaneity is a relative concept, these length measurements may be frame dependent. We do not need to take this precaution (of measuring both ends of the rod simultaneously) if we are measuring the length of the rod in the *rest* frame of the rod.

Consider, therefore, two events which occur simultaneously at the ends of the rod in S. Take the second equation in (59), set $\Delta t = 0$ (because the ends of the rod are measured at the same instant), to conclude that:

(74) $\Delta x' = \gamma \Delta x$.

If we replace Δx by the symbol L, and $\Delta x'$ by the symbol L_0, we obtain:

(74) $L = \frac{L_0}{\gamma} = L_0\sqrt{1 - \frac{v^2}{c^2}}$.

We see, therefore, that *the length L of a body in the direction of its motion with uniform velocity* v *is reduced by a factor* $(1 - v^2/c^2)^{1/2}$. This is the phenomenon of "length contraction".

Evidently, the greatest length ascribed to a uniformly moving body is in its *rest frame*, i.e. the frame of reference in which the said body is stationary. This length, which we have denoted L_0, is known as the "rest length" or "proper length" of the body.

Note, for future reference, that a "proper" measure of a quantity is that which is taken in the instantaneous rest frame.

13. Time Dilation

Consider frames S and S' in the usual standard configuration. Suppose we have a clock rigidly attached to the origin $(x', y', z') = (0,0,0)$ of the S' axis. Suppose further that, in S', two consecutive ticks of the clock are associated with the events:

$$\mathsf{P}' = (t' = 0, x' = 0, y' = 0, z' = 0)$$

and:

$$\mathsf{Q}' = (t' = T_0, x' = 0, y' = 0, z' = 0).$$

Thus, in S', one clock tick lasts for T_0 seconds (this length of a clock tick is equal to the difference between the time coordinate of Q and the time coordinate of P).

Using the inverse Lorentz transformations (57), we see that the events P' and Q' associated with the two consecutive ticks of the clock transform according to:

(68)

$$\mathsf{P}' = (0,0,0,0) \xrightarrow{\text{Inverse Lorentz trans.}} \mathsf{P} = (0,0,0,0)$$

$$\mathsf{Q}' = (T_0,0,0,0) \xrightarrow{\text{Inverse Lorentz trans.}} \mathsf{Q} = (\gamma T_0, \gamma v T_0, 0, 0)$$

---

[13] This chapter of the lecture course is closely based on W. Rindler, *Introduction to Special Relativity*, Oxford University Press, Oxford (1991), chapter 2, together with A. P. French, *Special Relativity*, Chapman and Hall, London (1968), chapter 4.

[14] The average distance from the earth to the moon is 384,403 km, so light reflected from the moon takes a little over a second to reach us.

[15] See http://www.atlasoftheuniverse.com/galaxy.html, which is very useful for getting a feel for the various length scales associated with the universe.

Thus, relative to the observer in S, one tick of the moving clock lasts for γT_0 seconds.

We conclude that a clock moving uniformly with velocity v through an inertial frame S *runs slow* by a factor of $\gamma = 1/(1 - v^2/c^2)^{1/2}$ relative to a stationary clock in S. (The fastest rate is evidently ascribed to a clock in its rest frame; this is called the "proper rate", and the time as measured in the rest frame of a particle is known as its "proper time").

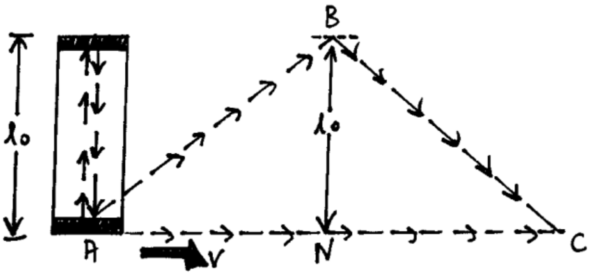


Figure 12: Use of the "Einstein clock" to determine the formula for time dilation

This concludes our *mathematical* derivation of time dilation. The same result can be derived using the following intuitive argument, known as the "Einstein clock".

Consider the "Einstein clock" which is shown in Figure 12. This clock consists of a pair of parallel plane mirrors, whose reflective surfaces face one another at a distance l_0. One "tick" of this clock is given by the time taken for a light pulse to go from the bottom mirror to the top mirror and back again.

Now introduce the frames S and S' which are in the standard configuration. When the clock is attached to the moving frame S', it is evident that one tick $\Delta t'$ of the clock lasts:

$$(69) \quad \Delta t' = 2l_0/c$$

seconds relative to a co-moving observer attached to S' (this is the time light takes to travel a distance of $2l_0$).

Relative to an observer in S, however, light travels the path ABC during one tick of the clock attached to S' (see Figure 12). Suppose that one tick of the clock lasts Δt relative to S. Then the distance AN is equal to the distance covered at speed v in $\Delta t/2$ seconds, i.e. $AN = v\Delta t/2$. Therefore, by Pythagoras' theorem,

$$(70) \quad AB = \sqrt{AN^2 + NB^2} = \sqrt{\left(\frac{v\Delta t}{2}\right)^2 + l_0^2}\ .$$

The total distance covered by the light in one tick of the moving clock is, in S, equal to 2AB. Therefore, in S, the time Δt elapsed during one tick of the moving clock is:

$$(71) \quad \Delta t = \frac{2AB}{c} = \frac{2}{c}\sqrt{\left(\frac{v\Delta t}{2}\right)^2 + l_0^2}\ .$$

Therefore, dividing (69) and (71), we obtain:

$$(72) \quad \begin{aligned} \frac{\Delta t}{\Delta t'} &= \sqrt{\left(\frac{v\Delta t}{2l_0}\right)^2 + 1} \\ &= \sqrt{\left(\frac{v\Delta t}{c} \times \left\{\frac{c}{2l_0}\right\}\right)^2 + 1} \end{aligned}\ .$$

Now, by equation (69), the quantity in braces is $1/\Delta t'$. If we substitute (69) into (72) and then re-arrange, we obtain the time-dilation formula once more:

$$(73) \quad \Delta t = \frac{\Delta t'}{\sqrt{1 - \frac{v^2}{c^2}}}\ .$$

★ *Exercise 21: Discuss, without the use of any equations, why the Einstein clock, together with Einstein's axioms, imply the existence of time dilation.*

Since the phenomenon of time dilation is so contrary to "common sense", we close this section with three brief examples of the experimental vindication of time dilation[16].

(i) In October, 1971, the reality of the time dilation phenomenon was experimentally demonstrated by Hafele and Keating, using extremely accurate atomic clocks which were flown around the world in commercial airliners. (Note that this experiment is complicated by the fact that both Special Relativity and General Relativity need to be taken into account in this experiment).[17]

---

[16] For further examples of time dilation experiments, click on the "time dilation" link available at http://hyperphysics.phy-astr.gsu.edu/hbase/relativ/relcon.html#relcon. See also http://www.mpi-hd.mpg.de/ato/Rel/ and http://math.ucr.edu/home/baez/physics/Relativity/SR/experiments.html#4. tests of time dilation.

[17] J.C. Hafele and R. E. Keating, Science **177**, 166 (1972).

(ii) Subatomic particles called muons are created by cosmic ray interaction with the upper atmosphere. At rest, they disintegrate in about 2×10^{-6} seconds and should not have time to reach the Earth's surface. Because they travel at close to the speed of light, however, *time dilation extends their life span* as seen from Earth so they can be observed reaching the surface before they disintegrate.

(iii) The Global Positioning System (GPS) utilises the corrections of both Special Relativity (such the corrections due to time dilation) *and* General Relativity (due to the curvature of space-time in the vicinity of GPS satellites; this curvature is quantified via the Schwarzschild solution to the Einstein field equations of General Relativity, which is also used to describe the curvature of spacetime in the vicinity of an uncharged non-rotating black hole). Even though GPS satellites do *not* move at "relativistic" speeds relative to the Earth, *the extremely high accuracy required of GPS necessitates the corrections of both Special and General Relativity* [a similar comment applies to point (i) above!]. For more information, see e.g. the articles by Neil Ashby at the following two web links:

http://relativity.livingreviews.org/Articles/lrr-2003-1/index.html

http://www.ipgp.jussieu.fr/~tarantola/Files/Professional/GPS/Neil_Ashby_Relativity_GPS.pdf

14. Twin "Paradox"

The Twin "Paradox" is probably the most famous "paradox" of Special Relativity.

Suppose a pair of identical twins hold synchronised clocks A and B at rest at some point P of an inertial frame S. The first twin stays at point P. The other twin hops in a rocket, is briefly accelerated to velocity v, and subsequently travels to some distant point Q in S. The rocket then decelerates and returns with velocity -v to P. According to time dilation, the B-twin will be younger than the A-twin, for each twin ages at the same rate as the clock they are carrying.

The apparent paradox is as follows: can we not invoke the principle of relativity and say that B remained stationary, while A went on a round trip? We would then conclude that A should be younger when the twins meet once again, which is in contradiction with our earlier statement.

The apparent paradox is resolved by pointing out that *no such symmetry exists*. Twin A remained at rest in a single inertial frame during the entire duration of the experiment, while B was accelerated from her rest frame at P, at Q, and once again at P. The "symmetry" argument of the previous paragraph is therefore incorrect, and the apparent paradox is resolved[18].

We will undertake a fuller discussion of the twin "paradox", together with some other apparent paradoxes, during a tutorial.

15. Velocity Transformation

In the past few sections, we have spoken a lot about the behaviour of *events* under the Lorentz transformations, together with various associated implications. In this section, we look at how *velocities* transform under the Lorentz transformations.

To this end, consider the usual inertial frames S and S' in standard configuration. In S at some instant of time, suppose a particle has the velocity vector $\mathbf{u}$. What is the velocity in frame S'?

Define:

$$(74) \quad \begin{cases} \mathbf{u} = (u_1, u_2, u_3) = \left(\dfrac{dx}{dt}, \dfrac{dy}{dt}, \dfrac{dz}{dt} \right) \\ \mathbf{u}' = (u_1', u_2', u_3') = \left(\dfrac{dx'}{dt'}, \dfrac{dy'}{dt'}, \dfrac{dz'}{dt'} \right) \end{cases}$$

Now, substitute from equations (60) into the second equation of (74):

$$(75) \quad \mathbf{u}' = \left(\frac{\gamma(dx - vdt)}{\gamma\left(dt - \dfrac{vdx}{c^2}\right)}, \frac{dy}{\gamma\left(dt - \dfrac{vdx}{c^2}\right)}, \frac{dz}{\gamma\left(dt - \dfrac{vdx}{c^2}\right)} \right).$$

Divide each numerator and denominator by dt, to give:

---

[18] More detail on the twin paradox is available online at http://www.phys.unsw.edu.au/~jw/twin.html. I also recommend B. E. Schutz, *A First Course in General Relativity*, Cambridge University Press, Cambridge (1985), pp. 28-30.

(76)
$$\begin{aligned}\mathbf{u}' &= \left(\frac{\frac{dx}{dt} - v}{1 - \frac{v}{c^2}\frac{dx}{dt}}, \frac{\frac{dy}{dt}}{\gamma\left(1 - \frac{v}{c^2}\frac{dx}{dt}\right)}, \frac{\frac{dz}{dt}}{\gamma\left(1 - \frac{v}{c^2}\frac{dx}{dt}\right)} \right) \\ &= \left(\frac{u_1 - v}{1 - \frac{v}{c^2}u_1}, \frac{u_2}{\gamma\left(1 - \frac{v}{c^2}u_1\right)}, \frac{u_3}{\gamma\left(1 - \frac{v}{c^2}u_1\right)} \right)\end{aligned}.$$

Therefore we arrive at the following velocity transformation formulae for Special Relativity:

(77)
$$\begin{cases} u_1' = \dfrac{u_1 - v}{1 - u_1 v/c^2} \\ u_2' = \dfrac{u_2}{\gamma\left(1 - u_1 v/c^2\right)} \\ u_3' = \dfrac{u_3}{\gamma\left(1 - u_1 v/c^2\right)} \end{cases}$$

We can use a "v reversal" (see Exercise 12) to yield the inverse velocity transformations:

(78)
$$\begin{cases} u_1 = \dfrac{u_1' + v}{1 + u_1' v/c^2} \\ u_2 = \dfrac{u_2'}{\gamma\left(1 + u_1' v/c^2\right)} \\ u_3 = \dfrac{u_3'}{\gamma\left(1 + u_1' v/c^2\right)} \end{cases}$$

Equations (78) are the *velocity addition formulae of Special Relativity*. Of particular importance is the first member of (78), which gives the result of adding two collinear velocities u_1' and v.

Note also that equations (77) and (78) reduce to their Galilean counterparts if the speed of light is taken to infinity.

We close this section by deriving a number of relativistic identities which shall be of use to us later in this course.

Denote by:

(79a) $u \equiv \sqrt{u_1^2 + u_2^2 + u_3^2}$ and (79b) $u' \equiv \sqrt{u_1'^2 + u_2'^2 + u_3'^2}$

the magnitudes of corresponding velocities in S and S'. Now, we learned in section 10.4 of the Lorentz invariance of the squared interval (see formula (62)). As a special case of this result, the *infinitesimal* interval is invariant, and therefore:

(80)
$$\begin{aligned} &c^2 dt'^2 - dx'^2 - dy'^2 - dz'^2 \\ &\quad = c^2 dt^2 - dx^2 - dy^2 - dz^2. \end{aligned}$$

Factor out dt'^2 from the left side of this equation, and factor out dt^2 from the right side, to give:

(81)
$$\begin{aligned} &dt'^2\left(c^2 - \left(\frac{dx'}{dt'}\right)^2 - \left(\frac{dy'}{dt'}\right)^2 - \left(\frac{dz'}{dt'}\right)^2\right) \\ &\quad = dt^2\left(c^2 - \left(\frac{dx}{dt}\right)^2 - \left(\frac{dy}{dt}\right)^2 - \left(\frac{dz}{dt}\right)^2\right) \end{aligned}$$

Now use equations (79) to write this in the form:

(82) $dt^2\left(c^2 - u^2\right) = dt'^2\left(c^2 - u'^2\right)$.

Put this equation to one side for the moment.

Take the first member of equations (60), then square it:

(83) $dt'^2 = \gamma^2(v)\left(dt - \dfrac{vdx}{c^2}\right)^2$

and factor out dt^2 from the right side:

(84)
$$\begin{aligned} dt'^2 &= \gamma^2(v) dt^2\left(1 - \frac{v}{c^2}\frac{dx}{dt}\right)^2 \\ &= \gamma^2(v) dt^2\left(1 - \frac{u_1 v}{c^2}\right)^2 \end{aligned}.$$

This may now be substituted into (82), to give:

(85) $dt^2\left(c^2 - u^2\right) = \gamma^2(v) dt^2\left(1 - \dfrac{u_1 v}{c^2}\right)^2\left(c^2 - u'^2\right)$,

Cancel the dt^2 from both sides, isolate $c^2 - u'^2$ on the left-hand side, write out the γ factor explicitly as $(1 - v^2/c^2)^{-1/2}$, and then multiply the right side by c^4/c^4, to give:

(86) $c^2 - u'^2 = \dfrac{c^2\left(c^2 - u^2\right)\left(c^2 - v^2\right)}{\left(c^2 - u_1 v\right)^2}$.

★ *Exercise 22: Use formula (86) to show that, if* $u' < c$ *and* $v < c$*, then* $u < c$*. Interpret this result.*

★ *Exercise 23: Using (86) as a starting point, prove the equations for the transformation of the gamma factors between inertial frames:*

$$\text{(87a)}\quad \frac{\gamma(u')}{\gamma(u)} = \gamma(v)\left(1 - \frac{u_1 v}{c^2}\right) \qquad \text{(87b)}\quad \frac{\gamma(u)}{\gamma(u')} = \gamma(v)\left(1 + \frac{u_1' v}{c^2}\right).$$

CHAPTER 5: RELATIVISTIC OPTICS[19]

16. The drag effect

Consider our usual pair of inertial reference frames S and S', in the standard configuration. Suppose that there is a large fish-tank filled with water, which is attached to S', as shown in Figure 13 below.

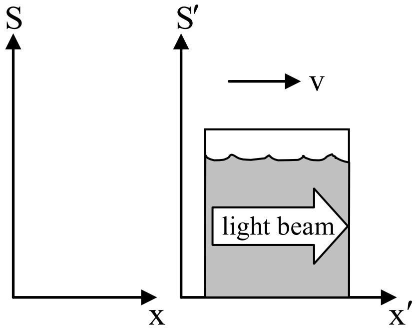


Figure 13: The drag effect

Now imagine that a beam of light travels in the x' direction within the fish-tank attached to S'. The velocity of the light in the frame S' is $u_1' = c/n$.

Make use of the first member of the Einstein velocity addition formulae (78), to write down the following expression for the velocity u_1 of the light in frame S:

$$\text{(88)}\quad \begin{aligned} u_1 &= \frac{u_1' + v}{1 + u_1' v/c^2} \\ &= \frac{c/n + v}{1 + v/(nc)} \\ &= \left(\frac{c}{n} + v\right)\left(1 + \frac{v}{nc}\right)^{-1} \\ &\approx \left(\frac{c}{n} + v\right)\left(1 - \frac{v}{nc}\right), \end{aligned}$$

where we have made use of the binomial approximation (16) in the last line of (88). Next, expand the last line of (88) and hence write down:

$$\text{(89)}\quad \begin{aligned} u_1 &\approx \frac{c}{n} + v\left(1 - \frac{1}{n^2}\right) - \frac{v^2}{nc} \\ &\approx \frac{c}{n} + v\left(1 - \frac{1}{n^2}\right) \end{aligned}$$

Rewrite this as:

$$\text{(90)}\quad u_1 \approx \frac{c}{n} + v\kappa, \qquad \kappa \equiv 1 - \frac{1}{n^2}.$$

The so-called "drag coefficient" κ is a number between zero and one which indicates what fraction of its own velocity the moving medium imparts to the light wave which is travelling within it.

17. The relativistic Doppler effect, Hubble's Law and Quasars

If you stand by the side of the road and a car happens to drive past while it is blaring its horn, you hear a slightly higher pitch while the car is coming towards you, and a slightly lower pitch while the car is receding. This is an example of the Doppler effect, a detailed treatment of which should have been given in at least one of your earlier courses. The Doppler effect also exists in Special Relativity, but the Doppler formulae need to be modified to take into account the effects of time dilation.

Consider the situation shown in Figure 14. Here, we consider a light source P which is travelling through some inertial frame S. At a given instant of time, the velocity of the light source P is given by $\mathbf{u}$, as marked in the Figure. The radial component of this velocity is denoted by u_r relative to an observer who is fixed to the origin O of frame S.

---

[19] The theoretical part of this chapter is closely based on W. Rindler, *Introduction to Special Relativity*, Oxford University Press, Oxford (1991), chapter 3. The experiments and their descriptions are taken from (i) A. P. French, *Special Relativity*, Chapman and Hall, London (1968), chapter 5; (ii) The Sloan Digital Sky Survey, http://www.sdss.org/news/releases/20000413.qso.img2.html.

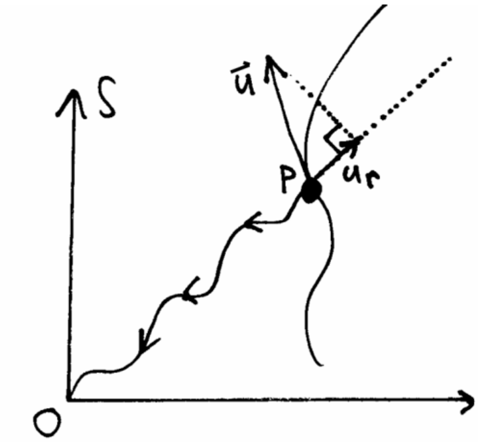


Figure 14: Diagram for derivation of the relativistic Doppler formula

Suppose the light source P emits pulses of light at intervals of dt_0 seconds as judged in its rest frame. By time dilation, we conclude that the time between light pulses will be $\gamma(u)dt_0$ as judged in S. In that time interval, the observer in S will deem that the radial distance of the source has increased by an amount equal to the said time interval multiplied by the radial velocity; therefore the radial distance of the source will have increased by $\gamma(u)u_r dt_0$. Therefore, according to the observer at the origin O of the frame S, the time interval between pulses is:

$$(91)\quad dt = dt_0\gamma(u) + \frac{\gamma(u)u_r dt_0}{c}\,,$$

where the first term on the right side accounts for time dilation and the second term accounts for the fact that the radial motion of the source adds a time lag to the pulses. Divide both sides of (91) by dt_0. Next, note that $dt/dt_0 = \nu_0/\nu$, where ν_0 is the frequency of the light pulses in the rest frame of the source, and ν is the frequency of the source perceived by the observer at the origin O of frame S. Therefore equation (91) becomes:

$$(92)\quad \frac{\nu_0}{\nu} = \left(\gamma(u) + \frac{\gamma(u)u_r}{c}\right) = \frac{1+u_r/c}{\sqrt{1-u^2/c^2}}\,.$$

Note that when the motion of the source is purely radial, $u = u_r$ and so (92) becomes:

$$\begin{aligned} \frac{\nu_0}{\nu} &\rightarrow \frac{1+u_r/c}{\sqrt{1-u_r^2/c^2}} \\ (93)\qquad &= \frac{1+u_r/c}{\sqrt{(1-u_r/c)(1+u_r/c)}} \\ &= \sqrt{\frac{1+u_r/c}{1-u_r/c}}\,. \end{aligned}$$

A famous example of the Doppler effect for light is the so-called "red-shift" of distant galactic spectra. The spectrum of most galaxies is pretty close to being a continuous smear. However, even in such smeared-out spectra, one can distinguish certain dark lines due to the *absorption* of light created by the galaxy as it passes through cooler gases before leaving the galaxy. Two such lines, the H and K lines of ionised calcium, can be seen in the three galactic spectra of Figure 15, where the horizontal axis increases with wavelength. We can see that the spectra are all *shifted towards the red (i.e. towards longer wavelengths)*. Using the Doppler formula, these so-called "red-shifts" can be converted to recession velocities, which are given in Figure 15.

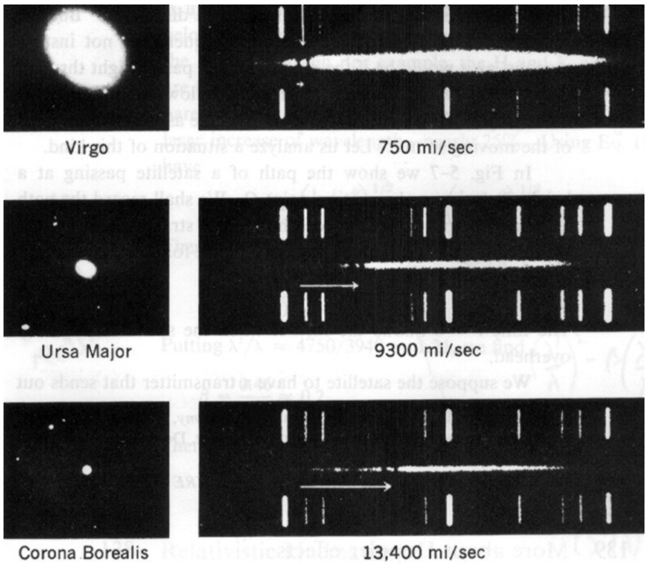


Figure 15: Spectra of three different galaxies, all of which exhibit a significant red-shift from which a recession velocity may be calculated

Hubble noted the linear relation between recession velocity and distance for remote galaxies; this relation is now known as *Hubble's Law*, and is illustrated in Figure 16. Note that the three spectra in Figure 15 yield the first three data points of Figure 16.

One might rightly object that the recession velocities in Figure 16 are hardly relativistic, so let me give an example where relativistic effects make an essential contribution to the Doppler effect. In Figure 17, we see part of the spectrum of one of the most distant astrophysical objects currently known, with a red-shift of z=7.085. (We define the red-shift z scale such that the observed wavelength is 1+z times the emitted wavelength). The large peak is the Lyman-alpha emission peak, and is a distinguishing feature of quasar spectra. This spectral feature would normally appear in the ultraviolet, with a wavelength of about 1200 Å. For this quasar, the Lyman-alpha peak appears in the infrared, representing a nearly tenfold increase in the wavelength due to the relativistic Doppler effect.

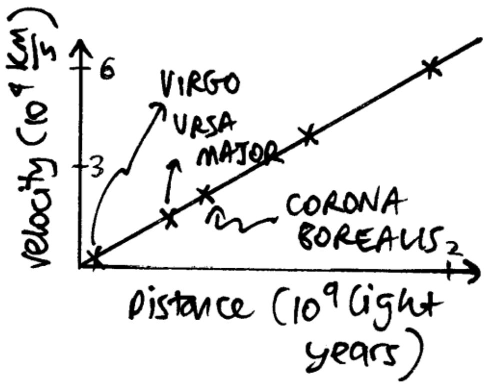


Figure 16: Hubble's Law for the linear relationship between recession velocity and distance

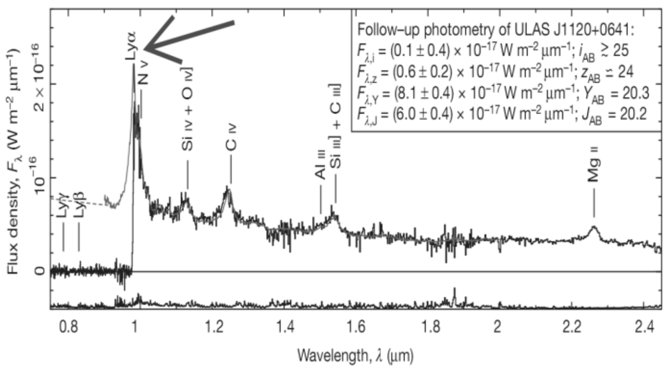


Figure 17: Spectrum of quasar ULASJ112001.481064124.3, with red-shift z=7.085. (Source: Mortlock et al., Nature 474, 617 (2011).[20] The large arrow marks the Lyman alpha line, which would have a wavelength of approximately 1200Å (in the ultraviolet band) for a source which is stationary with respect to the receiver. Due to the relativistic Doppler effect for a source (quasar) which is receding from the observer (on earth), the wavelength of the Lyman alpha line is now approximately 10,000 Å (in the infrared band).

★ *Exercise 24: As the first Russian Sputnik satellite whizzed around the earth it emitted a radio-frequency signal that was observed by many tracking stations. In the rest frame of Sputnik, this signal was of constant frequency. One observation of Sputnik's signal, as observed on Earth, is given in Figure 18a. The object of this exercise is to calculate a formula for the solid line which runs through the data points. To this end,(i) identify all relevant parameters in the problem; (ii) study the material in section 17 to see if there are any techniques which can be adapted to the present problem; (ii) obtain a formula for the measured frequency versus time, as a function of all relevant parameters in the problem. Hint: see Figure 18b.*

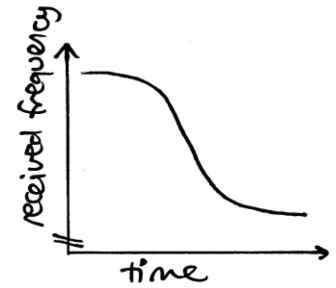


Figure 18a: Radio-frequency signal emitted by Sputnik I, relative to an observing station fixed to the earth[21].

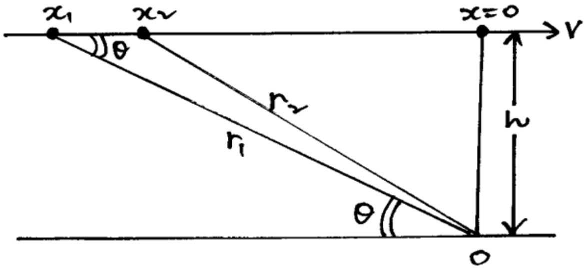


Figure 18b: Diagram to assist in calculation of Doppler effect on signals which are emitted at some angle with respect to the line of motion of the source

18. Aberration and visual appearance of moving objects

If you are driving through the rain, then the angle of the raindrops relative to the car will depend on the state of motion of the car. A similar effect occurs in relativistic optics: two inertial observers in different reference frames will not in general agree on the angle which an incoming ray of light makes with their relative line of motion.

---

[20] For an excellent introduction to quasar red-shifts, see: http://www.2dfquasar.org/restframe.html.

[21] Data taken by M.I.T. Lincoln Laboratory, 7/10/1957.

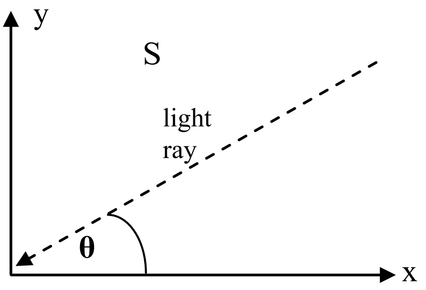


Figure 19: Diagram to assist in derivation of the aberration formulae of relativistic optics

Consider, then, the diagram shown in Figure 19. Here we see an incoming ray of light whose negative direction makes an angle of θ with the x axis of an inertial frame S. What value does this angle take in S', where the inertial frames S and S' are in the usual standard configuration with respect to each other?

The answer to this question is provided by the velocity transformation formulae discussed in section 15. Take the first member of the velocity transformation equations (77), setting $u_1 = -c\cos\theta$ and $u_1' = -c\cos\theta'$, to give:

$$u_1' = \frac{u_1 - v}{1 - \frac{u_1 v}{c^2}}$$

$$(94) \quad \Rightarrow -c\cos\theta' = \frac{-c\cos\theta - v}{1 - \frac{(-c\cos\theta)v}{c^2}}$$

$$\Rightarrow \cos\theta' = \frac{\cos\theta + \frac{v}{c}}{1 + \frac{v}{c}\cos\theta}.$$

Similarly, one can show that:

$$(95) \quad \sin\theta' = \frac{\sin\theta}{\gamma\left(1 + \frac{v}{c}\cos\theta\right)}.$$

If we now substitute from equations (94) and (95) into the trigonometric identity:

$$(96) \quad \tan\tfrac{1}{2}\theta' = \frac{\sin\theta'}{1 + \cos\theta'},$$

then the resulting expression can be manipulated into the following form:

$$(97) \quad \tan\tfrac{1}{2}\theta' = \sqrt{\frac{c - v}{c + v}}\tan\tfrac{1}{2}\theta.$$

★ *Exercise 25: Derive equations (95) to (97).*

Our aberration formula (97) implies the existence of *distortions in the visual appearance of extended moving objects*. From the viewpoint of the observer, this is because light from different parts of the moving object has taken different times to reach her eye at a given instant, and thus was emitted at different past times.

A nice example of the use of (97) is given by Figure 20. This shows one frame in a video computed by C.M. Savage and A.C. Searle; it shows the distorted appearance of the shadow of a tram moving from left to right at 90% of the speed of light.

★ *Exercise 26: A particle moves uniformly in a frame* S *with velocity* $\mathbf{u}$ *making an angle* α *with the positive* x *axis. If* α' *is the corresponding angle in the usual second frame* S'*, prove the "particle aberration formula":*

$$(98) \quad \tan\alpha' = \frac{\sin\alpha}{\gamma(v)(\cos\alpha - v/u)}.$$

★ *Exercise 27: Show that the ratio of solid angles subtended in* S *and* S' *by a thin pencil of light-rays converging on the coincident origins of these frames, its negative direction making angles* α, α' *with the respective* x *axes, is given by:*

$$(99) \quad \frac{d\Omega}{d\Omega'} = \left(\frac{d\alpha}{d\alpha'}\right)^2 = \gamma^2(v)\left(1 + \frac{v}{c}\cos\alpha\right)^2.$$

Figure 20: Distorted appearance of the shadow of a tram moving from left to right at 90% of the speed of light (image created by C. Savage and A. Searle)[22]

---

[22] For more detail, see animations and paper by C. Savage and A. Searle, *The Relativistic Raytracer*, at http://www.anu.edu.au/Physics/Searle/paper2.html. See also realtimerelativity.org.

CHAPTER 6: SPACETIME[23]

The remainder of the course will draw heavily on "Tensors for Special Relativity", from W. Rindler, *Introduction to Special Relativity*, Oxford University Press, Oxford (1991). Before proceeding further, we discuss salient material from this appendix.

19. Spacetime and four-tensors

We saw, in exercise 15, that the Lorentz transformations leave invariant the so-called "squared interval" $\mathbf{\Delta s^2}$, which was defined in equation (62) to be:

$$(62)\quad \mathbf{\Delta s^2} \equiv c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 .$$

Passing over to infinitesimals, we have:

$$(100)\quad \mathbf{ds}^2 \equiv c^2 dt^2 - dx^2 - dy^2 - dz^2 .$$

Now, we know from Pythagoras' theorem that the infinitesimal squared length $dl^2 = dx^2 + dy^2 + dz^2$ in three-dimensional space is invariant under rotations and translations of orthogonal Cartesian axes (e.g. rotating or translating your coordinate system does not change the length of a short toothpick which you are holding in your hand). Similarly, equation (100) gives an invariant in the four-dimensional "space-time" of events (t,x,y,z). A crucial difference between the "Euclidean" metric of three-dimensional space and the "pseudo-Euclidean" metric of four-dimensional space is the presence of the negative signs in (100): space and time are *not* treated on exactly the same footing, and space-time accordingly possesses non-isotropic properties utterly unlike Euclidean space.

Recall Einstein's first axiom from section 7 of the course, which stated that "The laws of physics are identical in all inertial frames". We also saw that the Lorentz transformations may be viewed as a change of coordinate system from one inertial reference frame to another inertial reference frame. We are therefore drawn to consider a tensor description of physical laws, since tensor equations have the property of being true or false independently of the coordinate system. Thus tensor analysis is the natural mathematical framework in which to incorporate Einstein's first axiom.

More precisely, we will consider tensorial objects defined on the four-space of events, which are coordinatized by:

$$(101)\quad x^0 = ct,\quad x^1 = x,\quad x^2 = y,\quad x^3 = z$$

and which behave as tensors under the Lorentz transformations. Note that the time variable is ct, since we want all coordinates to have the same dimensions (i.e. length).

We use Greek indices $\mu, \nu, ...$ for the range 0,1,2,3 while Latin indices $i, j, ...$ can take the range 1,2,3. Four-tensors will be denoted by capital letters (A, B^μ, C^μ_ν etc.) and three-tensors will be denoted by lowercase letters.

Equation (100) can be written as the tensor equation:

$$(102)\quad \mathbf{ds}^2 = g_{\mu\nu} dx^\mu dx^\nu ,$$

where the metric tensor is:

$$(103)\quad g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} .$$

★ *Exercise 28: Show that the metric tensor* $g_{\mu\nu}$ *is numerically equal to its conjugate* $g^{\mu\nu}$*, i.e.:*

$$(104)\quad g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} .$$

The scalar product of two four-vectors is:

$$(105)\quad \begin{aligned} \mathbf{A} \bullet \mathbf{B} &= g_{\mu\nu} A^\mu B^\nu \\ &= A^0B^0 - A^1B^1 - A^2B^2 - A^3B^3 \end{aligned} .$$

Two four-vectors **A** and **B** are orthogonal if:

$$(106)\quad \mathbf{A} \bullet \mathbf{B} = 0 .$$

As a special case of (105), the square of a four-vector is:

$$(107)\quad \mathbf{A}^2 = (A^0)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2 .$$

---

[23] W. Rindler, *Introduction to Special Relativity*, Oxford University Press, Oxford (1991), chapter 4.

The space of events x^μ, equipped with the metric tensor $g_{\mu\nu}$, is a metric space[24] (i.e. a space with a metric) known as "Minkowski space" or "flat space-time". This space is the stage upon which the tensorial formulation of Special Relativity is played out. The tensorial formulation of a physical law ensures its invariance under Lorentz transformations, in keeping with Einstein's first axiom.

20. World-lines and light cones

In the course of its time evolution the path of a point particle will trace out a line in space-time; this line is known as the "world-line" of the particle. This world-line $x^\mu(\tau)$ may be parameterized by the proper time τ of the particle. Such a world-line through space time is shown in Figure 21, for the case of two spatial dimensions and one time dimension (this results in a three-dimensional space-time which is easier to visualize that the full four-dimensional space-time!).

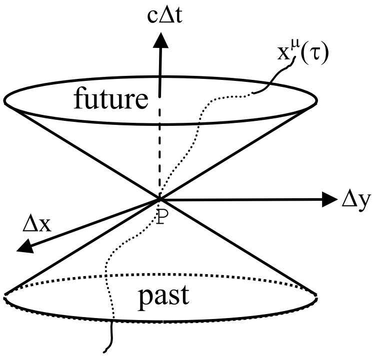


Figure 21: The light cone

In addition to the concept of a world-line, Figure 21 introduces the concept of the "light cone", which has the vertical axis as time and the horizontal directions as two spatial coordinates. Suppose that the event `P` lies at the origin of space-time coordinates for a given inertial observer. Suppose furthermore that a spherical burst of light is emitted from the event `P`; space-time points which lie on this spherical burst of expanding light will obey:

$$(108)\quad c^2\Delta t^2 = \Delta x^2 + \Delta y^2 + \Delta z^2 .$$

With the help of (62), this can be expressed in the form:

$$(109)\quad \mathbf{\Delta s^2} \equiv c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 = 0,$$

an equation which we have already seen to be invariant under Lorentz transformation (see exercise 15). For $\Delta t>0$, equation (109) represents an expanding sphere of light centered on the event `P`; for $\Delta t<0$, equation (109) represents a contracting sphere of light which collapses to `P` at time $\Delta t=0$. Since light represents the ultimate speed limit, the maximal set of events that can be causally linked to event `P` are precisely those that lie within the spherical burst of light defined by equation (109): more precisely, events which satisfy $\mathbf{\Delta s^2} > 0$ and $\Delta t>0$ lie inside the expanding sphere of light and constitute the maximal set of space-time events ("absolute future") which can be causally influenced by event `P`, while events which satisfy $\mathbf{\Delta s^2} > 0$ and $\Delta t<0$ lie inside the contracting sphere of light constitute the maximal set of space-time events ("absolute past") which can causally influence the event `P`. Events which lie *outside* the expanding shell of light (109) can have no causal relation with event `P`. Note that the notions of "absolute future" and "absolute past", for space-time events lying inside the expanding sphere of light, are invariant under Lorentz transformations (cf. section 10.5).

In this context, note the following terminology: Space-time displacement vectors satisfying $\mathbf{\Delta s^2} > 0$ are called "time-like"; space-time displacement vectors satisfying $\mathbf{\Delta s^2} < 0$ are called "space-like"; space-time displacement vectors satisfying $\mathbf{\Delta s^2} = 0$ are called "light-like" or "null". The set of displacements satisfying $\mathbf{\Delta s^2} = 0$ form a three-dimensional hyper-surface through space-time which is known as the "light cone". For the case of two spatial dimensions, this light cone is illustrated in Figure 21. Evidently, by the arguments regarding causality which have just been outlined, and since a particle always moves at speeds less than that of light, its world-line not only locally but also globally lies within every light cone that has its vertex on the said particle.

21. Manipulation of four-tensors

As we saw in our earlier studies on tensor analysis, if we simply say that an "object is a tensor" then it is understood that the object behaves as a tensor under all non-singular differentiable coordinate transformations. An object which behaves as a tensor only under a certain subgroup of non-singular differentiable coordinate transformations, like the Lorentz transformations, may be called a "qualified tensor", and its name should be qualified by an adjective which recalls the subgroup in question, as in

---

[24] Metric spaces may be fruitfully studied in their own right. If interested in following up this point, I recommend W. A. Sutherland, *Introduction to Metric and Topological Spaces*, Oxford Science Publications, Oxford (1975).

"Lorentz tensor" = "four tensor". It is these tensors which are used in Special Relativity. We will occasionally lapse from the proper terminology, and simply speak of "tensor" when we really should use the term "four-tensor".

Suppose that a tensor transformation has to be computed for some contravariant four-vector A^μ, this transformation being between two inertial reference frames which are in standard configuration with respect to each other. Any contravariant four-vector A^μ transforms according to the same scheme as the space-time coordinates x^μ themselves, namely:

$$(110)\begin{cases} x'^0 = \gamma\left(x^0 - \dfrac{vx^1}{c}\right) \\ x'^1 = \gamma\left(x^1 - \dfrac{vx^0}{c}\right) \\ x'^2 = x^2 \\ x'^3 = x^3 \end{cases} \qquad (111)$$

$$\begin{cases} x^0 = \gamma\left(x'^0 + \dfrac{vx'^1}{c}\right) \\ x^1 = \gamma\left(x'^1 + \dfrac{vx'^0}{c}\right) \\ x^2 = x'^2 \\ x^3 = x'^3 \end{cases}$$

Equation (110) results from using (101) to rewrite the forward Lorentz transformation (56); equation (111) results from applying the usual v-reversal transformation to (110) (cf. exercise 12).

As an example of the application of the above equations to a given contravariant four-vector A^μ, we write down formulae such as $A'^0 = \gamma(A^0 - vA^1/c)$, etc.

What about higher-order tensors? They transform according to the canonical transformation laws given in section A5 of Rindler's appendix, namely:

(112a) $A^{i'j'\ldots n'} = A^{ij\ldots n} p^{i'}_i p^{j'}_j \ldots p^{n'}_n$ (contravariant)

(112b) $A_{i'j'\ldots n'} = A_{ij\ldots n} p^i_{i'} p^j_{j'} \ldots p^n_{n'}$ (covariant)

(112c) $A^{i'\ldots k'}_{l'\ldots n'} = A^{i\ldots k}_{l\ldots n} p^{i'}_i \ldots p^{k'}_k p^l_{l'} \ldots p^n_{n'}$ (mixed)

(112d) $\dfrac{\partial x'^i}{\partial x^i} \equiv p^{i'}_i, \quad \dfrac{\partial x^i}{\partial x'^i} \equiv p^i_{i'}.$

We leave it as an exercise to derive the required transformation coefficients:

(113a) $p^{0'}_0 = p^{1'}_1 = \gamma, p^{0'}_1 = p^{1'}_0 = -\dfrac{\gamma v}{c}, p^{2'}_2 = p^{3'}_3 = 1$

(113b) $p^0_{0'} = p^1_{1'} = \gamma, p^0_{1'} = p^1_{0'} = \dfrac{\gamma v}{c}, p^2_{2'} = p^3_{3'} = 1$

(113c) All other "p" coefficients vanish.

★ *Exercise 29: Prove equations (113), and then derive the following transformation law for one of the components of the contravariant four-tensor* $T^{\mu\nu}$*:*

$$\begin{aligned} T^{1'2'} &= p^{1'}_\mu p^{2'}_\nu T^{\mu\nu} \\ &= p^{1'}_0 p^{2'}_2 T^{02} + p^{1'}_1 p^{2'}_2 T^{12} \\ &= \gamma\left(T^{12} - vT^{02}/c\right) \end{aligned} \tag{114}$$

22. Four-velocity and four-acceleration

Along time-like directions in spacetime, $\mathbf{\Delta s^2} > 0 \equiv ds^2$, and it is often convenient to work with the invariant quantity:

$$(115)\quad d\tau^2 = \frac{ds^2}{c^2} = dt^2 - \frac{dx^2 + dy^2 + dz^2}{c^2}$$

rather than with dt^2 itself. The symbol τ is chosen to be positive in the future-pointing direction, and is known as the "proper time". Now, at the very end of section 12 we stated that a "proper" measure of a quantity is that which is taken in the instantaneous rest frame of the particle under consideration. Hence we need to justify the use of the term "proper time" to refer to the τ defined above. Bearing this in mind, note the following three things: (i) In the instantaneous rest frame of a particle $dx = dy = dz = 0$ and so (115) shows that $d\tau$ is the time differential dt measured on an ideal clock attached to the moving particle, whether that particle is moving or not. (ii) If we write u for the speed of the particle, and divide (115) by dt^2, we obtain:

$$(116)\quad \begin{aligned} &\frac{d\tau^2}{dt^2} = 1 - \frac{\dfrac{dx^2}{dt^2} + \dfrac{dy^2}{dt^2} + \dfrac{dz^2}{dt^2}}{c^2} = 1 - \frac{u^2}{c^2} \\ &\Rightarrow \frac{dt}{d\tau} = \frac{1}{\sqrt{1 - \dfrac{u^2}{c^2}}} = \gamma(u) \end{aligned}$$

By the results of time dilation which we studied in section 13 (see especially equation (73)), equation (116) also bears out the assertion that τ is the proper time. (iii) The proper time is a very natural quantity with which to work and will appear in many relativistic formulae.

Consider now a particle with world-line $x^\mu(\tau)$ parameterized by the proper time elapsed for the moving particle. Space-time coordinates are the archetypal contravariant four-vectors, and so $x^\mu(\tau)$ is a four-vector. The differential of a four-vector is also a four-vector, and so $dx^\mu(\tau)$ is a four-vector, as is its derivative $dx^\mu(\tau)/d\tau$ and its second derivative $d^2x^\mu(\tau)/d\tau^2$. We are therefore led to the introduction of the *four-velocity* **U** and the *four-acceleration* **A**:

$$(117)\quad \mathbf{U} \equiv \frac{dx^\mu}{d\tau}, \quad \mathbf{A} \equiv \frac{d^2x^\mu}{d\tau^2} = \frac{d\mathbf{U}}{d\tau}.$$

What is the relation between the four-velocity and the more familiar three-velocity $\mathbf{u} = dx^i/dt$? Here it is:

$$(118)\quad \begin{aligned} \mathbf{U} &\equiv \frac{dx^\mu}{d\tau} = \frac{dt}{d\tau}\frac{dx^\mu}{dt} = \gamma(u)\frac{d}{dt}(ct,x,y,z) \\ &= \gamma(u)(c,\mathbf{u}) \end{aligned}.$$

The relation between the four-acceleration $\mathbf{A}$ and the three-acceleration $\mathbf{a}$ is worked out using similar logic:

$$(119)\quad \begin{aligned} \mathbf{A} &= \frac{d\mathbf{U}}{d\tau} = \gamma\frac{d\mathbf{U}}{dt} = \gamma\frac{d}{dt}\gamma(c,\mathbf{u}) = \gamma\frac{d}{dt}(\gamma c,\gamma\mathbf{u}) \\ &= \gamma(\frac{d\gamma}{dt}c, \frac{d\gamma}{dt}\mathbf{u} + \gamma\frac{d\mathbf{u}}{dt}) = \gamma(\frac{d\gamma}{dt}c, \frac{d\gamma}{dt}\mathbf{u} + \gamma\mathbf{a}) \end{aligned}$$

where $\gamma \equiv \gamma(u)$. In the instantaneous rest frame of the particle, this simplifies to:

$$(120)\quad \mathbf{A} = (0,\mathbf{a}).$$

★ *Exercise 30: Prove equation (120). Hint: The results of exercise 13 may be helpful!*

Thus **A** is equal to zero if and only if the magnitude of the three-acceleration in the rest frame vanishes. The four-velocity never vanishes.

★ *Exercise 31: (a) Show that the four-velocity never vanishes; (b) Show that the square of the four-velocity is always the same, namely:*

$$(121)\quad \mathbf{U}^2 = c^2.$$

Give two different proofs of this fact, one of which works in the instantaneous rest frame and one of which does not; (c) Show that the square of the four-acceleration is an invariant, namely:

$$(122)\quad \mathbf{A}^2 = -\alpha^2,$$

where α *is the magnitude of the three-acceleration in the rest frame. (d) Show that:*

$$(123)\quad \mathbf{U}\bullet\mathbf{A} = 0,$$

i.e. the four-acceleration is always orthogonal to the four-velocity. (e) Show, using (119) or otherwise, that:

$$(124)\quad \mathbf{A}^2 = \gamma^2\left(\left(\frac{d\gamma}{dt}\right)^2 c^2 - \left(\frac{d\gamma}{dt}\mathbf{u} + \gamma\mathbf{a}\right)^2\right).$$

(f) By making use of the results of exercise 13, show that:

$$(125)\quad \begin{aligned} \alpha^2 &= -\mathbf{A}^2 \\ &= \gamma^2\left(\left(\frac{d\gamma}{dt}\right)^2(u^2 - c^2) + 2\gamma\frac{d\gamma}{dt}u\frac{du}{dt} + \gamma^2 a^2\right) \end{aligned}$$

We close this section by looking into the absolute significance of the scalar product of two four-velocities, say **U** and **V**, which correspond to two uniformly moving particles. Since the dot product of four-vectors is an invariant, let us look at $\mathbf{U}\bullet\mathbf{V}$ in the rest frame of the second particle, in which the first particle has four-velocity $\mathbf{U} = \gamma(u)(c,\mathbf{u})$ and the second has four-velocity $\mathbf{V} = (c,\mathbf{0})$. Therefore:

$$(126)\quad \mathbf{U}\bullet\mathbf{V} = c^2\gamma(u),$$

i.e. $\mathbf{U}\bullet\mathbf{V}$ is c^2 multiplied by the Lorentz factor of the relative velocity of the corresponding particles.

★ *Exercise 32: All four-vectors in this problem are presumed to be real and non-zero. Let T, S, N, V respectively denote timelike, spacelike, null and general vectors. Prove: (i) any V orthogonal to a T or N (other than the N itself) is an S; (ii) the sum of two Ts, or of a T and N, which are isochronous (i.e. both pointing into the future or both pointing into the past) is a T isochronous with them.*

★ *Exercise 33: Prove the "zero component lemma" for four-vectors, which states that if a four-vector* V^μ *has a particular one of its four components zero in all inertial frames then the entire vector must be zero.*

CHAPTER 7: RELATIVISTIC PARTICLE MECHANICS 1[25]

23. Introduction

In the earliest sections of the lecture course, we saw some departures from Newtonian mechanics at relativistic velocities. We saw, furthermore, that Newtonian mechanics is inconsistent with some of the consequences of Einstein's axioms (e.g. Newtonian mechanics considers time to be an absolute quantity, and it allows particles to be accelerated to arbitrarily large speeds). While we made some headway in relativistic particle mechanics in the very first lecture, it is time to undertake a deeper formulation of the same using tensor analysis.

Accordingly, in chapters 7 and 8 of the lecture course we shall investigate some of the applications of the tensorial formulation of Special Relativity to the relativistic mechanics of *particles*. (We shall defer a discussion of relativistic *fields* such as the electromagnetic field until later in the course.)

24. Conservation of four-momentum

Our physics discussions so far have essentially been a detailed investigation of the consequences of Einstein's axioms of Special Relativity. This has taken us a long way, but we shall not be able to proceed further without the introduction of additional hypotheses. Such an additional hypothesis is that of *the conservation of four-momentum* which shall occupy us for the present section.

By "the conservation of four-momentum" we mean that *the sum of the four-momenta of all of the particles going into a collision is the same as the sum of the four-momenta of the particles leaving a collision*. Note that the said collision may be elastic or inelastic, and there may be more or fewer particles going in than coming out. Our *hypothesis* of the conservation of four-momentum is written in the form:

$$(127)\quad \sum{}^{*}\mathbf{P}_{(a)} = \mathbf{0},$$

where a=1,2,3... is *not* a tensor index but rather refers to the various particles both before and after the collision, with Σ^* being a sum that counts pre-collision terms positively and post-collision terms negatively. In order for us to accept hypothesis (127), it must: (a) reduce to the appropriate Newtonian results in the non-relativistic limit; (b) be consistent with the axioms of Special Relativity; (c) correctly predict the results of experiments undertaken using particles travelling at both relativistic and non-relativistic velocities.

We have hypothesized equation (127) for the conservation of four-momentum, so we had better define what four-momentum **P** actually is:

$$(128)\quad \begin{aligned} \mathbf{P} &\equiv m_0\mathbf{U} \ldots \text{definition of four-momentum} \\ &= m_0\gamma(u)(c,\mathbf{u}) \ldots \text{from equation (118)} \\ &= m(c,\mathbf{u}) \ldots \text{from equation (21)} \\ &= (mc,\mathbf{p}) \ldots \text{where } \mathbf{p} \equiv m\mathbf{u} \end{aligned}$$

where we have made use of section 3's distinction between the rest mass m_0 and the velocity-dependent mass m. Since **P** is the product of the four-tensor m_0 ("scalar" or "invariant") with the four-tensor **U**, we conclude that **P** is also a four-tensor/four-vector. Therefore (127) is a four-tensor equation; if it is valid in one inertial reference frame then in is valid in all inertial reference frames.

If we substitute (128) into our hypothesized law (127) for the conservation of four-momentum, and then separate the resulting vectorial expression into its temporal and spatial components, we end up with the following pair of conservation laws:

$$(129a)\quad \sum{}^{*}m = 0,$$

$$(129b)\quad \sum{}^{*}\mathbf{p} = \mathbf{0} \Rightarrow \sum{}^{*}m\mathbf{u} = \mathbf{0},$$

where we have left off the "(a)" subscripts for simplicity. Note that this pair of conservation laws is not independent, a fact which is a consequence of the "zero component lemma" examined in exercise 33; according to this lemma, the universal vanishing of one component of a four-vector - in this case $\Sigma^*\mathbf{P}$ - implies the vanishing of all of the other components.

In the non-relativistic limit, equations (129a) and (129b) are in agreement with Newtonian mechanics, for they assert the conservation of mass and momentum as two separate conservation laws (note that $m \to m_0$ in the non-relativistic limit).

For relativistic particles, equations (129) may *look* like their Newtonian counterparts in a formal sense, but we must remember that there is a profound

---

[25] W. Rindler, *Introduction to Special Relativity*, Oxford University Press, Oxford (1991), chapter 5.

conceptual and mathematical difference because of the velocity dependence of the mass. At the risk of over-emphasizing this point, it may be worth re-reading the discussions on the inertia of energy in section 2, to better appreciate this conceptual differences between (129a) and its Newtonian analogue.

We close this section by noting that, in view of Einstein's famous equation:

$$(8)\quad E = mc^2$$

for the equivalence of mass and energy, we may re-write equation (128) for the four-momentum as:

$$(130)\quad \mathbf{P} = (E/c, \mathbf{p}),$$

a form generally preferred by particle physicists.

25. Some four-momentum identities

Here, we derive three different four-momentum identities.

(a) Energy-momentum-mass relation. Since **P** is a four-tensor, its square must be the same in all reference frames (i.e. $\mathbf{P}^2$ is an invariant). Since (128) says that **P**=(mc,**p**), we can immediately write down:

$$(131)\quad \mathbf{P}^2 = m^2c^2 - p^2 = m_0^2c^2,$$

where the last step follows by evaluating $\mathbf{P}^2$ in the rest frame. Multiplying by c^2 and re-arranging, we obtain:

$$(132)\quad m^2c^4 = m_0^2c^4 + p^2c^2.$$

Since $E = mc^2$, the left side of (132) is the square of the energy and so we obtain the relativistic energy-momentum-mass relation:

$$(133)\quad E^2 = m_0^2c^4 + p^2c^2.$$

This is exactly the same as equation (19), which was derived using a less sophisticated approach.

(b) Some two-particle relationships. Suppose we have two particles with four-momenta $\mathbf{P}_1$ and $\mathbf{P}_2$. Then it can be shown that the following two-particle relationships hold:

$$(134)\quad \begin{aligned}\mathbf{P}_1 \bullet \mathbf{P}_2 &= c^2 m_{01} m_2 = c^2 m_1 m_{02} \\ &= c^2 m_{01} m_{02} \gamma(v)\end{aligned}$$

where m_{01} is the rest mass of the first particle, m_2 is the mass of the second particle in the rest frame of the first, etc., and v is their relative speed.

★ *Exercise 34: Prove each of the equations appearing in (134), by evaluating the invariant* $\mathbf{P}_1 \bullet \mathbf{P}_2$ *in the appropriate frames of reference.*

(c) Elastic collision lemma. Consider an elastic collision between two particles, i.e. one in which the individual rest masses are preserved. Write **P**,**Q** for the pre-collision four-momenta and $\mathbf{P}'$,$\mathbf{Q}'$ for the post-collision four-momenta. By conservation of four-momentum we have $\mathbf{P}+\mathbf{Q}=\mathbf{P}'+\mathbf{Q}'$, which may be squared to give $\mathbf{P}^2+\mathbf{Q}^2+2\mathbf{P}.\mathbf{Q}= \mathbf{P}'^2+\mathbf{Q}'^2+2\mathbf{P}'.\mathbf{Q}'$. Now $\mathbf{P}^2= \mathbf{P}'^2$ and $\mathbf{Q}^2= \mathbf{Q}'^2$, since (i) individual rest masses are preserved in an elastic collision and (ii) equation (131) says that the square of a four-momentum is equal to the product of the rest mass squared and c^2. Hence we end up with the following formula, which is known as the "elastic collision lemma":

$$(135)\quad \mathbf{P} \bullet \mathbf{Q} = \mathbf{P}' \bullet \mathbf{Q}'.$$

By the last member of equations (134), the elastic collision lemma is seen to be equivalent to the statement that the relative speed of the particles is the same before and after the elastic collision.

26. Relativistic billiards

Here, we derive the rules of relativistic billiards, by which we mean the elastic collision of two particles of equal rest mass, one of which is originally at rest. Our analysis will be applicable to a variety of situations in relativistic particle physics, such as high-energy electron-electron collisions, and high energy proton-proton collisions, where one of the said particles is moving at high speed and the other is initially stationary in some reference frame (usually called the "laboratory frame").

We consider the scenario shown in the left part of Figure 21 below ("laboratory frame" S), where a particle travelling to the right at speed V strikes an initially-stationary particle of the same rest mass which is located at the origin of coordinates. After the collision, the two particles fly off in the directions indicated by θ and ϕ.

Now, relativistic problems are often greatly simplified by working in the frame of reference that makes the mathematics as simple as possible. Rather than working in the "laboratory frame" S shown on the left of Figure 21, we instead choose to work in the so-called "centre of momentum" frame S', which is on the right side of Figure 21. As its name implies, the centre-of-momentum frame is that frame in which the two particles originally approach one another with equal and opposite velocities ±v along the x' axis.

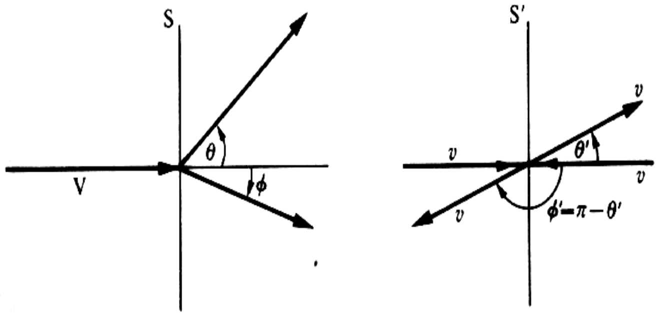


Figure 21: Two different inertial frames of reference. (a) Laboratory frame; (b) Centre-of-momentum frame. Note that these are in the usual "standard configuration" with respect to each other.[26]

By momentum conservation, the speed of the particles must be the same both before and after the collision, in the centre-of-momentum frame; we also explicitly assume that the nature (and hence the rest masses) of each particle is unchanged by the process of collision. Bearing all of the above in mind, Figure 21(b) shows both the pre-collision *and* post-collision velocities in the centre-of-momentum frame as ±v, but possibly along different lines. Note, furthermore, that the angles θ and φ in the laboratory frame S have been transformed to the angles θ' and $\phi'=\pi-\theta'$ in the centre-of-momentum frame S'.

As a next step, take the result of exercise 26 and apply a v-reversal transformation to it, yielding $\tan\alpha = \sin\alpha'/(\gamma(v)(\cos\alpha' + v/u))$. If we respectively let $\alpha=\theta$ and $\alpha=\phi$ in this formula, we end up with the following two expressions:

$$\text{(136a)}\quad \tan\theta = \frac{\sin\theta'}{\gamma(v)(\cos\theta'+1)},$$

$$\text{(136b)}\quad \begin{aligned}\tan\phi &= \frac{\sin\phi'}{\gamma(v)(\cos\phi'+1)}\\ &= \frac{\sin(\pi-\theta')}{\gamma(v)(\cos(\pi-\theta')+1)}\\ &= \frac{\sin(\pi)\cos(\theta')-\cos(\pi)\sin(\theta')}{\gamma(v)(\cos(\pi)\cos(\theta')+\sin(\pi)\sin(\theta')+1)}\\ &= \frac{\sin\theta'}{\gamma(v)(-\cos\theta'+1)}\end{aligned}$$

where we have made use of the facts that u=v and $\phi'=\pi-\theta'$. Multiply equations (136) together to give:

$$\text{(137)}\quad \begin{aligned}&\tan\theta\tan\phi\\ &= \frac{\sin\theta'}{\gamma(v)(\cos\theta'+1)}\times\frac{\sin\theta'}{\gamma(v)(-\cos\theta'+1)}\\ &= \frac{\sin^2\theta'}{\gamma^2(v)(1-\cos^2\theta')} = \frac{1}{\gamma^2(v)}\end{aligned}$$

Next, take equation (87b), which for the case of the incident billiard ball has $u' = u_1' = v$ and $u = V$, yielding, after a little algebra:

$$\text{(138)}\quad \frac{1}{\gamma^2(v)} = \frac{2}{\gamma(V)+1}.$$

This may be substituted into (137) to give our final result for relativistic billiards:

$$\text{(139)}\quad \tan\theta\tan\phi = \frac{2}{\gamma(V)+1}.$$

★ *Exercise 35: Derive equation (138).*

★ *Exercise 36: Show that, in the non-relativistic limit, where* $\tan\theta\tan\phi=1$, *we have a result that will be familiar to many of you who play billiards, where it is well known that the outgoing pair of billiard balls make a right angle with one another:*

$$\text{(140)}\quad \theta+\phi = 90^0,\ V \ll c.$$

For the relativistic case, show that we have:

$$\text{(141)}\quad \theta+\phi < 90^0.$$

---

[26] Figure taken from W. Rindler, *Introduction to Special Relativity*, Oxford University Press, Oxford (1991), p. 77.

Bearing the results of the previous exercise in mind, consider the experimental images shown in Figures 22 and 23. In both of these images, we see the tracks of an incident proton enter from the left; this incident proton then strikes a stationary proton, with the result that two protons emerge from the collision. For the low-energy case in Figure 22, we see that the opening angle is approximately a right angle; for the high-energy case, the opening angle is smaller than a right angle.

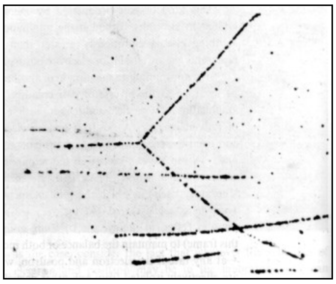


Figure 22: Elastic scattering of an incident proton of about 5 MeV by an initially stationary proton in a photographic emulsion. The incident proton enters from the left, and strikes an initially stationary proton.[27]

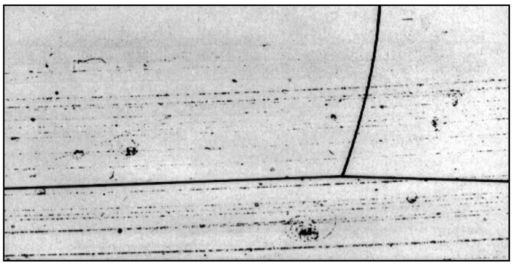


Figure 23: Elastic proton-proton collision in a liquid-hydrogen bubble chamber, using incident protons of about 3 GeV. The incident proton enters from the left, and strikes an initially stationary proton.[28]

★ *Exercise 37: If I told you that the image in Figure 23 corresponded to an elastic billiards-type collision between identical particles, and gave you no other information, how would you determine the incident velocity of the bullet particle?*

27. The centre-of-momentum frame

In the previous section we saw an example of the utility of working in the centre-of-momentum frame for a problem involving a two-particle collision. In that example, we defined the centre-of-momentum frame as that in which the total three-momentum vanishes. We will now generalise these arguments to the case of multiple particles which are involved in multiple collisions, together with other interactions such as splitting and fusion.

Consider some frame S in which there are a number of particles which are subject to no forces other than mutual collisions; we will also allow "self interactions" where a single particle may split into two pieces. We may then define the total mass $\overline{m}$, the total three-momentum $\overline{\mathbf{p}}$, and the total four-momentum $\overline{\mathbf{P}}$ as the sum, as a given instant of time, of the respective quantities for the individual particles:

$$\begin{aligned}\overline{m} &= \sum m, \quad \overline{\mathbf{p}} = \sum \mathbf{p}, \\ \overline{\mathbf{P}} &= \sum \mathbf{P} = \sum (mc, \mathbf{p}) = (\overline{m}c, \overline{\mathbf{p}})\end{aligned} \tag{142}$$

Because of the various conservation laws, all of the barred quantities are constant over time.

Now, since $\overline{\mathbf{P}}$ is a sum of four-vectors, $\overline{\mathbf{P}}$ is itself a four-vector. While this statement is correct, there is a subtlety worth discussing: consider Figure 24, which shows the world-lines in space-time for a group of particles which can undergo collisions, fusion and splitting. The planes Π and Π' represent planes of simultaneity in two different inertial frames S and S′ respectively. These two planes of simultaneity each represent a given instant of time in a given inertial frame; that these planes should be tilted with respect to one another can easily be inferred by looking at the Lorentz transformations. Now, in S, $\Sigma\mathbf{P}$ is summed over planes like Π; in S′, $\Sigma\mathbf{P}$ is summed over planes like Π'. We assert that the sum $\Sigma\mathbf{P}$ is the same for *whichever plane of simultaneity* (e.g. Π, Π' etc.) is used in the summation. To see this, imagine that we continuously rotate the plane Π into the plane Π'. As the plane of simultaneity is rotated, each individual four-momentum $\mathbf{P}$ of each particle remains constant except when our rotating plane of simultaneity sweeps over a collision; but then the sub-sum of $\Sigma\mathbf{P}$ which enters into the collision remains constant, by the conservation of four-momentum. Thus $\overline{\mathbf{P}} = \Sigma\mathbf{P}$ is the same for all inertial observers.

---

[27] Source: A. P. French, *Special Relativity*, Chapman and Hall, London (1968), page 192.

[28] Source: A. P. French, *Special Relativity*, Chapman and Hall, London (1968), page 192.

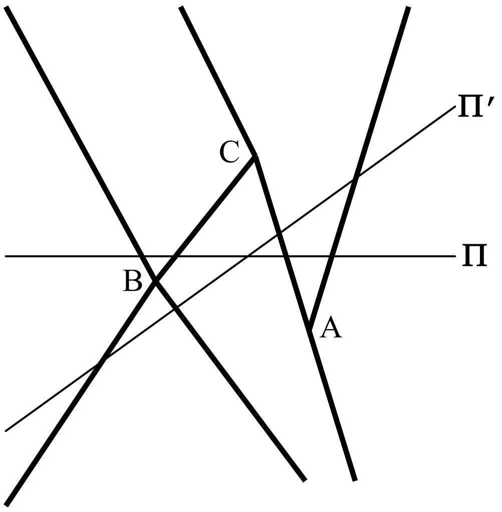


Figure 24: A tangle of world-lines (bold) for particles in space-time. The time axis is vertical. A collision event occurs at A, a "splitting" event[29] occurs at B, and a "fusion" event occurs at C. Π and Π′ are planes of simultaneity in two different frames of references.

We are now ready to introduce the *centre of momentum frame* (denoted by "S_{CM}") which forms the title of the present article. This frame S_{CM} moves with velocity $\mathbf{u}_{CM} = \overline{\mathbf{p}}/\overline{m}$ relative to S, and is therefore the frame in which $\overline{\mathbf{P}}$ has no spatial components.

Having defined the centre-of-momentum frame, it is natural to want to define the centre of mass frame. Now, in Special Relativity, the concept of centre of mass is frame-dependent; I have asked you to demonstrate this fact in the next exercise.

★ *Exercise 38: Show that, in Special Relativity, the concept of centre of mass is frame dependent. One way in which you might choose to do this is by considering two identical particles moving with equal and opposite velocities along parallel lines in some inertial frame S.*

Since the centre-of-mass is frame dependent, we can define the "proper" centre of mass as the centre of mass of a given system in its centre of momentum frame. By taking the third member of equation (142) as a starting point, we have:

$$(143)\quad \begin{aligned} \overline{\mathbf{P}} &= (\overline{m}c, \overline{\mathbf{p}}) \\ &= (\overline{m}c, \overline{m}\mathbf{u}_{CM}) \text{ since } \mathbf{u}_{CM} = \overline{\mathbf{p}}/\overline{m} \\ &= \overline{m}\gamma^{-1}(u_{CM})\gamma(u_{CM})(c, \mathbf{u}_{CM}) \\ &= \overline{m}\gamma^{-1}(u_{CM})\mathbf{U}_{CM} \text{ by equation (118)} \end{aligned}$$

where $u_{CM} \equiv |\mathbf{u}_{CM}|$. Now, by the quotient rule of tensor analysis, the quantity $\overline{m}\gamma^{-1}(u_{CM})$ in (143) must be an invariant (i.e. a rank zero Lorentz tensor). Denoting this invariant by $\overline{m}_{CM}$, we therefore have:

$$(144)\quad \overline{m} = \gamma(u_{CM})\overline{m}_{CM}$$

and:

$$(145)\quad \overline{\mathbf{P}} = \overline{m}_{CM}\mathbf{U}_{CM}.$$

28. Threshold energies

Consider a collision problem where a fast proton ("p"), of rest mass M, strikes a stationary proton. After the collision, we end up with the two said protons plus a pion ("π^0") of mass m. The reaction may be denoted thus: $p + p \rightarrow p + p + \pi^0$. One might ask, "What is the minimum energy of the incident proton, in the laboratory frame, which is needed in order for a pion to be created from a collision involving a moving proton colliding with a stationary proton?". This minimum energy is an example of a "threshold" energy, and will form the subject of the present article.

★ *Exercise 39: Regarding the example given in the previous paragraph, we assert that it is not enough for the energy of the incident proton to equal the sum of the rest mass energies of the proton and the pion, in order for the reaction* $p + p \rightarrow p + p + \pi^0$ *to take place. Why is this?*

Consider the collision of two particles, a "bullet" and a stationary "target", the pre-collision momenta of which are $\mathbf{P}_1$ and $\mathbf{P}_2$ respectively. The total four-momentum of the pre-collision system is $\overline{\mathbf{P}} = \mathbf{P}_1 + \mathbf{P}_2$, an expression which may be squared to give:

$$(146)\quad \mathbf{P}_1^2 + \mathbf{P}_2^2 + 2\mathbf{P}_1 \bullet \mathbf{P}_2 = \overline{\mathbf{P}}^2.$$

Now, from (128), the invariant $\mathbf{P}^2 = m^2c^2 - p^2$; since this is an invariant, we may evaluate it in the rest frame to give $\mathbf{P}^2 = m_0^2c^2$. Therefore the first two terms on the left side of (146) are equal to $m_{01}^2c^2$ and $m_{02}^2c^2$ respectively, where m_{01} denotes the rest mass of particle #1 and m_{02} denotes the rest mass of particle #2. By using (134), we see that the third term on the left side of (146) is equal to $2c^2m_{01}m_{02}\gamma(v)$, where v is the relative speed of the two particles. Lastly, we note that the right side of (146) is equal to

---

[29] As an example of such an event, a heavy nucleus might spontaneously disintegrate into a pair of lighter nuclei.

$\overline{m}_{CM}^2 c^2$, according to a similar argument to that which was used to derive the expressions for the first two terms on the left side of (146). Substituting all of these pieces into (146), and cancelling a common factor of c^2, we end up with:

$$(147) \quad m_{01}^2 + m_{02}^2 + 2m_{01}m_{02}\gamma(v) = \overline{m}_{CM}^2 .$$

Evidently, the minimum allowable gamma factor corresponds to the minimum allowable value for the right side of (147); this minimum allowed value for the right side will occur when all product particles *are at rest in the CM frame*[30], in which case $\overline{m}_{CM}$ is simply the sum of the rest masses of the post-collision particles.

Thus, for the case of the proton-proton example with which we opened this article, we have:

$$(148) \quad M^2 + M^2 + 2M^2\gamma(v_{thresh}) = (2M + m)^2 ,$$

from which we see that:

$$(149) \quad \gamma(v_{thresh}) = 1 + \frac{2m}{M} + \frac{m^2}{2M^2} .$$

Thus the required threshold speed is:

$$(150) \quad v_{thresh} = c\sqrt{1 - \frac{1}{\left(1 + \frac{2m}{M} + \frac{m^2}{2M^2}\right)^2}} .$$

★ *Exercise 40: Consider a "bullet" particle of rest mass M which strikes a stationary particle of rest mass M; after this collision, we have the original two particles plus a new particle of rest mass m. Define the efficiency "k" of this particle-creation process to be the ratio of the rest mass energy of the created particle to the kinetic energy of the bullet particle.*

(a) Show that this efficiency is equal to:

$$(151) \quad k = \frac{2}{4 + \frac{m}{M}}$$

and hence conclude that, if the rest mass m of the particle to be created is much less than M, then the efficiency is about 50%.

(b)[31] *Suppose that I asked you to create a so-called ψ particle by colliding electrons with anti-electrons*[32]*. Show that, for this case, the efficiency is:*

$$(152) \quad k = \frac{2}{\frac{m}{M} - 4\frac{M}{m}}$$

Evaluate this efficiency for the collision of a moving anti-electron with a stationary electron (note that the rest mass of the ψ is about 3700 time that of the electron). Since this efficiency is miniscule, can you think laterally and devise a simple way of making the process close to 100% efficient?

CHAPTER 8: RELATIVISTIC PARTICLE MECHANICS 2[33]

29. Photons

We have seen (see e.g. section 26) that the use of four vectors was useful in considering collisions involving material particles. Four vectors are also useful in the study of collisions which involve photons.

So, what is the expression for the four-momentum of a photon? Well, we saw in (130) that **P**=(E/c,**p**). Now, for a photon, the energy E is given by $E = h\nu$ where ν is the photon frequency and h is Planck's constant. Also, we saw in (3) that E=pc, i.e. $|\mathbf{p}|$=E/c. Bearing these points in mind, we wee that our formula for the four-momentum of a photon becomes:

$$(153) \quad \mathbf{P} = \left(\frac{E}{c}, \mathbf{p}\right) = \left(\frac{E}{c}, \frac{E}{c}\mathbf{n}\right) = \frac{E}{c}(1, \mathbf{n}) = \frac{h\nu}{c}(1, \mathbf{n})$$

where **n** is a unit vector which points in the direction of the photon's momentum.

Before considering an example of a collision problem involving photons, let us look at two simpler scenarios.

(a) Consider two isolated photons floating around through space-time; the respective four-momenta of this pair of photons will be denoted $\mathbf{P}_1$ and $\mathbf{P}_2$. In a given inertial reference frame, the first photon has

---

[30] This will occur if the post-collision system travels as a single lump.

[31] The topic of this exercise led to the award of the 1976 Nobel Prize in Physics, to Richter and Ting.

[32] Note that the electron-positron pair annihilate one another; you will need to make use of this fact in the derivation of (152).

[33] W. Rindler, *Introduction to Special Relativity*, Oxford University Press, Oxford (1991), chapter 5.

frequency ν_1 and travels in the direction $\mathbf{n}_1$, which the second photon has frequency ν_2 and travels in the direction $\mathbf{n}_2$. Now, by making use of (153), we see that:

$$\begin{aligned} \mathbf{P}_1 \bullet \mathbf{P}_2 &= \frac{h\nu_1}{c}(1,\mathbf{n}_1)\bullet\frac{h\nu_2}{c}(1,\mathbf{n}_2) \\ &= \frac{h^2\nu_1\nu_2}{c^2}(1-\mathbf{n}_1\bullet\mathbf{n}_2) \\ &= \frac{h^2\nu_1\nu_2}{c^2}(1-\cos\theta) \end{aligned} \tag{154}$$

where θ is the angle between $\mathbf{n}_1$ and $\mathbf{n}_2$.

(b) Next, consider a photon and a particle of non-zero rest mass m_0. Suppose the photon to have the four-momentum $\mathbf{P}=h\nu c^{-1}(1,\mathbf{n})$ (from equation (153)), and the particle to have the four-momentum $\mathbf{Q}=(mc,\mathbf{p})$ (from equation (128)) . Then $\mathbf{P}\bullet\mathbf{Q}$ is an invariant, which means that we can calculate it in the rest frame of the particle, and so we see:

$$\mathbf{P}\bullet\mathbf{Q} = \frac{h\nu}{c}m_0 c = h\nu m_0, \tag{155}$$

where ν is the frequency of the photon in the rest frame of the particle.

Compton Scattering. We are now ready to consider a collision problem involving a photon and a material particle. To this end, let us return to the problem of billiards shown in the left panel of Figure 21. In frame S, we will consider the "bullet" particle to be a photon of frequency ν which is incident along the x-axis of S. This incident photon strikes a stationary electron located at the origin of S. After the collision, the photon travels along the direction θ with a reduced frequency ν' and the recoiling electron travels along the direction ϕ. Denote by $\mathbf{P}$ and $\mathbf{P}'$ the pre-collision and post-collision four-momenta of the photon; similarly, denote by $\mathbf{Q}$ and $\mathbf{Q}'$ the pre-collision and post-collision four-momenta of the electron. By the conservation of four-momentum, we have $\mathbf{P}+\mathbf{Q}=\mathbf{P}'+\mathbf{Q}'$, which may be re-arranged to give $\mathbf{P}+\mathbf{Q}-\mathbf{P}'=\mathbf{Q}'$. This expression can then be squared:

$$\begin{aligned} &(\mathbf{P}+\mathbf{Q}-\mathbf{P}')\bullet(\mathbf{P}+\mathbf{Q}-\mathbf{P}') = \mathbf{Q}'^2 \\ &\Rightarrow \mathbf{P}^2+\mathbf{P}\bullet\mathbf{Q}-\mathbf{P}\bullet\mathbf{P}'+\mathbf{Q}\bullet\mathbf{P}+\mathbf{Q}^2 \\ &-\mathbf{Q}\bullet\mathbf{P}'-\mathbf{P}'\bullet\mathbf{P}-\mathbf{P}'\bullet\mathbf{Q}+\mathbf{P}'^2=\mathbf{Q}'^2 \\ &\Rightarrow \mathbf{P}^2+2\mathbf{P}\bullet\mathbf{Q}-2\mathbf{P}\bullet\mathbf{P}'+\mathbf{Q}^2 \\ &-2\mathbf{P}'\bullet\mathbf{Q}+\mathbf{P}'^2=\mathbf{Q}'^2 \end{aligned} \tag{156}$$

and use made of the facts that $\mathbf{P}^2=\mathbf{P}'^2=0$ (see (153)) and $\mathbf{Q}^2=\mathbf{Q}'^2$ (see (131)), to leave:

$$\begin{aligned} &2\mathbf{P}\bullet\mathbf{Q}-2\mathbf{P}\bullet\mathbf{P}'-2\mathbf{P}'\bullet\mathbf{Q}=0 \\ &\Rightarrow \mathbf{P}\bullet\mathbf{P}'=\mathbf{Q}\bullet(\mathbf{P}-\mathbf{P}') \end{aligned} \tag{156}$$

We now use (154) and (155) to rewrite (156):

$$\begin{aligned} &\frac{h^2\nu\nu'}{c^2}(1-\cos\theta)=h\nu m_0-h\nu' m_0 \\ &\Rightarrow \frac{h\nu\nu'}{c^2}(1-\cos\theta)=m_0(\nu-\nu') \end{aligned} \tag{157}$$

This expression is readily manipulated into the more usual form:

$$\lambda'-\lambda=\frac{2h}{cm_0}\sin^2\left(\frac{\theta}{2}\right), \tag{158}$$

which was experimentally verified by Compton in 1922. Since the right side of (158) is evidently positive, we see that the photon loses energy (increases its wavelength) in scattering from a stationary electron. Intuitively, this makes sense: if the photon has made the electron accelerate from rest to some non-zero speed, then the said photon must have lost energy in the process.

★ *Exercise 41: We have just considered the famous problem of Compton scattering, whereby a photon strikes a stationary electron; after the collision, we saw that the electron recoiled and that the photon lost energy (because its wavelength increases according to (158)). In this exercise, we consider what is called "Inverse Compton Scattering". Here, a photon collides with a moving electron; for such a collision, the photon may actually gain energy. Such a process is an important source of extragalactic x-rays. Using the same terminology that was introduced earlier, we have* $\mathbf{Q}=\gamma m(1,\mathbf{u})$, $\mathbf{P}=h\nu(1,\mathbf{n})$, *and* $P'=h\nu'(1,\mathbf{n}')$, *where we work in the so-called "natural units" that have* c=1. *For the extreme case of a head-on collision, where* $\mathbf{u}\bullet\mathbf{n}=-u$, $\mathbf{u}\bullet\mathbf{n}'=u$, $\mathbf{n}\bullet\mathbf{n}'=-1$, *show that:*

$$(159)\quad h\nu' = \frac{\gamma m(1+u)}{2+\left(\frac{\gamma m(1-u)}{h\nu}\right)}$$

and then briefly interpret your result.

30. Three-force and four-force

We have come a long way in Special Relativity with an absolutely minimal use of the concept of force. We will remedy this situation in the present section!

The four-force **F** is defined to be the proper-time rate-of-change of four-momentum:

$$(160)\quad \mathbf{F} \equiv \frac{d\mathbf{P}}{d\tau} = \frac{d}{d\tau}(m_0\mathbf{U}).$$

It is clear from the definition that **F** is indeed a Lorentz tensor. Now, proceeding from the definition (160), we can immediately write down:

$$(161)\quad \begin{aligned} \mathbf{F} &\equiv \frac{d\mathbf{P}}{d\tau} = \frac{dt}{d\tau}\frac{d}{dt}\mathbf{P} \underset{(116)}{=} \gamma\frac{d}{dt}\mathbf{P} \underset{(128)}{=} \gamma\frac{d}{dt}(mc,\mathbf{p}) \\ &= \gamma\left(c\frac{dm}{dt},\mathbf{f}\right) \underset{(8)}{=} \gamma\left(\frac{1}{c}\frac{dE}{dt},\mathbf{f}\right) \end{aligned}$$

where **f** is the relativistic three-force d**p**/dt and dE/dt is the power (rate at which the force transfers energy to the particle). Equation (161) relates the four-force **F** to the three-force **f** and the power.

The appearance of γ**f** as the spatial part of a four-vector let us work out how the three-force **f** transforms between different inertial reference frames. The result, the derivation of which is left to an exercise, is as follows:

$$(162)\quad \begin{cases} f_1' = \dfrac{f_1 - v\,dm/dt}{1-u_1 v/c^2} \\ f_2' = \dfrac{f_2}{\gamma(1-u_1 v/c^2)} \\ f_3' = \dfrac{f_3}{\gamma(1-u_1 v/c^2)} \end{cases}$$

★ *Exercise 42: Prove equations (162). Evidently, the three-force **f** is not invariant among frames, and depends both on the velocity of the particle on which it acts, and on the relative velocity of the frames.*

CHAPTER 9: RELATIVISTIC FIELDS: CLASSICAL ELECTRODYNAMICS[34]

31. Scalar and vector potentials for classical electromagnetic fields[35]

For sources in vacuum, the Maxwell equations for the electromagnetic field are:

$$(163a)\quad \nabla\bullet\vec{E} = 4\pi\rho$$

$$(163b)\quad \nabla\times\vec{B} - c^{-1}\partial_t\vec{E} = 4\pi c^{-1}\vec{j}$$

$$(163c)\quad \nabla\times\vec{E} + c^{-1}\partial_t\vec{B} = 0$$

$$(163d)\quad \nabla\bullet\vec{B} = 0$$

where $\vec{E}$ and $\vec{B}$ are the electric and magnetic fields, ρ is the charge density, and $\vec{J}$ is the current density. All of these quantities will in general be functions of both position and time. We are working in Gaussian units (see the appendix of Jackson's text if you are interested).

For later reference, note that both (163a) and (163b) are often called the *inhomogeneous Maxwell equations* due to the presence of the source terms for the charge density and the current density. Equations (163c) and (163d) are often called the *homogeneous Maxwell equations*, on account of the absence of source terms in these two equations.

On account of both (163d) and the vector identity $\nabla\bullet\nabla\times\vec{C} = \vec{0}$ for a suitably-well-behaved vector function $\vec{C}$, we can write the magnetic field $\vec{B}$ in terms of a vector potential $\vec{A}$ using the following formula:

$$(164)\quad \vec{B} = \nabla\times\vec{A},$$

which allows Maxwell equation (163c) to be written in the form:

$$(165)\quad \begin{aligned} &\nabla\times\vec{E} + c^{-1}\partial_t\left(\nabla\times\vec{A}\right) = 0 \\ &\Rightarrow \nabla\times\left(\vec{E} + c^{-1}\partial_t\vec{A}\right) = 0 \end{aligned}$$

---

[34] J.D.Jackson, *Classical Electrodynamics (second edition)*, John Wiley, New York (1965).

[35] J.D.Jackson, *Classical Electrodynamics (second edition)*, John Wiley, New York (1965), sections 1.1 and 6.4. See also the Appendix to the book.

Now, bearing in mind the vector identity $\nabla\times\nabla\Phi = 0$, we see that equation (165) implies that:

$$(166)\quad \begin{aligned} &\vec{E} + c^{-1}\partial_t\vec{A} = -\nabla\Phi \\ &\Rightarrow \vec{E} = -\nabla\Phi - c^{-1}\partial_t\vec{A}. \end{aligned}$$

Equations (164) and (166) show that, rather than working in terms of the *fields* $(\vec{E},\vec{B})$, we can instead choose to work in terms of the *potentials* $(\Phi,\vec{A})$. The fields can be derived from the potentials using (164) and (166). In the context of classical electrodynamics, the introduction of the potentials $(\Phi,\vec{A})$ is merely a mathematical artifice which makes our equations easier to work with[36].

Now, we have constructed our potentials so that the homogeneous Maxwell equations are *automatically satisfied* if one choose to work in terms of the said potentials. This is a useful simplification, since it means that we need only deal with two Maxwell equations rather than four. The question naturally arises, though, of what the inhomogeneous Maxwell equations look like when written in terms of the potentials rather than the fields.

Well, if we substitute (166) into (163a) then we arrive at:

$$(167)\quad \begin{aligned} &\nabla\bullet\left(-\nabla\Phi - c^{-1}\partial_t\vec{A}\right) = 4\pi\rho \\ &\Rightarrow \nabla^2\Phi + c^{-1}\partial_t\nabla\bullet\vec{A} = -4\pi\rho \end{aligned}$$

and if we substitute both (164) and (166) into (163b) we get:

$$(168)\quad \begin{aligned} &\nabla\times\left(\nabla\times\vec{A}\right) - c^{-1}\partial_t\left(-\nabla\Phi - c^{-1}\partial_t\vec{A}\right) \\ &\quad = 4\pi c^{-1}\vec{J} \end{aligned}$$

Next, recall the vector identity:

$$(169)\quad \nabla\times\left(\nabla\times\vec{A}\right) = \nabla\left(\nabla\bullet\vec{A}\right) - \nabla^2\vec{A}$$

to re-write (168) as:

$$(170)\quad \begin{aligned} &\nabla\left(\nabla\bullet\vec{A}\right) - \nabla^2\vec{A} \\ &\quad - c^{-1}\partial_t\left(-\nabla\Phi - c^{-1}\partial_t\vec{A}\right) = 4\pi c^{-1}\vec{J} \\ &\Rightarrow \nabla\left(\nabla\bullet\vec{A}\right) - \nabla^2\vec{A} + c^{-1}\partial_t\nabla\Phi + c^{-2}\partial_t^2\vec{A} = 4\pi c^{-1}\vec{J} \\ &\Rightarrow \left(\nabla^2 - c^{-2}\partial_t^2\right)\vec{A} - \nabla\left(\nabla\bullet\vec{A} + c^{-1}\partial_t\Phi\right) = -4\pi c^{-1}\vec{J} \end{aligned}$$

Equations (167) and (170) are the desired results, being the form that the inhomogeneous Maxwell equations take when we choose to work in terms of the potentials rather than in terms of the fields. However, these equations are often rather awkward to work with, because they are a pair of *coupled* partial differential equations for the scalar potential Φ and the vector potential $\vec{A}$.

Before simplifying equations (167) and (170), we need to take a brief tangent through the theory of gauge freedom. The fields $(\vec{E},\vec{B})$ can, as we have seen, be obtained from the potentials $(\Phi,\vec{A})$ using (164) and (166). These potentials are not unique; indeed, it can be shown that if the potentials $(\Phi,\vec{A})$ are replaced by the potentials $(\Phi',\vec{A}')$ according to the following prescription:

$$(171a)\quad \vec{A} \to \vec{A}' = \vec{A} + \nabla\Lambda$$

$$(171b)\quad \Phi \to \Phi' = \Phi - c^{-1}\partial_t\Lambda$$

then the fields $(\vec{E},\vec{B})$ remain *unchanged*. Here, Λ is an arbitrary function of position and time. The transformation of the potentials given by (171) is known as a “gauge transformation”; the associated freedom in choice of potentials is known as “gauge freedom”. The gauge freedom of the potentials serves to highlight the fact that the said potentials are mere mathematical artifices, with the fields themselves representing the physically-meaningful quantity.

★ *Exercise 43: Show that the gauge transformation (171) of the potentials leaves both the electric and magnetic fields unchanged.*

We can make use of gauge freedom by demanding that the scalar and vector potentials satisfy the so-called “Lorentz condition”:

$$(172)\quad \nabla\bullet\vec{A} + c^{-1}\partial_t\Phi = 0.$$

We are now ready to return to the question of simplifying equations (167) and (170). For, making use of the Lorentz condition, these equations become:

---

[36] This is not so for quantum mechanics ... but that is a story for another day! If interested, please check out some material on the Aharonov-Bohm effect.

(173a) $(\nabla^2 - c^{-2}\partial_t^2)\Phi = -4\pi\rho$,

(173b) $(\nabla^2 - c^{-2}\partial_t^2)\vec{A} = -4\pi c^{-1}\vec{J}$.

The good thing about this pair of partial differential equations is that they are both uncoupled (ie. we have one equation for the scalar potential Φ and another equation for the vector potential $\vec{A}$) and relatively simple (both are inhomogeneous D'Alembert-type wave equations).

★ *Exercise 44: Suppose that I give you a scalar and vector potential which do not satisfy the Lorentz condition. Show that a gauge transformation can always be made such that the transformed potentials now satisfy the Lorentz condition.*

32. Lorentz covariance of classical electrodynamics[37]

It is somewhat ironic that an "anomalous" experiment using *electromagnetic* radiation (Michelson-Morley experiment) should have ushered in Special Relativity, when in fact *the Maxwell and Lorentz-force equations which underpin such electromagnetic phenomena are already consistent with the formulations of Special Relativity*. In this section we shall demonstrate this Lorentz covariance (ie. invariance in form under Lorentz transformations) of the Maxwell and Lorentz-force equations for the classical electromagnetic field, by writing these equations in four-tensor form.

Let us begin with the well-known expression for the Lorentz force felt by a moving particle with charge q that moves with velocity $\vec{v}$ through a region of space in which there exist both electric and magnetic fields:

(174) $\vec{f} = \frac{d\vec{p}}{dt} = q\left(\vec{E} + \frac{1}{c}\vec{v}\times\vec{B}\right)$.

Now let us replace the derivative with respect to the time t by the derivative with respect to the proper time τ using the trick that we learned in equation (116), to give:

$$\frac{d\tau}{dt}\frac{d\vec{p}}{d\tau} = q\left(\vec{E} + \frac{1}{c}\vec{v}\times\vec{B}\right)$$

(175) $\overset{(116)}{\Longrightarrow} \frac{1}{\gamma}\frac{d\vec{p}}{d\tau} = q\left(\vec{E} + \frac{1}{c}\vec{v}\times\vec{B}\right)$.

$$\Longrightarrow \frac{d\vec{p}}{d\tau} = \frac{q}{c}\left(\gamma c\vec{E} + \gamma\vec{v}\times\vec{B}\right)$$

Now note from (118) that $\gamma c = U^0$, which is the temporal component of the four-velocity **U**. We can also see from (118) that $\gamma\vec{v}$ is the spatial component of the four-velocity **U**; we denote this spatial component of the four-velocity by $\vec{U}$, so that $\mathbf{U} \equiv (U^0, \vec{U})$. Bearing these comments in mind, we see that (175) can be written as:

(176) $\frac{d\vec{p}}{d\tau} = \frac{q}{c}\left(U^0\vec{E} + \vec{U}\times\vec{B}\right)$.

Now, the left side of (176) is simply the proper-time derivative of the spatial part of the energy-momentum four-vector **P**. It is natural, therefore, for us to see what the proper-time derivative of the temporal part of the energy-momentum four-vector looks like:

(177) $\frac{dP^0}{d\tau} \overset{(130)}{=} \frac{d}{d\tau}\frac{E}{c} = \frac{1}{c}\frac{dE}{dt}\frac{dt}{d\tau} \overset{(116)}{=} \frac{1}{c}\frac{dE}{dt}\gamma$.

Now, the power dE/dt is equal to the dot product of the force $q\vec{E}$ on the particle and the velocity $\vec{v}$ of the particle, which allows us to proceed further:

(178) $\frac{dP^0}{d\tau} = \frac{1}{c}\frac{dE}{dt}\gamma = \frac{q}{c}\vec{E}\bullet\vec{v}\gamma \overset{(118)}{=} \frac{q}{c}\vec{U}\bullet\vec{E}$.

We can see that the left sides of (176) and (178) are equal to the proper-time derivative of the spatial and temporal parts of the four-momentum, respectively. Thus, if this pair of laws is to form the two pieces of a single four-tensor law, then the right sides of these equations must form the temporal and spatial components of some four vector. This is indeed the case, but we shall leave the proof of this fact until the end of this section, for reasons that shall become clear in due course.

For now, change tack by noting that the famous continuity equation:

(179) $\partial_t\rho + \nabla\bullet\vec{j} = 0$

is implicit in the Maxwell equations.

---

[37] J.D.Jackson, *Classical Electrodynamics (second edition)*, John Wiley, New York (1965), section 11.9.

★ *Exercise 45: Show that the continuity equation (179) can be derived from the Maxwell equations. Briefly interpret this result.*

We now introduce a new four-vector, namely the four-current **J**, via the definition:

(180) $\mathbf{J} \equiv \rho^{(0)}\mathbf{U}$,

where **U** is the four-velocity and $\rho^{(0)}$ is the proper charge density at a given point (ie the charge density in the instantaneous rest frame of the piece of current). Since both terms on the right side of (180) are Lorentz tensors, the four-current must also be a Lorentz tensor, and so we are justified in calling it a "four vector".

Now, we can write down the following expression which relates the proper charge density $\rho^{(0)}$ to the charge density ρ:

(181) $\rho = \gamma\rho^{(0)}$.

This equation follows from (i) the experimentally-observed invariance of total charge under Lorentz transformation, together with (ii) length contraction. We can now make use of (181) to rework expression (180) for the four-current:

(182) $\mathbf{J} = \frac{\rho}{\gamma}\mathbf{U} \overset{(118)}{=} \frac{\rho}{\gamma}(\gamma c, \gamma\vec{v}) = (\rho c, \rho\vec{v}) = (c\rho, \vec{J})$.

The continuity equation (179) can now be written in the four-tensor form:

(183) $\partial_\alpha J^\alpha = 0$,

which is manifestly Lorentz covariant (ie invariant in form under a Lorentz transformation).

★ *Exercise 46: Show that (179) = (183).*

If we now introduce yet another four-vector[38], namely the four-potential **A**:

(184) $\mathbf{A} = (\Phi, \vec{A})$

then the Lorentz condition (172) becomes:

(185) $\partial_\alpha A^\alpha = 0$,

and the wave equations (173) become:

(186) $\Box\,\mathbf{A} = 4\pi c^{-1}\mathbf{J}$,

where the D'Alembertian $\Box$ is:

(187) $\Box \equiv \partial_\alpha \partial^\alpha = \partial_0^2 - \nabla^2 = c^{-2}\partial_t^2 - \nabla^2$.

★ *Exercise 47: Obtain (185) and (186).*

Now, equation (186) *shows* that **A** is a four-vector, because all other quantities appearing in this equation are four-tensors. Hence we are justified in calling **A** a four-vector, and both (185) and (186) are four-tensor equations.

As a mid-flight summary, let us point out that we have demonstrated the Lorentz covariance of:

- the continuity equation (see (179) & (183));
- the Lorentz condition (see (172) & (185));
- the wave equations which are satisfied by the electromagnetic potentials in the Lorentz gauge (see (173) & (186)).

We close this section by demonstrating the Lorentz covariance of two further things:

- The Maxwell equations (163);
- The Lorentz force equations ((176) & (178)).

Let us introduce the following second-rank, anti-symmetric field strength tensor:

(188) $F^{\alpha\beta} \equiv \partial^\alpha A^\beta - \partial^\beta A^\alpha$.

The tensorial nature of $F^{\alpha\beta}$ follows at once from the tensorial nature of A^α, together with the fact that the derivative of a Lorentz tensor is also a Lorentz tensor. Also, the antisymmetric nature of the field-strength tensor $F^{\alpha\beta}$, ie the fact that:

(189) $F^{\alpha\beta} = -F^{\beta\alpha}$,

follows directly from (188).

We claim that the various elements of the field strength tensor are directly related to the various components of the electric and magnetic field via:

(190) $F^{\alpha\beta} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}$.

---

[38] We are not yet justified in calling this a four-vector, but that justification will follow shortly.

★ *Exercise 48: Obtain (190) from (188).*

We also claim the Maxwell equations can be written in the four-tensorial form:

$$\text{(191a)}\quad \partial_\alpha F^{\alpha\beta} = 4\pi c^{-1} J^\beta \ \text{(inhomogeneous)}$$

$$\text{(191b)}\quad \partial^\alpha F^{\beta\gamma} + \partial^\gamma F^{\alpha\beta} + \partial^\beta F^{\gamma\alpha} = 0 \ \text{(homogeneous)}$$

and leave the proof of this to an exercise.

★ *Exercise 49: Show that (191a) is equivalent to the inhomogeneous Maxwell equations, and that (191b) is equivalent to the homogeneous Maxwell equations.*

Since all terms in equations (191) are Lorentz tensors, we conclude that the *Maxwell equations are invariant in form under the Lorentz transformations.* Classical electrodynamics is therefore consistent with Special Relativity[39], with the electric and magnetic field being "mixed up" into a single tensorial object $F^{\alpha\beta}$.

We close this section with the promised "cleaning up" of the calculation which led to equations (176) and (178). We have already pointed out that the left sides of (176) are the proper-time derivatives of the spatial and temporal components of the energy-momentum four-vector **P**. Now that we have introduced the concept of the field-tensor $F^{\alpha\beta}$, we can unite both (176) and (178) into the single tensor equation:

$$\text{(192)}\quad \frac{dP^\alpha}{d\tau} = \frac{q}{c} F^{\alpha\beta} U_\beta \,.$$

★ *Exercise 50: Show that (192) is equivalent to (176) and (178).*

33. Lorentz transformation of electromagnetic fields[40]

We have seen that the electric and magnetic fields are "mixed up" in the elements of a second-rank Lorentz tensor $F^{\alpha\beta}$. This field-strength tensor transforms according to:

$$\text{(193)}\quad F'^{\alpha\beta} = \frac{\partial x'^\alpha}{\partial x^\bullet} \frac{\partial x'^\beta}{\partial x^\blacksquare} F^{\bullet\blacksquare},$$

where the matrices $\partial x'^\alpha / \partial x^\bullet$ and $\partial x'^\beta / \partial x^\blacksquare$ can be determined from the Lorentz transformations. Once these matrixes have been determined, equation (193) can be used to obtain the transformation properties of the electric and magnetic fields. For transformations between the usual pair of inertial frames S and S′ which are in the standard configuration with respect to each other, we have:

$$\text{(194a)}\quad \begin{cases} E_{x'} = E_x \\ E_{y'} = \gamma\left(E_y - vB_z / c\right) \\ E_{z'} = \gamma\left(E_z + vB_y / c\right) \end{cases}$$

$$\text{(194b)}\quad \begin{cases} B_{x'} = B_x \\ B_{y'} = \gamma\left(B_y + vE_z / c\right) \\ B_{z'} = \gamma\left(B_z - vE_y / c\right) \end{cases}$$

★ *Exercise 51: Derive equations (194).*

The inverse versions of the transformations (194) are readily obtained via a v-reversal transformation.

Equations (194) show that the electric and magnetic fields have no independent existence. A field which is purely electric or purely magnetic in one inertial reference frame will appear as a mixture of electric and magnetic fields in a different inertial reference frame. Since the electric and magnetic fields are so intertwined, one should more properly speak of "the electromagnetic field $F^{\alpha\beta}$" rather than the separate electric and magnetic fields.

To make these remarks more concrete, let us consider the example of a point charge which lies at the origin of some moving inertial reference frame S′ (see Figure 25).

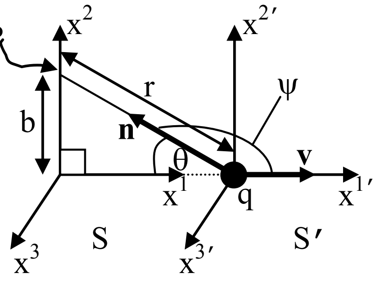


Figure 25: Diagram to help calculate the electromagnetic field for a moving charge q which is nailed to the origin of the moving inertial frame S′

---

[39] Note, in this context, that Einstein's "Special Relativity" paper of 1905 is called *On the Electrodynamics of Moving Bodies.* An Adobe PDF file of this paper is available at: http://www.fourmilab.ch/etexts/einstein/specrel/specrel.pdf.

[40] J.D.Jackson, *Classical Electrodynamics (second edition)*, John Wiley, New York (1965), section 11.10.

Here, the inertial frames S and S' are in the usual standard configuration with respect to one another. The point charge q is nailed to the origin of frame S', which moves with speed v to the right. We will consider field measurements made by an observer at P, who is attached to the vertical S axis at a height b above the origin. Denote by **n** the unit vector which (according to S) links the charge to the observation point P; in S', this vector will be called $\mathbf{n}'$. Similarly, we have the angles (θ, ψ) in S which become (θ', ψ') in S'.

From the point of view of someone in the inertial frame S', the charge is stationary and so space is filled with the purely electric field due to the stationary charge q. In this frame, P is a moving point of observation, but this does not change the fact that, according to S', the field at point P is a purely electric field due to a point charge. Thus, according to S', the electric and magnetic field at P is:

$$\text{(195a)}\quad \begin{cases} E_{x'} = -\dfrac{q\cos\theta'}{r'^2} = \dfrac{-qvt'}{r'^3} \\ E_{y'} = \dfrac{q\sin\theta'}{r'^2} = \dfrac{qb}{r'^3} \\ E_{z'} = 0 \end{cases}$$

$$\text{(195b)}\quad \begin{cases} B_{x'} = 0 \\ B_{y'} = 0 \\ B_{z'} = 0 \end{cases}$$

If we re-write the right sides of each of these equations in terms of the coordinates of S, we end up with the following expression for the non-vanishing components of the electric and magnetic fields:

$$\text{(196)}\quad \begin{cases} E_{x'} = \dfrac{-q\gamma vt}{\left(b^2+\gamma^2v^2t^2\right)^{3/2}}, \\ E_{y'} = \dfrac{qb}{\left(b^2+\gamma^2v^2t^2\right)^{3/2}}, \end{cases}$$

where we have made use of the fact that $t' = \gamma\left(t - vx/c^2\right) = \gamma t$ at P, since x=0 for the point P in the frame S.

By making use of the v-reversed form of (194) we can take (196) and derive the following formula, the proof of which is left to an exercise:

$$\text{(197)}\quad \begin{cases} E_x = E_{x'} = -\dfrac{qv\gamma t}{\left(b^2+\gamma^2v^2t^2\right)^{3/2}} \\ E_y = \gamma E_{y'} = \dfrac{\gamma qb}{\left(b^2+\gamma^2v^2t^2\right)^{3/2}} \\ B_z = \gamma v E_{y'}/c = vE_y/c \\ \text{All other components vanish} \end{cases}$$

★ *Exercise 52: Derive equations (197).*

We see that a purely electric field in S' appears as a combined electric and magnetic field in S.

From the second member of (197), the peak transverse electric field is evidently $E_y(t=0)$, which has the value:

$$\text{(198)}\quad E_y(t=0) = \frac{\gamma qb}{\left(b^2+0\right)^{3/2}} = \frac{\gamma q}{b^2}.$$

This value for the peak transverse electric field is γ times larger than its non-relativistic value. Also, the value of the peak transverse electric field becomes arbitrarily large as $\gamma \to \infty$, i.e. as $v \to c$.

Having investigated the *peak* value of the transverse electric field of the moving charge relative to an observer at P, let us look at the *duration* of this field felt by the observer. To determine this, think of the second member of (197) as a graph of E_y versus time (see Figure 26). The peak height of this graph has already been calculated in (198) as $\gamma q/b^2$; the full-width at half maximum is easily calculated to be:

$$\text{(199)}\quad \Delta t = \frac{2b\sqrt{2^{2/3}-1}}{\gamma v} \approx \frac{b}{\gamma v}.$$

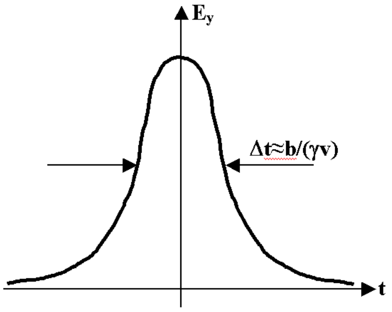


Figure 26: Electric field at the observation point P (see Fig. 25), as a function of time, for a uniformly-moving charged particle

We see, therefore, that as the gamma factor increases, the *peak* transverse field increases in proportion, but the *duration* goes in inverse proportion, to the said gamma factor. As the velocity of the uniformly-moving charged particle approaches that of light, the observer at P sees nearly equal transverse and mutually perpendicular electric and magnetic fields, which are indistinguishable from the fields of a pulse of plane polarized radiation propagating in the x direction.

★ *Exercise 53: In the limit of high velocity, the E_x component given by the first member of (197) becomes irrelevant for any realistic detector. Why? Also, justify the statement that, as the velocity of the uniformly-moving charged particle approaches that of light, the observer at P sees nearly equal transverse mutually perpendiular electric and magnetic fields, which are indistinguishable from the fields of a pulse of plane polarized radiation propagating in the x direction.*

Another way of looking at the problem of a uniformly-moving charge is to obtain the spatial variation of the fields relative to the instantaneous present position of the charge in the laboratory (i.e. frame S). Now, from the first two members of (197) we see that $E_x / E_y = -vt / b$. With reference to Figure 25, we conclude that the electric field is observed to be directed along **n**, which is a radial unit vector from the charge's present position to the observation point (this is as it would be for the case of a static Coulomb field). By expressing the denominator in (197) in terms of r, which is the radial distance from the present position of the charge to the observer (see Fig. 25), and the angle ψ, we obtain the following for the electric field in terms of the charge's present position:

$$(200)\quad \mathbf{E} = \frac{q\mathbf{r}}{r^3\gamma^2\left(1-\beta^2\sin^2\psi\right)^{3/2}}, \quad \beta \equiv v/c.$$

★ *Exercise 54: Derive equation (200).*

The field is radial, but the lines of force are isotropically distributed only for $\beta = 0$. Along the direction of motion ($\psi = 0, \pi$), the field strength is down by a factor of γ^{-2} relative to isotropy, while in the transverse directions ($\psi = \pi/2$) it is larger by a factor of γ. The whiskbroom pattern of the lines of force, shown in Figure 27, is the spatial "snapshot" equivalent of the temporal behaviour sketched in Fig. 26.

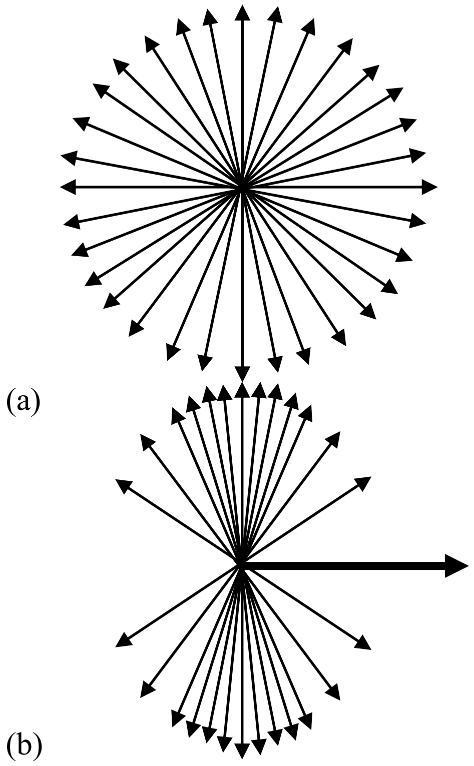


Figure 27: Lines of electric force for a charged particle (a) at rest and (b) in motion (γ≈3)

The compression of the lines of force in the transverse direction can be viewed as a consequence of the Lorentz contraction.

We end this section on a historical note. Our length contraction result is purely a consequence of Maxwell's equations and can be obtained without explicit use of Special Relativity. Lorentz himself so obtained it, and thereon based his "explanation" of the Lorentz contraction of material bodies (which *predates* Special Relativity) which he invoked in an attempt to understand the null result of the Michelson-Morley experiment. Lorentz argued that, if charges "contract", then so must all matter.

34. The Euler-Lagrange field equations[41]

The Euler-Lagrange equations, derived in the present section, are exceptionally important in both the classical and quantum theory of fields. They give a means for deriving field equations (be they classical or quantum-mechanical) from a so-called *Lagrangian density.* Here, we give a brief introduction to the

---

[41] F. Mandl and G.Shaw, *Quantum Field Theory (Revised Edition)*, John Wiley and Sons, Chichester (1993), section 2.2; J.D.Jackson, *Classical Electrodynamics (second edition)*, John Wiley, New York (1965), section 12.8. Note that the book by Mandl and Shaw also points out how *classical* Lagrangian field theory (i.e., the subject of section 34 of our course) is of great utility in developing a corresponding *quantum* theory of fields.

Euler-Lagrange equations. Many field equations in physics can be derived using the formalism presented here. Examples include the Klein-Gordon equation of relativistic quantum mechanics (exercise 55), the Maxwell equations of classical electrodynamics (exercise 56), the Dirac equation of relativistic quantum mechanics, the so-called "Standard Model" of Particle Physics, and the Einstein Field Equations of General Relativity. Note also that a variety of non-relativistic field equations (such as the Schrödinger equation) can also be derived using the Euler-Lagrange formalism.

In view of the very great generality of the formalism we are about to derive, we consider a system which requires several fields $\phi_r(x), r = 1,2,..., n$ to specify it, where x is a position vector in space-time. In the context of Lagrangian field theory, we restrict ourselves to theories which can be derived by means of a variational principle from an action integral involving the Lagrangian density $\mathcal{L} = \mathcal{L}(\phi_r, \phi_{r,\alpha})$, where:

(201a) $\partial\phi / \partial x^\mu \equiv \partial_\mu \phi \equiv \phi_{,\mu}$

(201b) $\partial\phi / \partial x_\mu \equiv \partial^\mu \phi \equiv \phi_,{}^\mu$.

The Lagrangian density $\mathcal{L}$ is related to the more familiar Lagrangian L by the formula:

(202) $L(t) = \int \mathcal{L}(\phi_r, \phi_{r,\alpha}) dx\, dy\, dz$.

Note that the Lagrangian density is considered to depend on the fields and their first derivatives only, which is why we write $\mathcal{L} = \mathcal{L}(\phi_r, \phi_{r,\alpha})$.

Define the action integral $S(\Omega)$ for an arbitrary region Ω of the space-time continuum by:

(203) $S(\Omega) = \int_\Omega d^4x\, \mathcal{L}(\phi_r, \phi_{r,\alpha})$,

where d^4x stands for the four-dimensional volume element $dx^0 dx^1 dx^2 dx^3$. The Lorentz-invariant nature of the action integral follows if the Lagrangian density is a Lorentz scalar (because the four-dimensional volume element is invariant).

For an arbitrary region Ω, we consider a variation of the fields which vanishes on the surface $\Gamma(\Omega)$ bounding the region Ω:

(204a) $\phi_r(x) \rightarrow \phi_r(x) + \delta\phi_r(x)$,

where

(204b) $\delta\phi_r(x) = 0$ on $\Gamma(\Omega)$.

We demand that for an arbitrary region Ω, this variation should have a stationary value, i.e. that:

(205) $\delta S(\Omega) = 0$.

We shall now show that this demand leads to the famous Euler-Lagrange equations:

(206) $$\frac{\partial\mathcal{L}}{\partial\phi_r} - \frac{\partial}{\partial x^\alpha}\left(\frac{\partial\mathcal{L}}{\partial\phi_{r,\alpha}}\right) = 0, \quad r = 1,2,..., N .$$

The proof of (206) begins by substituting expression (203) into equation (205):

$$\begin{aligned} 0 &= \delta S(\Omega) \\ (207) \qquad &= \delta \int_\Omega d^4x\, \mathcal{L}(\phi_r, \phi_{r,\alpha}) \text{... from (203)} \\ &= \int_\Omega d^4x \left\{ \frac{\partial\mathcal{L}}{\partial\phi_r}\delta\phi_r + \frac{\partial\mathcal{L}}{\partial\phi_{r,\alpha}}\delta\phi_{r,\alpha} \right\} \end{aligned}$$

To proceed further, we need to work out how to do the four-dimensional version of an integration by parts. To this end, consider the following application of the chain rule:

$$\begin{aligned} &\frac{\partial}{\partial x^\alpha}\left(\frac{\partial\mathcal{L}}{\partial\phi_{r,\alpha}}\delta\phi_r\right) \\ (208) \qquad &= \left(\frac{\partial}{\partial x^\alpha}\frac{\partial\mathcal{L}}{\partial\phi_{r,\alpha}}\right)\delta\phi_r + \frac{\partial\mathcal{L}}{\partial\phi_{r,\alpha}}\left(\frac{\partial}{\partial x^\alpha}\delta\phi_r\right) \\ &\equiv \left(\frac{\partial}{\partial x^\alpha}\frac{\partial\mathcal{L}}{\partial\phi_{r,\alpha}}\right)\delta\phi_r + \frac{\partial\mathcal{L}}{\partial\phi_{r,\alpha}}\delta\phi_{r,\alpha} \end{aligned}$$

which may be re-arranged to give:

$$\begin{aligned} &\frac{\partial\mathcal{L}}{\partial\phi_{r,\alpha}}\delta\phi_{r,\alpha} \\ (209) \qquad &= \frac{\partial}{\partial x^\alpha}\left(\frac{\partial\mathcal{L}}{\partial\phi_{r,\alpha}}\delta\phi_r\right) - \left(\frac{\partial}{\partial x^\alpha}\frac{\partial\mathcal{L}}{\partial\phi_{r,\alpha}}\right)\delta\phi_r . \end{aligned}$$

The left side of this equation is equal to the second term in braces in (207). Therefore equation (207) can be re-written as:

$$0=\int_{\Omega} d^4x\left\{\begin{array}{c}\dfrac{\partial\mathcal{L}}{\partial\phi_r}\delta\phi_r+\dfrac{\partial}{\partial x^{\alpha}}\left(\dfrac{\partial\mathcal{L}}{\partial\phi_{r,\alpha}}\delta\phi_r\right)\\ -\left(\dfrac{\partial}{\partial x^{\alpha}}\dfrac{\partial\mathcal{L}}{\partial\phi_{r,\alpha}}\right)\delta\phi_r\end{array}\right\}$$

$$(210)\qquad =\int_{\Omega} d^4x\left\{\frac{\partial\mathcal{L}}{\partial\phi_r}-\left(\frac{\partial}{\partial x^{\alpha}}\frac{\partial\mathcal{L}}{\partial\phi_{r,\alpha}}\right)\right\}\delta\phi_r+\int_{\Omega} d^4x\frac{\partial}{\partial x^{\alpha}}\left(\frac{\partial\mathcal{L}}{\partial\phi_{r,\alpha}}\delta\phi_r\right)$$

This completes our four-dimensional integration by parts.

Now, using the four-dimensional version of the Gauss divergence theorem, the last volume integral in equation (210) can be converted into a surface integral over the surface $\Gamma(\Omega)$; this surface integral vanishes on account of (204b). Therefore (210) becomes:

$$(211)\qquad \int_{\Omega} d^4x\left\{\frac{\partial\mathcal{L}}{\partial\phi_r}-\left(\frac{\partial}{\partial x^{\alpha}}\frac{\partial\mathcal{L}}{\partial\phi_{r,\alpha}}\right)\right\}\delta\phi_r=0\,.$$

Since Ω and $\delta\phi_r$ are both arbitrary, the term in braces must be zero. This leads directly to (206), thus completing our derivation of the Euler-Lagrange equations.

★ *Exercise 55: Consider the following Lagrangian density:*

$$(212)\qquad \mathcal{L}=\frac{1}{2}\left(\phi_{,\alpha}\phi_{,}{}^{\alpha}-\mu^2\phi^2\right)$$

for a single real field $\phi(x)$*, where* μ *is a constant. Show that the equation of motion for this field is the famous Klein-Gordon equation of relativistic quantum mechanics:*

$$(213)\qquad (\Box+\mu^2)\phi(x)=0\,,$$

where:

$$(214)\qquad \Box\equiv c^{-2}\partial_t^2-\nabla^2=\partial^{\mu}\partial_{\mu}\,.$$

★ *Exercise 56: Consider the following Lagrangian density:*

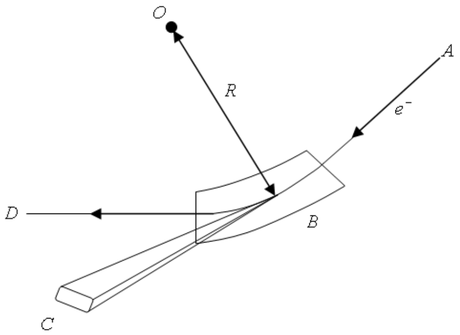
$$(215)\qquad \mathcal{L}=-\frac{1}{16\pi}F_{\alpha\beta}F^{\alpha\beta}-\frac{1}{c}J_{\alpha}A^{\alpha}\,.$$

Show that this leads to the tensorial form (191a) of the inhomogeneous Maxwell equations. Show also that (183) follows directly from (191a).

APPENDIX: Why is synchrotron radiation so directional? [42]

"Synchrotron radiation" may be defined as "the electromagnetic radiation emitted by relativistic charged particles that are accelerated along curved trajectories".[43]

This radiation, which may be produced by both terrestrial sources (such as the Australian Synchrotron) and astrophysical sources, is known to be highly directional.



To illustrate this "highly directional nature" of synchrotron radiation, consider the figure above.[44] Here, a charged point particle (electron) enters a magnetic field B from the point A, with the associated Lorentz ("**v** cross **B**") force serving to curve the trajectory of the electron as shown. When traversing the curved region of the trajectory, a narrow cone C of synchrotron radiation is emitted.

Here, we address a question raised in today's tutorial, regarding what insight our Special Relativity course may give to the question of why X-ray synchrotron radiation is so highly directional.

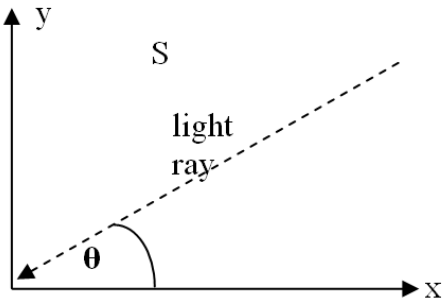
To this end, consider Figure 19 of the Special Relativity notes:

---

[42] Thanks to Nichol Furey, from the Monash Special Relativity class of 2003, for developing the line of analysis given here.

[43] D.M. Paganin, *Coherent X-Ray Optics*, Oxford University Press, New York (2006), p. 139.

[44] Ibid., p. 142.



Here we see an incoming ray of light whose negative direction makes an angle of θ with the x axis of an inertial frame S. Introduce the usual second inertial frame S' in "standard configuration" relative to S; let θ' denote the angle made by the negative direction of the light ray with the x' axis of S'.

In the lecture notes, we obtained the following formula relating θ and θ':

$$(95)\quad \sin\theta' = \frac{\sin\theta}{\gamma\left(1+\frac{v}{c}\cos\theta\right)},$$

where $\gamma = 1/(1-v^2/c^2)^{1/2}$ is the usual "gamma factor".

If we apply a v-reversal to this formula [see Exercise #12] we immediately obtain:

$$(A)\quad \sin\theta = \frac{\sin\theta'}{\gamma\left(1-\frac{v}{c}\cos\theta'\right)}.$$

Now let $\theta' = \pi/2$, which corresponds to the electron emitting a photon in the negative y' direction. Making use of the fact that $\sin\theta' = \sin(\pi/2) = 1$ and $\cos\theta' = \cos(\pi/2) = 0$, equation (A) becomes:

$$(B)\quad \sin\theta = \frac{\sin\theta'}{\gamma\left(1-\frac{v}{c}\cos\theta'\right)} \to \frac{1}{\gamma\left(1-\frac{v}{c}.0\right)} = \gamma^{-1}.$$

Now assume that the electron is highly relativistic, so that $v = c - \varepsilon$, where ε is a positive number that is much smaller than c. This implies that

$$\begin{aligned} \gamma^{-1} &= (1-v^2/c^2)^{1/2} \\ &= (1-(c-\varepsilon)^2/c^2)^{1/2} \\ &= (1-(c-\varepsilon)^2/c^2)^{1/2} \\ (C)\qquad &= (1-(1-\tfrac{\varepsilon}{c})^2)^{1/2} \qquad . \\ &\approx (1-(1-2\tfrac{\varepsilon}{c}))^{1/2} \\ &= \sqrt{2\varepsilon/c} \\ &<< 1. \end{aligned}$$

Since (B) says that $\sin\theta = \gamma^{-1}$, and the highly relativistic nature of the electron implies that $\gamma^{-1} << 1$, we conclude that we can make the small-angle approximation $\sin\theta \approx \theta$ in equation (B). Hence equation (B) becomes:

(D) $\theta \approx \gamma^{-1} << 1$ [note that θ is in radians].

Thus, photons which are emitted at right angles to the direction x' of propagation in the "instantaneous rest frame" S', will (from the perspective of the "lab frame" S) make a small angle $\theta \approx \gamma^{-1}$ radians with respect to the x axis.

Hence we see that our studies on "Aberration and visual appearance of moving objects", in section 18 of the Special Relativity notes, may be used to show that *synchrotron radiation is strongly concentrated in the direction of the charge's velocity, in a forward-pointing cone of half-angle* $\theta \approx \gamma^{-1}$.

Example: For the Australian synchrotron's 3 GeV beam, $\gamma = 5871$.[45] Hence $\theta \approx 1/5871 \approx 200$ microradians, which is indeed highly directional.

* * *

Note: Some of the equation numbers have gone haywire in chapter 3 and 4. Specifically, at the end of chapter 3 the last equation is number 67, and in the next chapter we jump to 74 for the next equation. The equation after that is then numbered 68, and the numbering increases from there on (including another equation 74).

---

[45] http://accelconf.web.cern.ch/AccelConf/e08/papers/thpc135.pdf