Chapter 9. Special Relativity

This chapter starts with a review of special relativity's basics, including its very convenient 4-vector formalism. This background is then used for the analysis of the relation between the electromagnetic field's values measured in different inertial reference frames moving relative to each other. The results enable us to discuss relativistic particle dynamics in the electric and magnetic fields, and the analytical mechanics of the particles – and of the electromagnetic field as such.

9.1. Einstein postulates and the Lorentz transform

As was emphasized at the derivation of expressions for the dipole and quadrupole radiation in the last chapter, they are only valid for systems of non-relativistic particles moving with velocities **u** much lower than c. In order to generalize these results to particles moving with arbitrary **u**, we need help from the relativity theory. Moreover, an analysis of the motion of charged relativistic particles in electric and magnetic fields is also a natural part of electrodynamics. This is why I will follow the tradition of using this course for a (by necessity, brief) introduction to the special relativity theory. This theory is based on the fundamental idea that measurements of physical variables (including the spatial and even temporal intervals between two events) may give different results in different reference frames, in particular in two inertial frames moving relative to each other translationally (i.e. without rotation), with a certain constant velocity **v** (Fig. 1).

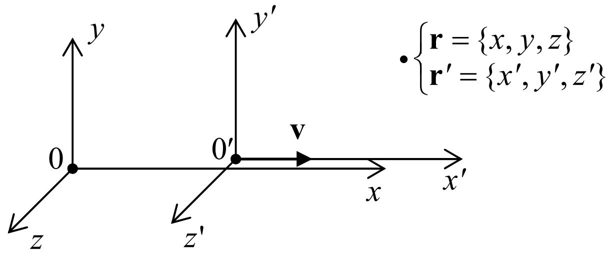


Fig. 9.1. Translational mutual motion of two reference frames.

In the non-relativistic (Newtonian) mechanics the problem of transfer between such reference frames has a simple solution at least in the limit $v << c$, because the basic equation of particle dynamics (the 2nd Newton law) [1]

$$m_k \ddot{\mathbf{r}}_k = -\nabla_k \sum_{k'} U(\mathbf{r}_k - \mathbf{r}_{k'}), \tag{9.1}$$

where U is the potential energy of inter-particle interactions, is invariant with respect to the so-called *Galilean transformation* (or just "transform" for short).[2] Choosing the coordinates in both frames so that their axes x and x' are parallel to the vector **v** (as in Fig. 1), the transform may be represented as

[1] Let me hope that the reader does not need a reminder that for Eq. (1) to be valid, the reference frames 0 and 0' have to be inertial – see, e.g., CM Sec. 1.2.

[2] It had been first formulated by Galileo Galilei, if only rather informally, as early as 1638 – four years before Isaac Newton was *born*! Note also the very unfortunate term "boost" used sometimes to describe such translational transformations. (It is especially unnatural in the special relativity, not describing accelerations.) In my course, this term is avoided, with the equivalent "transform" used instead.

$$\boxed{x = x' + vt', \qquad y = y', \qquad z = z', \qquad t = t',} \tag{9.2a}$$

Galilean transform

and plugging Eq. (2a) into Eq. (1), we get an absolutely similarly looking equation of motion in the "moving" reference frame 0'. Since the reciprocal transform,

$$x' = x - vt, \qquad y = y', \qquad z' = z, \qquad t' = t, \tag{9.2b}$$

is similar to the direct one, with the replacement of (+v) with (–v), we may say that the Galilean invariance means that there is no "master" (*absolute*) spatial reference frame in classical mechanics, although the spatial and temporal intervals between different instant events are absolute, i.e. reference-frame invariant: $\Delta x = \Delta x'$,…, $\Delta t = \Delta t'$.

However, it is straightforward to use Eq. (2) to check that the form of the wave equation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) f = 0, \tag{9.3}$$

describing, in particular, the electromagnetic wave propagation in free space,[3] is *not* Galilean-invariant.[4] For the "usual" (say, elastic) waves, which obey a similar equation albeit with a different speed,[5] this lack of Galilean invariance is natural and is compatible with the invariance of Eq. (1), from which the wave equation originates. This is because the elastic waves are essentially the oscillations of interacting particles of a certain medium (e.g., an elastic solid), making the reference frame connected to this medium, special. So, if the electromagnetic waves were oscillations of a certain special medium (which was first called the "luminiferous aether"[6] and later *aether* – or just "ether"), similar arguments might be applicable to reconcile Eqs. (2) and (3).

The detection of such a medium was the goal of the measurements carried out between 1881 and 1887 (with better and better precision) by Albert Abraham Michelson and Edward Williams Morley, which are sometimes called "the most famous failed experiments in physics". Figure 2 shows a crude scheme of these experiments.

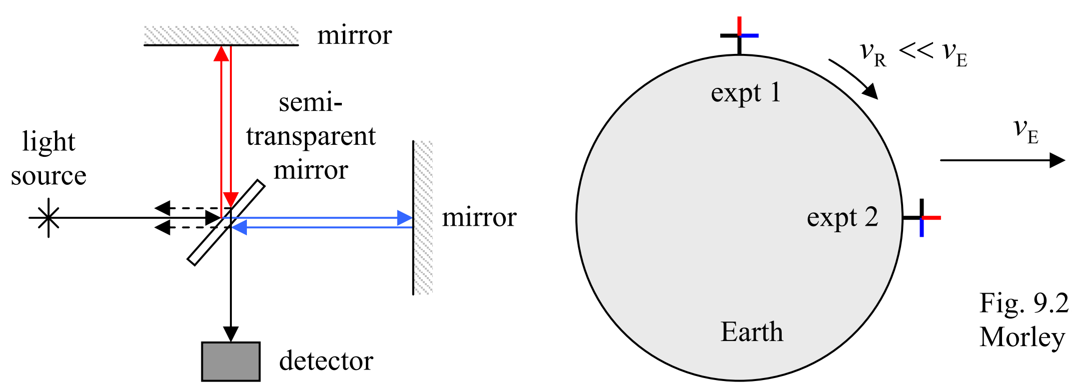


Fig. 9.2. The Michelson-Morley experiment.

[3] The discussions in this chapter and most of the next chapter will be restricted to the free-space (and hence dispersion-free) case; some media effects on the radiation by relativistic particles will be discussed in Sec.10.4.

[4] It is interesting that the usual (non-relativistic) Schrödinger equation, whose fundamental solution for a free particle is a similar monochromatic wave (albeit with a different dispersion law), *is* Galilean-invariant, with a certain change of the wavefunction's phase – see, e.g., QM Chapter 1.

[5] See, e.g., CM Secs. 6.5 and 7.7.

[6] In ancient Greek mythology, aether is the clean air breathed by the gods residing on Mount Olympus.

A nearly monochromatic wave from a light source is split into two parts (optimally, of equal intensity), using a semi-transparent mirror tilted by the angle $\pi/4$ to the incident wave direction. These two partial waves are reflected back by two fully-reflecting mirrors and arrive at the same semi-transparent mirror again. Here half of each wave is directed toward the light source (they vanish there without affecting the source), but another half is passed toward an intensity detector, forming, with its counterpart, an interference pattern similar to that in the Young experiment. Thus each of the interfering waves has traveled twice (back and forth) each of two mutually perpendicular "arms" of the interferometer. Assuming that the aether, in which light propagates with speed c, moves with speed $v < c$ along one of the arms, of length l_l, it is straightforward (and hence left for the reader's exercise :-) to get the following expression for the difference between the light roundtrip times:

$$\Delta t = \frac{2}{c}\left[\frac{l_t}{\left(1 - v^2/c^2\right)^{1/2}} - \frac{l_l}{1 - v^2/c^2}\right] \approx \frac{l}{c}\left(\frac{v}{c}\right)^2, \qquad (9.4)$$

where l_t is the length of the second, "transverse" arm of the interferometer (perpendicular to $\mathbf{v}$), and the last, approximate expression is valid at $l_t \approx l_l \equiv l$ and $v << c$.

Since the Earth moves around the Sun with a speed $v_E \approx 30$ km/s $\approx 10^{-4}\ c$, the arm positions relative to this motion alternate, due to the Earth's rotation about its axis, every 6 hours – see the right panel of Fig. 2. Hence if we assume that the aether rests in the Sun's reference frame, then Δt (and the corresponding shift of the interference fringes), has to change its sign with this half-period as well. The same alternation may be achieved, at a smaller time scale, by a deliberate rotation of the instrument by $\pi/2$. In the most precise version of the Michelson-Morley experiment (circa 1887), this shift was expected to be close to 0.4 of the interference pattern period. The results of the search for such a shift were negative, with the error bar about 0.01 of the period.[7]

The most prominent immediate explanation for this zero result[8] was suggested in 1889 by George Francis FitzGerald and (independently and more qualitatively) by H. Lorentz in 1892: as evident from Eq. (4), if the longitudinal arm of the interferometer itself experiences the so-called *length contraction*:

$$l_l(v) = l_l(0)\left(1 - \frac{v^2}{c^2}\right)^{1/2}, \qquad (9.5)$$

while the transverse arm's length is not affected by its motion through the aether, this effect kills the shift Δt. This radical idea received strong support from the proof, in 1887-1905, that the Maxwell equations, and hence the wave equation (3), are form-invariant under the so-called *Lorentz transform*,[9] which in particular describes Eq. (5). For the choice of coordinates shown in Fig. 1, the transform reads

---

[7] Through the 20th century, the Michelson-Morley-type experiments were repeated using more and more refined experimental techniques, always with zero results for the apparent aether motion speed. For example, recent experiments using cryogenically cooled optical resonators have reduced the upper limit for such speed to just $3\times10^{-15}\ c$ –see H. Müller *et al.*, *Phys. Rev. Lett.* **91**, 020401 (2003).

[8] The zero result of a slightly later experiment, namely a precise measurement of the torque that should be exerted by the moving aether on a charged capacitor, carried out in 1903 by F. Trouton and H. Noble (following G. FitzGerald's suggestion), seconded the Michelson and Morley's conclusions.

[9] The theoretical work toward this result included important contributions by Woldemart Voigt (in 1887), Hendrik Lorentz (in 1892-1904), Joseph Larmor (in 1897 and 1900), and Henri Poincaré (in 1900 and 1905).

$$\boxed{x = \frac{x' + vt'}{\left(1 - v^2/c^2\right)^{1/2}}, \quad y = y', \quad z = z', \quad t = \frac{t' + (v/c^2)x'}{\left(1 - v^2/c^2\right)^{1/2}}.} \qquad (9.6a)$$

Lorentz transform

It is elementary to solve these equations for the primed coordinates to get the reciprocal transform

$$x' = \frac{x - vt}{\left(1 - v^2/c^2\right)^{1/2}}, \quad y' = y, \quad z' = z, \quad t' = \frac{t - (v/c^2)x}{\left(1 - v^2/c^2\right)^{1/2}}. \qquad (9.6b)$$

(I will soon represent Eqs. (6) in a more elegant form – see Eqs. (19) below.)

The Lorentz transform relations (6) are evidently reduced to the Galilean transform formulas (2) at $v^2 << c^2$. However, all attempts to give a reasonable interpretation of these equalities while keeping the notion of the aether have failed, in particular because of the restrictions imposed by results of earlier experiments carried out in 1851 and 1853 by Hippolyte Fizeau – which were repeated with higher accuracy by the same Michelson and Morley in 1886. These experiments have shown that if one sticks to the aether concept, this hypothetical medium has to be partially "dragged" by any moving dielectric material with a speed proportional to ($\kappa - 1$). Such local drag would be irreconcilable with the assumed continuity of the aether.

In his famous 1905 paper, Albert Einstein suggested a bold resolution of this contradiction, essentially removing the concept of the aether altogether.[10] Moreover, he argued that the Lorentz transform is the general property of time and space, rather than of the electromagnetic field alone. He started with two postulates, the first one essentially repeating the relativity principle formulated a bit earlier (in 1904) by H. Poincaré in the following form:

> "*…the laws of physical phenomena should be the same, whether for an observer fixed or for an observer carried along in a uniform movement of translation; so that we have not and could not have any means of discerning whether or not we are carried along in such a motion*."[11]

The second Einstein postulate was that the speed of light c, in free space, should be constant in all reference frames. (This is essentially a denial of the aether's existence.)

Then, Einstein showed that the Lorenz transform relations (6) naturally follow from his postulates, with a few (very natural) additional assumptions. Let a point source emit a short flash of light, at the moment $t = t' = 0$ when the origins of the reference frames shown in Fig. 1 coincide. Then, according to the second of Einstein's postulates, in each of the frames, the spherical wave propagates with the same speed c, i.e. the coordinates of points of its front, measured in the two frames, have to obey the following equalities:

$$\begin{aligned} (ct)^2 - (x^2 + y^2 + z^2) &= 0, \\ (ct')^2 - (x'^2 + y'^2 + z'^2) &= 0. \end{aligned} \qquad (9.7)$$

[10] In hindsight, this was much relief, because the aether had been a very awkward construct to start with. In particular, according to the basic theory of elasticity (see, e.g., CM Ch. 7), in order to carry such transverse waves as the electromagnetic ones, this medium would need to have a non-zero shear modulus, i.e. behave as an elastic solid – rather than as a rarified gas hypothesized initially by C. Huygens.

[11] Note that though the relativity principle excludes the notion of the special ("absolute") spatial reference frame, its quoted verbal formulation still leaves the possibility of the Galilean "absolute time" $t = t'$ open. The quantitative relativity theory kills this option – see Eqs. (6) and their discussion below.

What may be the general relation between the combinations in the left-hand side of these equations – not for this particular wave's front, but in general? A very natural (essentially, the only justifiable) choice is

$$\left[(ct)^2 - (x^2 + y^2 + z^2)\right] = f(v^2)\left[(ct')^2 - (x'^2 + y'^2 + z'^2)\right]. \tag{9.8}$$

Now, according to the first postulate, the same relation should be valid if we swap the reference frames ($x \leftrightarrow x'$, etc.) and replace v with $(-v)$. This is only possible if $f^2 = 1$, so excluding the option $f = -1$ (which is incompatible with the Galilean transform in the limit $v/c \to 0$), we are left with $f = +1$, i.e.

$$(ct)^2 - (x^2 + y^2 + z^2) = (ct')^2 - (x'^2 + y'^2 + z'^2). \tag{9.9}$$

For the line with $y = y' = 0$ and $z = z' = 0$, Eq. (9) is reduced to

$$(ct)^2 - x^2 = (ct')^2 - x'^2. \tag{9.10}$$

It is very illuminating to interpret this relation as the one resulting from a mutual rotation of the reference frames (that now have to include clocks to measure time) on the plane of the coordinate x and the so-called *imaginary time* $\tau \equiv ict$ – see Fig. 3.

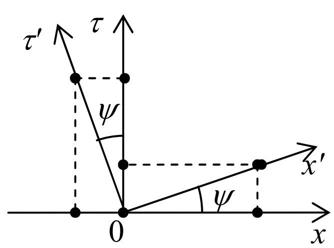


Fig. 9.3. The Lorentz transform as a mutual rotation of two reference frames on the $[x, \tau]$ plane.

Indeed, rewriting Eq. (10) as

$$\tau^2 + x^2 = \tau'^2 + x'^2, \tag{9.11}$$

we may consider it as the invariance of the squared radius at the rotation shown in Fig. 3 and described by the following geometric relations:

$$\begin{aligned} x &= x'\cos\psi - \tau'\sin\psi, \\ \tau &= x'\sin\psi + \tau'\cos\psi, \end{aligned} \tag{9.12a}$$

with the reciprocal relations

$$\begin{aligned} x' &= x\cos\psi + \tau\sin\psi, \\ \tau' &= -x\sin\psi + \tau\cos\psi. \end{aligned} \tag{9.12b}$$

So far, the angle ψ has been arbitrary. In the spirit of Eq. (8), a natural choice is $\psi = \psi(v)$, with the requirement $\psi(0) = 0$. To find this function, let us write the definition of the velocity v of frame 0', as measured in frame 0 (which was implied above): for $x' = 0$, $x = vt$. In the variables x and τ, this means

$$\frac{x}{\tau}\Big|_{x'=0} \equiv \frac{x}{ict}\Big|_{x'=0} = \frac{v}{ic}. \tag{9.13}$$

On the other hand, for the same point $x' = 0$, Eqs. (12a) yield

$$\left.\frac{x}{\tau}\right|_{x'=0} = -\tan\psi . \qquad (9.14)$$

These two expressions are compatible only if

$$\tan\psi = \frac{iv}{c}, \qquad (9.15)$$

so

$$\sin\psi \equiv \frac{\tan\psi}{\left(1+\tan^2\psi\right)^{1/2}} = \frac{iv/c}{\left(1-v^2/c^2\right)^{1/2}} \equiv i\beta\gamma, \qquad \cos\psi \equiv \frac{1}{\left(1+\tan^2\psi\right)^{1/2}} = \frac{1}{\left(1-v^2/c^2\right)^{1/2}} \equiv \gamma, \qquad (9.16)$$

where β and γ are two very convenient and commonly used dimensionless parameters defined as

Relativistic parameters β and γ

$$\boxed{\boldsymbol{\beta} \equiv \frac{\mathbf{v}}{c}, \qquad \gamma \equiv \frac{1}{\left(1-v^2/c^2\right)^{1/2}} \equiv \frac{1}{\left(1-\beta^2\right)^{1/2}} .} \qquad (9.17)$$

(The vector β is called the *normalized velocity*, while the scalar γ is the *Lorentz factor*.)[12]

Using the above relations for ψ, Eqs. (12) become

$$x = \gamma(x' - i\beta\tau'), \qquad \tau = \gamma(i\beta x' + \tau'), \qquad (9.18a)$$

$$x' = \gamma(x + i\beta\tau), \qquad \tau' = \gamma(-i\beta x + \tau). \qquad (9.18b)$$

Now returning to the real variables [x, ct], we get the Lorentz transform relations (6), in a more compact form:

Lorentz transform – again

$$\boxed{\begin{aligned} x &= \gamma(x' + \beta\, ct'), \quad y = y', \quad z = z', \quad ct = \gamma(ct' + \beta\, x'), \\ x' &= \gamma(x - \beta\, ct), \quad y' = y, \quad z' = z, \quad ct' = \gamma(ct - \beta\, x). \end{aligned}} \qquad (9.19a), (9.19b)$$

An immediate corollary of Eqs. (19) is that for γ to stay real, we need $v^2 \le c^2$, i.e. that the speed of any physical body (to which we could connect a meaningful reference frame) cannot exceed the speed of light, as measured in *any* other meaningful reference frame.[13]

9.2. Relativistic kinematic effects

Before proceeding to other corollaries of Eqs. (19), let us spend a few minutes discussing what these relations actually mean. Evidently, they are trying to tell us that the spatial and temporal intervals are not absolute (as they are in the Newtonian space), but do depend on the reference frame they are measured in. So, we have to understand very clearly what exactly may be measured – and thus may be discussed in a meaningful physics theory. Recognizing this necessity, A. Einstein introduced the notion of numerous imaginary *observers* that may be distributed all over each reference frame. Each observer

[12] Note the following identities: $\gamma^2 \equiv 1/(1-\beta^2)$ and $(\gamma^2 - 1) \equiv \beta^2/(1-\beta^2) \equiv \gamma^2\beta^2$, which are frequently handy in relativity-related algebra. One more function of β, the *rapidity* $\varphi \equiv \tanh^{-1}\beta$ (so that $\psi = i\varphi$), is also useful for some calculations.

[13] All attempts to rationally conjecture particles moving with $v > c$ (called *tachyons*) have failed – so far, at least. Possibly the strongest objection against their existence is the fact that the tachyons could be used to communicate back in time, thus violating the causality principle – see, e.g., G. Benford *et al*., *Phys. Rev. D* **2**, 263 (1970).

has a clock and may use it to measure the instants of *local* events, taking place at the observer's location. He also conjectured, very reasonably, that:

(i) all observers within the same reference frame may agree on a common length measure ("a scale"), i.e. on their relative positions in that frame, and synchronize their clocks,[14] and

(ii) the observers belonging to different reference frames may agree on the nomenclature of *world events* (e.g., short flashes of light) to which their respective measurements refer.

Actually, these additional postulates have been already implied in our "derivation" of the Lorentz transform in Sec. 1. For example, by the set $\{x, y, z, t\}$ we mean the results of space and time measurements of a certain world event, about that all observers belonging to frame 0 agree. Similarly, all observers of frame 0' have to agree about the results $\{x', y', z', t'\}$. Finally, when the origin of frame 0' passes by some sequential points x_k of frame 0, the observers in the latter frame may measure its passage times t_k without a fundamental error, and know that all these times belong to $x' = 0$.

Now we can analyze the major corollaries of the Lorentz transform, which are rather striking from the point of view of our everyday (rather non-relativistic) experience.

(i) Length contraction. Let us consider a thin rigid rod oriented along the x-axis, with its length $l \equiv x_2 - x_1$, where $x_{1,2}$ are the coordinates of the rod's ends, as measured in its rest frame 0, at any instant t (Fig. 4). What would be the rod's length l' measured by the Einstein observers in the moving frame 0'?

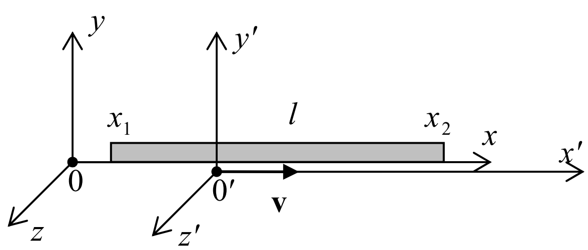


Fig. 9.4. The relativistic length contraction.

At a time instant t' agreed upon in advance, the observers who find themselves exactly at the rod's ends, may register that fact, and then subtract their coordinates $x'_{1,2}$ to calculate the apparent rod length $l' \equiv x_2' - x_1'$ in the moving frame. According to Eq. (19a), l may be expressed via this l' as

$$l \equiv x_2 - x_1 = \gamma(x_2' + \beta ct') - \gamma(x_1' + \beta ct') = \gamma(x_2' - x_1') \equiv \gamma l' . \tag{9.20a}$$

Hence, the rod's length, as measured in the *moving* reference frame is

Length contraction

$$\boxed{l' = \frac{l}{\gamma} = l\left(1 - \frac{v^2}{c^2}\right)^{1/2} \leq l,} \tag{9.20b}$$

in accordance with the FitzGerald-Lorentz hypothesis (5). This is the *relativistic length contraction* effect: an object is always the longest (has the so-called *proper length l*) if measured in its *rest frame*.

[14] A posteriori, the Lorenz transform may be used to show that consensus-creating procedures (such as clock synchronization) are indeed possible. The basic idea of the proof is that since at $v << c$, the relativistic corrections to space and time intervals are of the order of $(v/c)^2$, they have negligible effects on clocks being brought together into the same point for synchronization slowly, with a speed $u << c$. The reader interested in a detailed discussion of this and other fine points of special relativity may be referred to, e.g., either H. Arzeliès, *Relativistic Kinematics*, Pergamon, 1966, or W. Rindler, *Introduction to Special Relativity*, 2nd ed., Oxford U. Press, 1991.

Note that according to Eqs. (19), the length contraction takes place only in the direction of the relative motion of two reference frames. As was noted in Sec. 1, this result immediately explains the zero result of the Michelson-Morley-type experiments, so they give very convincing evidence (if not irrefutable proof) of Eqs. (18)-(19).

(ii) Time dilation. Now let us use Eqs. (19a) to find the time interval Δt, as measured in some reference frame 0, between two world events – say, two ticks of a clock moving with another frame 0' (Fig. 5), i.e. having fixed values of x', y', and z'.

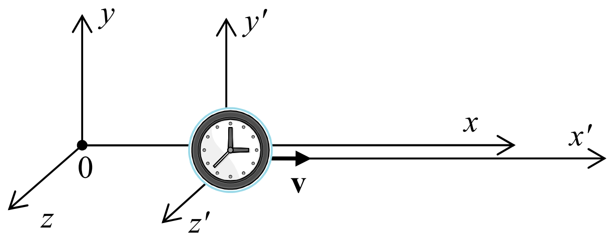


Fig. 9.5. The relativistic time dilation.

Let the time interval between these two events, measured in the clock's rest frame 0', be $\Delta t' \equiv t_2' - t_1'$. At these two moments, the clock would fly by two Einstein's observers at rest in frame 0, so they can record the corresponding moments $t_{1,2}$ shown by their clocks, and then calculate Δt as their difference. According to the last of Eqs. (19a),

$$c\Delta t \equiv ct_2 - ct_1 = \gamma\left[(ct_2' + \beta x') - (ct_1' + \beta x')\right] \equiv \gamma c \Delta t', \tag{9.21a}$$

so, finally,

$$\boxed{\Delta t = \gamma \Delta t' \equiv \frac{\Delta t'}{\left(1 - v^2/c^2\right)^{1/2}} \geq \Delta t'.} \tag{9.21b}$$

Time dilation

This is the famous *relativistic time dilation* (or "dilatation") effect: a time interval is *longer* if measured in a frame (in our case, frame 0) *moving relative to the clock*, while that in the clock's rest frame is the shortest possible – the so-called *proper time interval*.

This rather counter-intuitive effect is the everyday reality in experiments with high-energy elementary particles. For example, in a typical (and by no means record-breaking) experiment carried out in Fermilab, a beam of charged 200 GeV pions with $\gamma \approx$ 1,400 traveled a distance of l = 300 m with the measured loss of only 3% of the initial beam intensity due to the pion decay (mostly, into muon-neutrino pairs) with the proper lifetime $t_0 \approx 2.56\times10^{-8}$ s. Without the time dilation, only an $\exp\{-l/ct_0\}$ $\sim10^{-17}$ fraction of the initial pions would survive, while the relativity-corrected number, $\exp\{-l/ct\} = \exp\{-l/c\gamma t_0\} \approx 0.97$, was in full accordance with experimental measurements.

As another example, the global positioning systems (say, the GPS) are designed with the account of the time dilation due to the velocity of their satellites (and also some gravity-induced, i.e. general-relativity corrections, which I would not have time to discuss) and would give large errors without such corrections. So, there is no doubt that time dilation (21) is a reality, though the precision of its experimental tests I am aware of[15] has been limited to a few percent, because of the almost unavoidable involvement of less controllable gravity effects – which provide a time interval change of the opposite sign in most experiments near the Earth's surface.

[15] See, e.g., J. Hafele and R. Keating, *Science* **177**, 166 (1972).

Before the first reliable observation of time dilation (by B. Rossi and D. Hall in 1940), there had been serious doubts about the reality of this effect, the most famous being the *twin paradox* first posed (together with an immediate suggestion of its resolution) by P. Langevin in 1911. Let us send one of two twins on a long space roundtrip with the maximum speed approaching c. Upon his return to Earth, who of the twins would be older? The naïve approach is to say that due to the relativity principle, not one can be (and hence there is no time dilation) because each twin could claim that their counterpart rather than them, was moving, with the same speed but in the opposite direction. The resolution of the paradox is that one of the twins had to be accelerated to be brought back, and hence the reference frames have to be dissimilar: only one of them may stay inertial all the time. As a result, the twin who had been accelerated ("actually traveling") would be younger than their sibling when they finally came together. Constructive proof of this conclusion for the particular case of straight-line travel with a piecewise-constant acceleration, is simple and hence left for the reader's exercise.

(iii) Velocity transformation. Now let us calculate the velocity $\mathbf{u}$ of a moving point, as observed in reference frame 0, provided that its velocity, as measured in frame 0', is $\mathbf{u}'$ (Fig. 6).

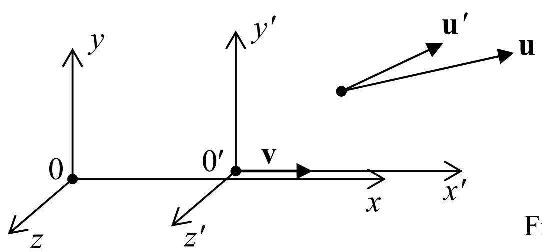


Fig. 9.6. The relativistic velocity addition.

Keeping the usual definition of velocity, but with due attention to the relativity of not only spatial but also temporal intervals, we may write

$$\mathbf{u} \equiv \frac{d\mathbf{r}}{dt}, \qquad \mathbf{u}' \equiv \frac{d\mathbf{r}'}{dt'}. \tag{9.22}$$

Plugging in the differentials of the Lorentz transform relations (6a) into these definitions, we get

$$u_x \equiv \frac{dx}{dt} = \frac{dx' + vdt'}{dt' + vdx'/c^2} \equiv \frac{u'_x + v}{1 + u'_x v/c^2}, \qquad u_y \equiv \frac{dy}{dt} = \frac{1}{\gamma}\frac{dy'}{dt' + vdx'/c^2} \equiv \frac{1}{\gamma}\frac{u'_y}{1 + u'_x v/c^2}, \tag{9.23}$$

with a similar formula for u_z. In the classical limit $v/c \to 0$, these relations are reduced to

$$u_x = u'_x + v, \qquad u_y = u'_y, \qquad u_z = u'_z, \tag{9.24a}$$

and may be merged into the familiar Galilean form

$$\mathbf{u} = \mathbf{u}' + \mathbf{v}, \qquad \text{for } v << c. \tag{9.24b}$$

In order to see how unusual the full relativistic rules (23) are at $u \sim c$, let us first consider a purely longitudinal motion, $u_y = u_z = 0$; then[16]

[16] With an account of the identity $\tanh(a + b) = (\tanh a + \tanh b)/(1 + \tanh a \tanh b)$, which readily follows from MA Eq. (3.5), Eq. (25) shows that rapidities $\varphi \equiv \tanh^{-1}\beta$ add up exactly as longitudinal velocities at non-relativistic motion, making that notion very convenient for the analysis of transfer between several frames.

$$u = \frac{u' + v}{1 + u'v/c^2},\qquad (9.25)$$

Longitudinal velocity addition

where $u \equiv u_x$ and $u' \equiv u'_x$. Figure 7 shows this u as the function of u', for several values of the reference frames' relative velocity v.

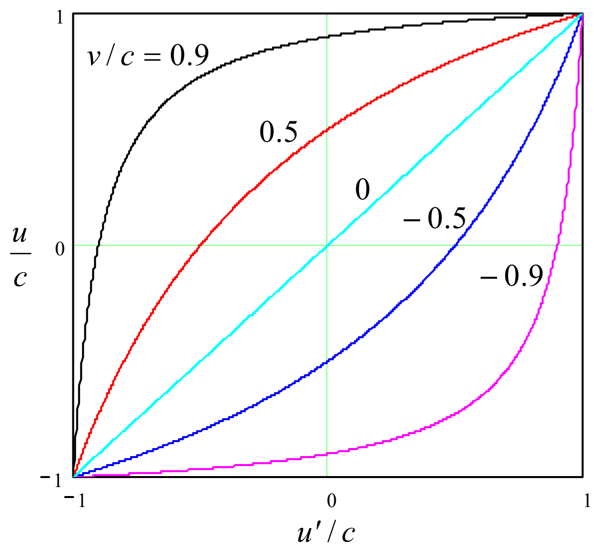


Fig. 9.7. The addition of longitudinal velocities.

The first sanity check is that if $v = 0$, i.e. if the reference frames are at rest relative to each other, then $u = u'$, as it should be – see the diagonal straight line in Fig. 7. Next, if magnitudes of u' and v are both below c, so is the magnitude of u. (Also good, because otherwise, ordinary particles in one frame would be tachyons in the other one, and the theory would be in big trouble.) Now strange things begin: even as u' and v are both approaching c, then u is also close to c, but does not exceed it. As an example, if we fired forward a bullet with the relative speed of $0.9c$, from a spaceship moving from the Earth also at $0.9c$, Eq. (25) predicts the speed of the bullet relative to the Earth to be just $[(0.9 + 0.9)/(1 + 0.9\times0.9)]c \approx 0.994c < c$, rather than $(0.9 + 0.9)\, c = 1.8\, c > c$ as in the Galilean kinematics. Actually, we could expect this strangeness, because it is necessary to fulfill the 2nd Einstein's postulate: the independence of the speed of light in any reference frame. Indeed, for $u' = \pm c$, Eq. (25) yields $u = \pm c$, regardless of v.

In the opposite case of a purely transverse motion, when a point moves across the relative motion of the frames (for example, at our choice of coordinates, $u'_x = u'_z = 0$), Eqs. (23) yield a much less spectacular result

$$u_y = \frac{1}{\gamma} u'_y \leq u'_y. \qquad (9.26)$$

This effect comes purely from the time dilation because the transverse spatial intervals are Lorentz-invariant.

In the case when both u_x' and u_y' are substantial (but u_z' is still zero), we may divide Eqs. (23) by each other to relate the angles θ of the point's propagation, as observed in the two reference frames:

$$\tan\theta \equiv \frac{u_y}{u_x} = \frac{u'_y}{\gamma(u'_x + v)} = \frac{\sin\theta'}{\gamma(\cos\theta' + v/u')}. \qquad (9.27)$$

Stellar aberration effect

This expression describes, in particular, the so-called *stellar aberration* effect: the dependence of the observed direction θ toward a star on the speed v of the telescope's motion relative to the star – see Fig. 8. (The effect is readily observable experimentally as the *annual aberration* due to the periodic change of speed v by $2v_E \approx 60$ km/s because of the Earth's rotation about the Sun. Since the aberration's main part is of the first order in $v_E/c \sim 10^{-4}$, this effect is very significant and has been known since the early 1700s.)

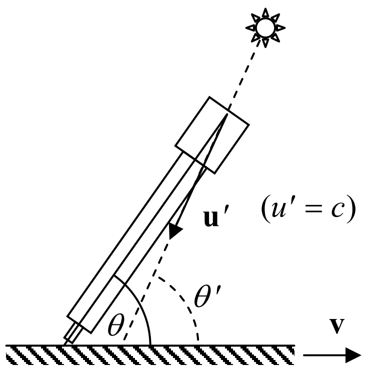


Fig. 9.8. The stellar aberration.

For the analysis of this effect, it is sufficient to take, in Eq. (27), $u' = c$, i.e. $v/u' = \beta$, and interpret θ' as the "proper" direction to the star, which would be measured at $v = 0$.[17] At $\beta << 1$, both Eq. (27) and the Galilean result (which the reader is invited to derive directly from Fig. 8),

$$\tan\theta = \frac{\sin\theta'}{\cos\theta' + \beta}, \tag{9.28}$$

may be well approximated by the first-order term

$$\Delta\theta \equiv \theta - \theta' \approx -\beta\sin\theta'. \tag{9.29}$$

Unfortunately, it is not easy to use the difference between Eqs. (27) and (28), of the second order in β, for special relativity's confirmation, because other components of the Earth's motion, such as its rotation, nutation, and torque-induced precession,[18] give masking first-order contributions to the aberration.

Finally, for a completely arbitrary direction of the vector $\mathbf{u}'$, Eqs. (22) may be readily used to calculate the velocity's magnitude. The most popular form of the resulting expression is the following expression for the square of the relative velocity (or rather the reduced relative velocity $\boldsymbol{\beta}$) of two points,

$$\beta^2 = \frac{(\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2)^2 - |\boldsymbol{\beta}_1 \cdot \boldsymbol{\beta}_2|}{(1 - \boldsymbol{\beta}_1 \cdot \boldsymbol{\beta}_2)^2} \leq 1. \tag{9.30}$$

where $\boldsymbol{\beta}_{1,2} \equiv \mathbf{v}_{1,2}/c$ are their normalized velocities as measured in the same reference frame.

[17] Strictly speaking, to reconcile the geometries shown in Fig. 1 (for which all our formulas, including Eq. (27), are valid) and Fig. 8 (giving the traditional scheme of the stellar aberration), it is necessary to invert the signs of $\mathbf{u}$ (and hence of $\sin\theta'$ and $\cos\theta'$) and $\mathbf{v}$, but as it is evident from Eq. (27), all the minus signs cancel, and the formula is valid "as is".

[18] See, e.g., CM Secs. 4.4-4.5.

(iv) The Doppler effect. Let us consider a monochromatic plane wave of some physical nature, traveling along the x-axis:

$$f = \mathrm{Re}\left[f_\omega \exp\{i(kx - \omega t\}\right] \equiv |f_\omega| \cos(kx - \omega t + \arg f_\omega) \equiv |f_\omega| \cos \Psi . \tag{9.31}$$

Its total phase, $\Psi \equiv kx - \omega t + \arg f_\omega$ (in contrast to its amplitude $|f_\omega|$ – see Sec. 5 below) cannot depend on the observer's reference frame, because the variable f vanishes completely at $\Psi = \pi(n + ½)$ (for all integer n), and such "world events" should be observable in all reference frames. The only way to keep $\Psi = \Psi'$ at all times is to have[19]

$$kx - \omega t = k'x' - \omega' t' . \tag{9.32}$$

First, let us use this general relation to consider the Doppler effect in the usual non-relativistic mechanical waves, e.g., oscillations of particles of a certain medium. Using the Galilean transform (2), we may rewrite Eq. (32) as

$$k(x' + vt) - \omega t = k'x' - \omega' t . \tag{9.33}$$

Since this transform leaves all space intervals (including the wavelength $\lambda = 2\pi/k$) intact, we can take $k = k'$, so Eq. (33) yields

$$\omega' = \omega - kv . \tag{9.34}$$

For a dispersion-free medium, the wave number k is the ratio of its frequency ω, as measured in the reference frame bound to the medium, and the wave velocity v_w. In particular, if the wave source rests in the medium, we may bind the reference frame 0 to the medium as well, and frame 0' to the wave's receiver (i.e. $v = v_r$), so

$$k = \frac{\omega}{v_w} , \tag{9.35}$$

and for the frequency perceived by the receiver, Eq. (34) yields

$$\omega' = \omega \frac{v_w - v_r}{v_w} . \tag{9.36}$$

On the other hand, if the receiver and the medium are at rest in the reference frame 0', while the wave source is bound to the frame 0 (so $v = -v_s$), Eq. (35) should be replaced with

$$k = k' = \frac{\omega'}{v_w} , \tag{9.37}$$

and Eq. (34) yields a different result:

$$\omega' = \omega \frac{v_w}{v_w - v_s} , \tag{9.38}$$

Finally, if both the source and detector are moving, it is straightforward to combine these two results to get the general relation

$$\omega' = \omega \frac{v_w - v_r}{v_w - v_s} . \tag{9.39}$$

[19] Strictly speaking, Eq. (32) is valid to an additive constant, but for notation simplicity, it may be always made equal to zero by selecting (as has already been done in all relations of Sec. 1) the reference frame origins and/or clock turn-on times so that at $t = 0$ and $x = 0$, $t' = 0$ and $x' = 0$ as well.

At low speeds of both the source and the receiver, this result simplifies,

$$\omega' \approx \omega(1-\beta), \qquad \text{with } \beta \equiv \frac{v_\text{r} - v_\text{s}}{v_\text{w}}, \tag{9.40}$$

but at speeds comparable to v_w we have to use the more general Eq. (39). Thus, the usual Doppler effect is generally affected not only by the relative speed ($v_\text{r} - v_\text{s}$) of the wave's source and detector but also by their speeds relative to the medium in which the waves propagate.

Somewhat counter-intuitively, for the electromagnetic waves the calculations are *simpler* because for them the propagation medium (aether) does not exist, the wave velocity equals $\pm c$ in any reference frame, and there are no two separate cases: we can always take $k = \pm\omega/c$ and $k' = \pm\omega'/c$. Plugging these relations, together with the Lorentz transform (19a), into the phase-invariance condition (32), we get

$$\pm\frac{\omega}{c}\gamma(x' + \beta ct') - \omega\gamma\frac{ct' + \beta x'}{c} = \pm\frac{\omega'}{c}x' - \omega't'. \tag{9.41}$$

This relation has to hold for any x' and t', so we may require that the net coefficients before these variables vanish. These two requirements yield the same equality:

$$\omega' = \omega\gamma(1 \mp \beta). \tag{9.42}$$

This result is already quite simple, but may be transformed further to be even more illuminating:

$$\omega' = \omega\frac{1 \mp \beta}{\left(1-\beta^2\right)^{1/2}} \equiv \omega\left[\frac{(1\mp\beta)(1\mp\beta)}{(1+\beta)(1-\beta)}\right]^{1/2}. \tag{9.43}$$

At any sign before β, one pair of parentheses cancels, so[20]

Longitudinal Doppler effect

$$\boxed{\omega' = \omega\left(\frac{1\mp\beta}{1\pm\beta}\right)^{1/2}.} \tag{9.44}$$

Thus the Doppler effect for electromagnetic waves depends only on the relative velocity $v = \beta c$ between the wave source and detector – as it should be, given the aether's absence. At velocities much lower than c, Eq. (44) may be approximated as

$$\omega' \approx \omega\frac{1\mp\beta/2}{1\pm\beta/2} \approx \omega(1\mp\beta), \tag{9.45}$$

i.e. in the first approximation in $\beta \equiv v/c$, it tends to the corresponding limit (40) of the usual Doppler effect.

If the wave vector **k** is tilted by angle θ to the vector **v** (as measured in frame 0), then we have to repeat the calculations, with k replaced by k_x, and components k_y and k_z left intact at the Lorentz transform. As a result, Eq. (42) is generalized as

[20] It may look like the reciprocal expression of ω via ω' is different, violating the relativity principle. However, in this case, we have to change the sign of β, because the relative velocity of the system is opposite, so we return to Eq. (44) again.

$$\omega' = \omega\gamma(1 - \beta\cos\theta). \tag{9.46}$$

For the case $\cos\theta = \pm 1$, Eq. (46) reduces to our previous result (42). However, at $\theta = \pi/2$ (i.e. $\cos\theta = 0$), the relation is rather different:

Transverse Doppler effect

$$\boxed{\omega' = \gamma\omega \equiv \frac{\omega}{\left(1-\beta^2\right)^{1/2}}.} \tag{9.47}$$

This is the *transverse Doppler effect* – which is absent in non-relativistic physics. Its first experimental evidence was obtained using electron beams (as had been suggested in 1906 by J. Stark), by H. Ives and G. Stilwell in 1938 and 1941. Later, similar experiments were repeated several times, but the first unambiguous measurements were performed only in 1979 by D. Hasselkamp *et al.* who confirmed Eq. (47) with a relative accuracy of about 10%. This precision may not look too spectacular, but besides the special tests discussed above, the Lorentz transform formulas have been also confirmed, less directly, by a huge body of other experimental data, especially in high energy physics, agreeing with calculations incorporating this transform as their part. This is why, with due respect to the spirit of challenging authority, I should warn the reader: if you decide to challenge the relativity theory (called "theory" by tradition only), you would also need to explain all these data. Best luck with that! [21]

9.3. 4-vectors, momentum, mass, and energy

Before proceeding to the relativistic dynamics, let us discuss the mathematical formalism that makes all calculations more compact – and more beautiful. We have already seen that the three spatial coordinates $\{x, y, z\}$ and the product ct are Lorentz-transformed similarly – see Eqs. (18)-(19) again. So it is natural to consider them as components of a single four-component vector (or, for short, *4-vector*),

$$\{x_0, x_1, x_2, x_3\} \equiv \{ct, \mathbf{r}\}, \tag{9.48}$$

with components

Space-time 4-vector

$$\boxed{x_0 \equiv ct, \quad x_1 \equiv x, \quad x_2 \equiv y, \quad x_3 \equiv z.} \tag{9.49}$$

According to Eqs. (19), its components are Lorentz-transformed as

Lorentz transform: 4-form

$$\boxed{x_j = \sum_{j'=0}^{3} L_{jj'} x'_{j'},} \tag{9.50}$$

where $L_{jj'}$ are the elements of the following 4×4 *Lorentz transform matrix*

Lorentz transform matrix

$$\boxed{\begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.} \tag{9.51}$$

Since such 4-vectors are a new notion for this course and will be used for many more purposes than just the space-time transform, we need to discuss the general mathematical rules they obey. Indeed,

[21] The same fact, ignored by crackpots, is also valid for other favorite directions of their attacks, including the Universe expansion, quantum measurement uncertainty, and entropy growth in physics, and the evolution theory in biology.

as was already mentioned in Sec. 8.9, the usual (three-component) vector is not just any ordered set (*string*) of three scalars $\{A_x, A_y, A_z\}$; if we want it to represent a reference-frame-independent physical reality, the vector's components have to obey certain rules at the transfer from one reference frame to another. In particular, in the non-relativistic limit the vector's *norm* (its magnitude squared),

$$A^2 = A_x^2 + A_y^2 + A_z^2, \tag{9.52}$$

should be invariant with respect to the transfer between different reference frames. However, a naïve extension of this approach to 4-vectors would not work, because, according to the calculations of Sec. 1, the Lorentz transform keeps intact the combinations of the type (7), with one sign negative, rather than the sum of all components squared. Hence for the 4-vectors, all the rules of the game have to be reviewed and adjusted – or rather redefined from the very beginning, for example as follows.[22]

An arbitrary 4-vector is a string of 4 scalars,[23]

General 4-vector

$$\boxed{\{A_0, A_1, A_2, A_3\},} \tag{9.53}$$

whose components A_j, as measured in the reference frames 0 and 0' shown in Fig. 1, obey the Lorentz transform relations similar to Eq. (50):

Lorentz transform: general 4-vector

$$\boxed{A_j = \sum_{j'=0}^{3} L_{jj'} A'_{j'} .} \tag{9.54}$$

As we have already seen in the example of the space-time 4-vector (48), this means in particular that

Lorentz invariance

$$\boxed{A_0^2 - \sum_{j=1}^{3} A_j^2 = (A'_0)^2 - \sum_{j=1}^{3} (A'_j)^2 .} \tag{9.55}$$

This is the so-called *Lorentz invariance* condition for the *4-vector's norm*. (The difference between this relation and Eq. (52), pertaining to Euclidian geometry, is the reason why the Minkowski space is called *pseudo-Euclidian*.) It is also straightforward to use Eqs. (51) and (54) to check that the evident generalization of the norm, the *scalar product* of two arbitrary 4-vectors,

Scalar 4-product

$$\boxed{A_0 B_0 - \sum_{j=1}^{3} A_j B_j ,} \tag{9.56}$$

is also Lorentz-invariant.

Now consider the 4-vector corresponding to a small *interval* between two close world events:

$$\{dx_0, dx_1, dx_2, dx_3\} = \{cdt, d\mathbf{r}\}; \tag{9.57}$$

its norm,

Interval

$$\boxed{(ds)^2 \equiv dx_0^2 - \sum_{j=1}^{3} dx_j^2 = c^2 (dt)^2 - (dr)^2 ,} \tag{9.58}$$

[22] The most prominent alternative, which has both advantages and drawbacks, is to use 4-vectors with one imaginary component – for example, the imaginary time *ict* instead of the real product *ct* in Eq. (48).

[23] Such vectors are said to reside in so-called 4D *Minkowski spaces* – called after Hermann Minkowski who was the first one to recast (in 1907) the special relativity relations in a form in which the spatial coordinates and time (or rather *ct*) are treated on an equal footing.

is of course also Lorentz-invariant. Since the speed of any particle (or signal) cannot be larger than c, for any pair of world events that are in a causal relation with each other, $(dr)^2$ cannot be larger than $(cdt)^2$, i.e. such *time-like* interval $(ds)^2$ cannot be negative. The 4D surface separating such intervals from *space-like* intervals $(ds)^2 < 0$ is called the light cone (Fig. 9).

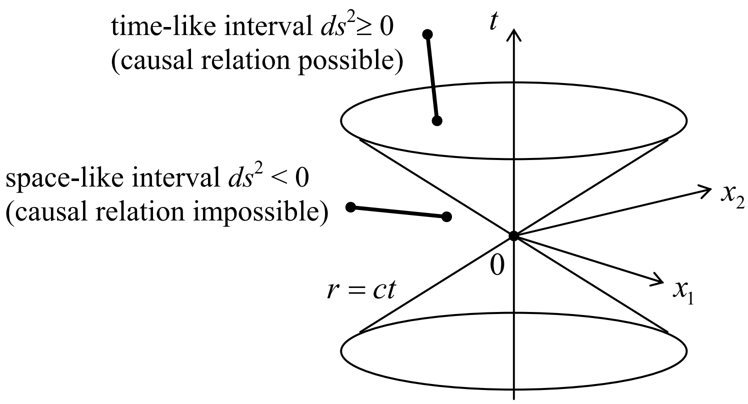


Fig. 9.9. A 2+1 dimensional image of the light cone – which is actually 3+1 dimensional.

Now let us consider two close world events that happen with the same point moving with velocity **u**. Then in the frame moving with the point (**v** = **u**), the last term on the right-hand side of Eq. (58) equals zero, while the involved time is the proper one, so

$$ds = cd\tau \,, \tag{9.59}$$

where $d\tau$ is the proper time interval. But according to Eq. (21), this means that we can write

$$d\tau = \frac{dt}{\gamma}, \tag{9.60}$$

where dt is the time interval in an *arbitrary* (besides being inertial) reference frame, while

$$\boldsymbol{\beta} \equiv \frac{\mathbf{u}}{c} \quad \text{and} \quad \gamma \equiv \frac{1}{\left(1-\beta^2\right)^{1/2}} = \frac{1}{\left(1-u^2/c^2\right)^{1/2}} \tag{9.61}$$

are the parameters (17) corresponding to the *point's* velocity (**u**) in that frame, so $ds = cdt/\gamma$.[24]

Let us use Eq. (60) to explore whether a 4-vector may be formed using the spatial Cartesian components of the point's velocity

$$\mathbf{u} = \left\{\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right\}. \tag{9.62}$$

Here we have a problem: per Eqs. (22), these components do not obey the Lorentz transform. However, let us use $d\tau \equiv dt/\gamma$, the proper time interval of the point, to form the following string:

$$\boxed{\left\{\frac{dx_0}{d\tau}, \frac{dx_1}{d\tau}, \frac{dx_2}{d\tau}, \frac{dx_3}{d\tau}\right\} \equiv \gamma\left\{c, \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right\} \equiv \gamma\{c, \mathbf{u}\}.} \tag{9.63}$$

4-velocity

[24] I have opted against using special indices (e.g., $\boldsymbol{\beta}_u$ and γ_u) to distinguish Eqs. (17) and (61) here and below, in a hope that the suitable velocity (of either a reference frame or a particle) will be always clear from the context.

As it follows from the comparison of the middle form of this expression with Eq. (48), since the time-space vector obeys the Lorentz transform, and τ is Lorentz-invariant, the string (63) is a legitimate 4-vector; it is called the *4-velocity* of a point – or of a point particle.

Now we are well equipped to proceed to relativistic dynamics. Let us start with such basic notions as the momentum **p** and the energy $\mathscr{E}$ – so far, for a free particle.[25] Perhaps the most elegant way to "derive" (or rather guess[26]) the expressions for **p** and $\mathscr{E}$ as functions of the particle's velocity **u**, is based on analytical mechanics. Due to the conservation of **v**, the trajectory of a free particle in the 4D Minkowski space $\{ct, \mathbf{r}\}$ is always a straight line. Hence, from the Hamilton principle,[27] we may expect its action $\mathscr{S}$, between points 1 and 2, to be a linear function of the space-time interval (59):

Free particle: action

$$\boxed{\mathscr{S} = \alpha \int_1^2 ds \equiv \alpha c \int_1^2 d\tau \equiv \alpha c \int_{t_1}^{t_2} \frac{dt}{\gamma},} \qquad (9.64)$$

where α is some constant. On the other hand, in analytical mechanics, the action is defined as

$$\mathscr{S} \equiv \int_{t_1}^{t_2} \mathscr{L} dt, \qquad (9.65)$$

where $\mathscr{L}$ is the particle's Lagrangian function.[28] Comparing these two expressions, we get

$$\mathscr{L} = \frac{\alpha c}{\gamma} \equiv \alpha c \left(1 - \frac{u^2}{c^2}\right)^{1/2}. \qquad (9.66)$$

In the non-relativistic limit ($u << c$), this function tends to

$$\mathscr{L} \approx \alpha c \left(1 - \frac{u^2}{2c^2}\right) \equiv \alpha c - \frac{\alpha u^2}{2c}. \qquad (9.67)$$

In order to correspond to the Newtonian mechanics,[29] the last (velocity-dependent) term should equal $mu^2/2$. From here we find $\alpha = -mc$, so, finally,

Free particle: Lagrangian function

$$\boxed{\mathscr{L} = -mc^2 \left(1 - \frac{u^2}{c^2}\right)^{1/2} \equiv -\frac{mc^2}{\gamma}.} \qquad (9.68)$$

Now we can find the Cartesian components p_j of the particle's momentum as the generalized momenta corresponding to the corresponding components r_j ($j = 1, 2, 3$) of the 3D radius-vector **r**:[30]

---

[25] I am sorry for using, just as in Sec. 6.3, the same traditional notation (**p**) for the particle's momentum as had been used earlier for the electric dipole moment. However, since the latter notion will be virtually unused in the balance of this course, this may hardly lead to confusion.

[26] Indeed, such a derivation uses additional assumptions, however natural (such as the Lorentz-invariance of $\mathscr{S}$), i.e. it can hardly be considered as a real proof of the final results, so they require experimental confirmation. Fortunately, such confirmations have been numerous – see below.

[27] See, e.g., CM Sec. 10.3.

[28] See, e.g., CM Sec. 2.1.

[29] See, e.g., CM Eq. (2.19b).

[30] See, e.g., CM Sec. 2.3, in particular Eq. (2.31).

$$p_j = \frac{\partial \mathscr{L}}{\partial \dot{r}_j} \equiv \frac{\partial \mathscr{L}}{\partial u_j} = -mc^2 \frac{\partial}{\partial u_j}\left(1 - \frac{u_1^2 + u_2^2 + u_3^2}{c^2}\right)^{1/2} = \frac{mu_j}{\left(1 - u^2/c^2\right)^{1/2}} \equiv m\gamma u_j. \tag{9.69}$$

Thus for the 3D vector of momentum, we can write the result in the same form as in non-relativistic mechanics,

Relativistic momentum

$$\boxed{\mathbf{p} = m\gamma\mathbf{u} \equiv M\mathbf{u},} \tag{9.70}$$

using the reference-frame-dependent scalar M (called the *relativistic mass*) defined as

Relativistic mass

$$\boxed{M \equiv m\gamma = \frac{m}{\left(1 - u^2/c^2\right)^{1/2}} \geq m,} \tag{9.71}$$

m being the non-relativistic mass of the particle. (More often, m is called the *rest mass*, because in the reference frame in which the particle rests, Eq. (71) yields $M = m$.)

Next, let us return to analytical mechanics to calculate the particle's energy $\mathscr{E}$ (which for a free particle coincides with its Hamiltonian function $\mathscr{H}$):[31]

$$\mathscr{E} = \mathscr{H} \equiv \sum_{j=1}^{3} p_j u_j - \mathscr{L} = \mathbf{p}\cdot\mathbf{u} - \mathscr{L} = \frac{mu^2}{\left(1 - u^2/c^2\right)^{1/2}} + mc^2\left(1 - \frac{u^2}{c^2}\right)^{1/2} \equiv \frac{mc^2}{\left(1 - u^2/c^2\right)^{1/2}}. \tag{9.72}$$

Thus, we have arrived at the most famous of Einstein's formulas – and probably of physics as a whole:

$\mathscr{E} = Mc^2$

$$\boxed{\mathscr{E} = m\gamma c^2 \equiv Mc^2,} \tag{9.73}$$

which expresses the relation between the free particle's mass and its energy.[32] In the non-relativistic limit, it reduces to

$$\mathscr{E} = \frac{mc^2}{\left(1 - u^2/c^2\right)^{1/2}} \approx mc^2\left(1 + \frac{u^2}{2c^2}\right) = mc^2 + \frac{mu^2}{2}, \tag{9.74}$$

the first term mc^2 being called the *rest energy* of a particle.

Now let us consider the following string of 4 scalars:

4-vector of energy-momentum

$$\boxed{\left\{\frac{\mathscr{E}}{c}, p_1, p_2, p_3\right\} \equiv \left\{\frac{\mathscr{E}}{c}, \mathbf{p}\right\}.} \tag{9.75}$$

Using Eqs. (70) and (73) to represent this expression as

$$\left\{\frac{\mathscr{E}}{c}, \mathbf{p}\right\} = m\gamma\{c, \mathbf{u}\}, \tag{9.76}$$

[31] See, e.g., CM Eq. (2.32).

[32] Let me hope that the reader understands that all the layman talk about the "mass to energy conversion" is only valid in a very limited sense of the word. While the Einstein relation (73) does allow the conversion of "massive" particles (with $m \neq 0$) into particles with $m = 0$, such as photons, each of the latter particles also has a non-zero relativistic mass M, and *simultaneously* the energy $\mathscr{E}$ related to this M by Eq. (73).

and comparing the result with Eq. (63), we immediately see that, since m is a Lorentz-invariant constant, this string is a legitimate 4-vector of *energy-momentum*. As a result, its norm,

$$\left(\frac{\mathscr{E}}{c}\right)^2 - p^2, \tag{9.77a}$$

is Lorentz-invariant, and in particular, has to be equal to the norm in the particle-bound frame. But in that frame, $p = 0$, and according to Eq. (73), $\mathscr{E} = mc^2$, and the norm is just

$$\left(\frac{\mathscr{E}}{c}\right)^2 = \left(\frac{mc^2}{c}\right)^2 \equiv (mc)^2, \tag{9.77b}$$

so in an arbitrary frame

$$\left(\frac{\mathscr{E}}{c}\right)^2 - p^2 = (mc)^2. \tag{9.78a}$$

This very important relation[33] between the relativistic energy and momentum (valid for free particles only!) is usually represented in the form[34]

Free particle: energy

$$\boxed{\mathscr{E}^2 = (mc^2)^2 + (pc)^2.} \tag{9.78b}$$

According to Eq. (70), in the so-called *ultra-relativistic limit* $u \to c$, p tends to infinity, while mc^2 stays constant, so $pc/mc^2 \to \infty$. As follows from Eq. (78), in this limit $\mathscr{E} \approx pc$. Though the above discussion was for particles with finite m, the 4-vector formalism allows us to consider compact objects with zero rest mass as ultra-relativistic particles for which the above energy-to-moment relation,

$$\mathscr{E} = pc, \qquad \text{for } m = 0, \tag{9.79}$$

is exact. Quantum electrodynamics[35] tells us that under certain conditions, the electromagnetic field quanta (photons) may be also considered as such *massless particles* with momentum $\mathbf{p} = \hbar\mathbf{k}$. Plugging (the modulus of) the last relation into Eq. (78), for the photon's energy we get $\mathscr{E} = pc = \hbar kc = \hbar\omega$. Please note again that according to Eq. (73), the relativistic mass of a photon is not equal to zero: $M = \mathscr{E}/c^2 = \hbar\omega/c^2$, so the term "massless particle" has a limited meaning: $m = 0$. For example, the relativistic mass of an optical phonon is of the order of 10^{-36} kg. On the human scale, this is not too much, but still, a noticeable (approximately one-millionth) part of the rest mass m_e of an electron.

The fundamental relations (70) and (73) have been repeatedly verified in numerous particle collision experiments, in which the total energy and momentum of a system of particles are conserved – at the same conditions as in non-relativistic dynamics. (For the momentum, this is the absence of external forces, and for the energy, the elasticity of particle interactions – in other words, the absence of alternative channels of energy escape.) Of course, generally only the total energy of the system is conserved, including the potential energy of particle interactions. However, at typical high-energy

[33] Please note one more simple and useful relation following from Eqs. (70) and (73): $\mathbf{p} = (\mathscr{E}/c^2)\mathbf{u}$.

[34] It may be tempting to interpret this relation as the perpendicular-vector-like addition of the rest energy mc^2 and the "kinetic energy" pc, but from the point of view of the total energy conservation (see below), a better definition of the kinetic energy is $T(u) \equiv \mathscr{E}(u) - \mathscr{E}(0)$.

[35] It is briefly reviewed in QM Chapter 9.

particle collisions, the potential energy vanishes so rapidly with the distance between them that we can use the momentum and energy conservation laws using Eq. (73).

As an example, let us calculate the minimum energy $\mathscr{E}_{\min}$ of a proton (p_a), necessary for the well-known high-energy reaction that generates a new proton-antiproton pair, $p_a + p_b \to p + p + p + \bar{p}$, provided that before the collision, proton p_b had been at rest in the lab frame. This minimum corresponds to the vanishing relative velocity of the reaction products, i.e. their motion with virtually the same velocity ($\mathbf{u}_{\text{fin}}$), as seen from the lab frame – see Fig. 10.

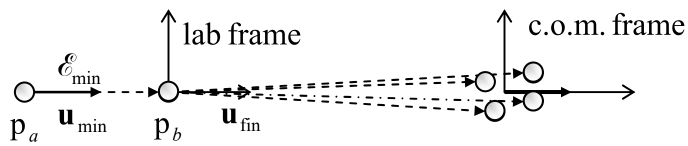


Fig. 9.10. A high-energy proton reaction at $\mathscr{E} \approx \mathscr{E}_{\min}$ – schematically.

Due to the momentum conservation, this velocity should have the same direction as the initial velocity ($\mathbf{u}_{\min}$) of proton p_a. This is why two scalar equations: for energy conservation,

$$\frac{mc^2}{\left(1-u_{\min}^2/c^2\right)^{1/2}} + mc^2 = \frac{4mc^2}{\left(1-u_{\text{fin}}^2/c^2\right)^{1/2}}, \tag{9.80a}$$

and for momentum conservation,

$$\frac{mu}{\left(1-u_{\min}^2/c^2\right)^{1/2}} + 0 = \frac{4mu_{\text{fin}}}{\left(1-u_{\text{fin}}^2/c^2\right)^{1/2}}, \tag{9.80b}$$

are sufficient to find both $u_{\min}$ and u_{fin}. After a rather tedious solution of this system of two nonlinear equations, we get

$$u_{\min} = \frac{4\sqrt{3}}{7}c \approx 0.990\,c, \qquad u_{\text{fin}} = \frac{\sqrt{3}}{2}c \approx 0.866\,c. \tag{9.81}$$

Finally, we can use Eq. (72) to calculate the required energy; the result is $\mathscr{E}_{\min} = 7\,mc^2$. (Note that at this threshold, only a minor $2mc^2$ part of the kinetic energy $T_{\min} = \mathscr{E}_{\min} - mc^2 = 6mc^2$ of the initially moving particle, goes into the "useful" proton-antiproton pair production.) The proton's rest mass, $m_p \approx 1.67\times10^{-27}$ kg, corresponds to $m_pc^2 \approx 1.502\times10^{-10}$ J ≈ 0.938 GeV, so $\mathscr{E}_{\min} \approx 6.57$ GeV.

The second, more intelligent way to solve the same problem is to use the center-of-mass (*c.o.m.*) reference frame that, in relativity, is defined as the frame in which the total momentum of the system vanishes.[36] In this frame, at $\mathscr{E} = \mathscr{E}_{\min}$, the velocity and momenta of all reaction products are vanishing, while the velocities of the protons p_a and p_b before the collision are equal and opposite, with an initially unknown magnitude u'. Hence the energy conservation law becomes

$$\frac{2mc^2}{\left(1-u'^2/c^2\right)^{1/2}} = 4mc^2, \tag{9.82}$$

[36] Note that according to this definition, the c.o.m.'s radius-vector is $\mathbf{R} = \Sigma_k M_k \mathbf{r}_k / \Sigma_k M_k \equiv \Sigma_k \gamma_k m_k \mathbf{r}_k / \Sigma_k \gamma_k m_k$, i.e. is generally different from the well-known non-relativistic expression $\mathbf{R} = \Sigma_k m_k \mathbf{r}_k / \Sigma_k m_k$.

readily giving $u'/c = \sqrt{3}/2$. (This is of course the same result as Eq. (81) gives for u_{fin}.) Now we can use the fact that the velocity of the proton p_a in the c.o.m. frame is $(-u')$, to find its lab-frame speed, using the velocity transform (25):

$$u_{\min} = \frac{2u'}{1+u'^2/c^2}. \tag{9.83}$$

With the above result for u', this relation gives the same result as the first method, $u_{\min}/c = 4\sqrt{3}/7$, but in a simpler way.

9.4. More on 4-vectors and 4-tensors

This is a good moment to introduce a formalism that will allow us, in particular, to solve the same proton collision problem in one more (and arguably, the most elegant) way. Much more importantly, this formalism will be virtually necessary for the description of the Lorentz transform of the electromagnetic field, and its interaction with relativistic particles – otherwise the formulas would be too cumbersome.

Let us call the 4-vectors we have used before,

Contravariant 4-vectors

$$\boxed{A^\alpha \equiv \{A_0, \mathbf{A}\},} \tag{9.84}$$

contravariant, and denote them with top indices, and introduce also *covariant* vectors,

Covariant 4-vectors

$$\boxed{A_\alpha \equiv \{A_0, -\mathbf{A}\},} \tag{9.85}$$

marked by bottom indices. Now if we form a scalar product of these two vectors using the *standard* (3D-like) rule, just as a sum of the products of the corresponding components, we immediately get

$$A_\alpha A^\alpha \equiv A^\alpha A_\alpha \equiv A_0^2 - A^2. \tag{9.86}$$

Note that the first and the second expressions may be understood as sums over four components of the product, with the summation sign dropped.[37] The scalar product (86) is just the norm of the 4-vector in our former definition, and as we already know, is Lorentz-invariant. Moreover, the scalar product of two different vectors (also a Lorentz invariant), may be rewritten in any of two similar forms:[38]

Scalar product's forms

$$\boxed{A_0 B_0 - \mathbf{A}\cdot\mathbf{B} \equiv A_\alpha B^\alpha = A^\alpha B_\alpha;} \tag{9.87}$$

again, the only caveat is to take one vector in the covariant, and the other one in the contravariant form.

Now let us return to our sample problem (Fig. 10). Since all components ($\mathscr{E}/c$ and $\mathbf{p}$) of the total 4-momentum of our system are conserved at the collision, its norm is conserved as well:

$$(p_a + p_b)_\alpha (p_a + p_b)^\alpha = (4p)_\alpha (4p)^\alpha. \tag{9.88}$$

[37] This compact notation may take some time to be accustomed to, but is very convenient (compact) and can hardly lead to any confusion, due to the following rule: the summation is implied when, and only when the same index is repeated twice, once on the top and another at the bottom. (It is frequently called *dummy index*, because its notation may be replaced with any other letter not used in the same formula.) In this course, this shorthand notation will be used only for 4-vectors, but not for the usual (3D spatial) vectors.

[38] Note also that, by definition, for any two 4-vectors, $A_\alpha B^\alpha = B^\alpha A_\alpha$.

Since now the vector product is the usual math construct, we know that the parentheses on the left-hand side of this equation may be multiplied as usual. We may also swap the operands and move constant factors through products as convenient. As a result, we get

$$(p_a)_\alpha (p_a)^\alpha + (p_b)_\alpha (p_b)^\alpha + 2(p_a)_\alpha (p_b)^\alpha = 16 p_\alpha p^\alpha . \qquad (9.89)$$

Thanks to the Lorentz invariance of each of the terms, we may calculate it in the reference frame we like. For the first two terms on the left-hand side, as well as for the right-hand side term, it is beneficial to use the frames in which that particular proton is at rest; as a result, according to Eq. (77b), each of the two left-hand-side terms equals $(mc)^2$, while the right-hand side equals $16(mc)^2$. On the contrary, the last term on the left-hand side is more easily evaluated in the lab frame, because in it, the three spatial components of the 4-momentum p_b vanish, and the scalar product is just the product of the scalars $\mathscr{E}/c$ for protons a and b. For the latter proton, being at rest, this ratio is just mc so we get a simple equation,

$$(mc)^2 + (mc)^2 + 2\frac{\mathscr{E}_{\min}}{c} mc = 16(mc)^2 , \qquad (9.90)$$

immediately giving the final result $\mathscr{E}_{\min} = 7\, mc^2$, already obtained earlier in two more complex ways.

Let me hope that this example was a convincing demonstration of the convenience of representing 4-vectors in the contravariant (84) and covariant (85) forms,[39] with Lorentz-invariant norms (86). To be useful for more complex tasks, this formalism should be developed a little bit further. In particular, it is crucial to know how the 4-vectors change under the Lorentz transform. For contravariant vectors, we already know the answer (54); let us rewrite it in our new notation:

Lorentz transform: contravariant vectors

$$\boxed{A^\alpha = L^\alpha_\beta A'^\beta .} \qquad (9.91)$$

where L^α_β is the matrix (51), generally called the *mixed Lorentz tensor*:[40]

Mixed Lorentz tensor

$$\boxed{L^\alpha_\beta = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},} \qquad (9.92)$$

Note that though the position of the indices α and β in the Lorentz tensor notation is not crucial, because this tensor is symmetric, it is convenient to place them using the general *index balance rule*: the difference of the numbers of the upper and lower indices should be the same in both parts of any 4-vector/tensor equality. (You may check that all the formulas above do satisfy this rule.)

[39] These forms are 4-vector extensions of the notions of contravariance and covariance, introduced in the 1850s by J. Sylvester (who also introduced the term "matrix" in its mathematical sense) for the description of the change of the usual 3-component spatial vectors at the transfer between different reference frames – e.g., resulting from the frame rotation. In this case, the contravariance or covariance of a vector is uniquely determined by its nature: if the Cartesian coordinates of a vector (such as the non-relativistic velocity $\mathbf{v} = d\mathbf{r}/dt$) are transformed similarly to the radius-vector $\mathbf{r}$, it is called contravariant, while the vectors (such as ∇f) that require the reciprocal transform, are called covariant. In the 4D Minkowski space, both forms may be used for any 4-vector.

[40] Just as the 4-vectors, 4-tensors with two top indices are called contravariant, and those with two bottom indices, are covariant. The tensors with one top and one bottom index are called *mixed*.

In order to rewrite Eq. (91) in a more general form that would not depend on the particular orientation of the coordinate axes (Fig. 1), let us use the contravariant and covariant forms of the 4-vector of the time-space interval (57),

$$dx^\alpha = \{cdt, d\mathbf{r}\}, \qquad dx_\alpha = \{cdt, -d\mathbf{r}\}; \tag{9.93}$$

then its norm (58) may be represented as[41]

$$(ds)^2 \equiv (cdt)^2 - (dr)^2 = dx^\alpha dx_\alpha = dx_\alpha dx^\alpha . \tag{9.94}$$

Applying Eq. (91) to the first, contravariant form of the 4-vector (93), we get

$$dx^\alpha = L^\alpha_\beta dx'^\beta . \tag{9.95}$$

But with our new shorthand notation, we can also write the usual rule of differentiation of each component x^α, considering it a function (in our case, linear) of four arguments x'^β, as follows:[42]

$$dx^\alpha = \frac{\partial x^\alpha}{\partial x'^\beta} dx'^\beta . \tag{9.96}$$

Comparing Eqs. (95) and (96), we can rewrite the general Lorentz transform rule (92) in a new form,

Lorentz transform: general form

$$\boxed{A^\alpha = \frac{\partial x^\alpha}{\partial x'^\beta} A'^\beta .} \tag{9.97a}$$

which does not depend on the coordinate axes' orientation.

It is straightforward to verify that the reciprocal transform may be represented as

Reciprocal Lorentz transform

$$\boxed{A'^\alpha = \frac{\partial x'^\alpha}{\partial x^\beta} A^\beta .} \tag{9.97b}$$

However, the reciprocal transform has to differ from the direct one only by the sign of the relative velocity of the frames, so for the coordinate choice shown in Fig. 1, its matrix is

[41] Another way to write this relation is $(ds)^2 = g_{\alpha\beta}\, dx^\alpha dx^\beta = g^{\alpha\beta} dx_\alpha\, dx_\beta$, where double summation over indices α and β is implied, and g is the so-called *metric tensor*,

$$g^{\alpha\beta} \equiv g_{\alpha\beta} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

which may be used, in particular, to transfer a covariant vector into the corresponding contravariant one and back: $A^\alpha = g^{\alpha\beta} A_\beta$, $A_\alpha = g_{\alpha\beta} A^\beta$. The metric tensor plays a key role in general relativity, in which it is affected by gravity – "curved" by particles' masses.

[42] Note that in the index balance rule, the top index in the denominator of a fraction is counted as a bottom index in the numerator, and vice versa.

$$\frac{\partial x'^{\alpha}}{\partial x^{\beta}} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{9.98}$$

Since according to Eqs. (84)-(85), covariant 4-vectors differ from the contravariant ones by the sign of their spatial components, their direct transform is given by matrix (98). Hence their direct and reciprocal transforms may be represented, respectively, as

Lorentz transform: covariant vectors

$$\boxed{A_{\alpha} = \frac{\partial x'^{\beta}}{\partial x^{\alpha}} A'_{\beta}, \qquad A'_{\alpha} = \frac{\partial x^{\beta}}{\partial x'^{\alpha}} A_{\beta},} \tag{9.99}$$

evidently satisfying the index balance rule. (Note that primed quantities are now multiplied, rather than divided as in the contravariant case.) As a sanity check, let us apply this formalism to the scalar product $A_{\alpha}A^{\alpha}$. As Eq. (96) shows, the implicit-sum notation allows us to multiply and divide any equality by the same partial differential of a coordinate, so we can write:

$$A_{\alpha}A^{\alpha} = \frac{\partial x'^{\beta}}{\partial x^{\alpha}} \frac{\partial x^{\alpha}}{\partial x'^{\gamma}} A'_{\beta}A'^{\gamma} = \frac{\partial x'^{\beta}}{\partial x'^{\gamma}} A'_{\beta}A'^{\gamma} = \delta_{\beta\gamma} A'_{\beta}A'^{\gamma} = A'_{\gamma}A'^{\gamma}, \tag{9.100}$$

i.e. the scalar product $A_{\alpha}A^{\alpha}$ (as well as $A^{\alpha}A_{\alpha}$) is Lorentz-invariant, as it should be.

Now, let us consider the 4-vectors of derivatives. Here we should be very careful. Consider, for example, the following 4-vector operator

$$\frac{\partial}{\partial x^{\alpha}} \equiv \left\{ \frac{\partial}{\partial (ct)}, \nabla \right\}, \tag{9.101}$$

As was discussed above, the operator is not changed by its multiplication and division by another differential, e.g., $\partial x'^{\beta}$ (with the corresponding implied summation over all four values of β), so

$$\frac{\partial}{\partial x^{\alpha}} = \frac{\partial x'^{\beta}}{\partial x^{\alpha}} \frac{\partial}{\partial x'^{\beta}}. \tag{9.102}$$

But, according to the first of Eqs. (99), this is exactly how the covariant vectors are Lorentz-transformed! Hence, we have to consider the derivative over a *contravariant* space-time interval as a *covariant* 4-vector, and vice versa.[43] (This result might be also expected from the index balance rule.) In particular, this means that the scalar product

$$\frac{\partial}{\partial x^{\alpha}} A^{\alpha} \equiv \frac{\partial A_0}{\partial (ct)} + \nabla \cdot \mathbf{A} \tag{9.103}$$

should be Lorentz-invariant for any legitimate 4-vector. A convenient shorthand for the covariant derivative, which complies with the index balance rule, is

$$\frac{\partial}{\partial x^{\alpha}} \equiv \partial_{\alpha}, \tag{9.104}$$

[43] As was mentioned above, this is also a property of the reference-frame transform of the "usual" 3D vectors.

so the invariant scalar product may be written just as $\partial_\alpha A^\alpha$. A similar definition of the contravariant derivative,

$$\partial^\alpha \equiv \frac{\partial}{\partial x_\alpha} = \left\{ \frac{\partial}{\partial(ct)}, -\nabla \right\}, \tag{9.105}$$

allows us to write the Lorentz-invariant scalar product (103) in any of the following two forms:

$$\frac{\partial A_0}{\partial(ct)} + \nabla \cdot \mathbf{A} = \partial^\alpha A_\alpha = \partial_\alpha A^\alpha . \tag{9.106}$$

Finally, let us see how the general Lorentz transform changes 4-tensors. A second-rank 4×4 matrix is a legitimate 4-tensor if the 4-vectors it relates obey the Lorentz transform. For example, if two legitimate 4-vectors are related as

$$A^\alpha = T^{\alpha\beta} B_\beta , \tag{9.107}$$

we should require that

$$A'^\alpha = T'^{\alpha\beta} B'_\beta , \tag{9.108}$$

where A^α and A'^α are related by Eqs. (97), while B_β and B'_β, by Eqs. (99). This requirement immediately yields

Lorentz transform of 4-tensors

$$\boxed{T^{\alpha\beta} = \frac{\partial x^\alpha}{\partial x'^\gamma} \frac{\partial x^\beta}{\partial x'^\delta} T'^{\gamma\delta}, \qquad T'^{\alpha\beta} = \frac{\partial x'^\alpha}{\partial x^\gamma} \frac{\partial x'^\beta}{\partial x^\delta} T^{\gamma\delta},} \tag{9.109}$$

with the implied summation over two indices, γ and δ. The rules for the covariant and mixed tensors are similar.[44]

9.5. Maxwell equations in the 4-form

This 4-vector formalism background is sufficient to analyze the Lorentz transform of the electromagnetic field. Just to warm up, let us consider the continuity equation (4.5),

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0 , \tag{9.110}$$

which expresses the electric charge conservation, and as we already know, is compatible with the Maxwell equations. If we now define the contravariant and covariant *4-vectors of electric current* as

4-vector of electric current

$$\boxed{j^\alpha \equiv \{\rho c, \mathbf{j}\}, \qquad j_\alpha \equiv \{\rho c, -\mathbf{j}\},} \tag{9.111}$$

then Eq. (110) may be represented in the form

Continuity equation: 4-form

$$\boxed{\partial^\alpha j_\alpha = \partial_\alpha j^\alpha = 0 ,} \tag{9.112}$$

showing that the continuity equation is *form-invariant*[45] with respect to the Lorentz transform.

---

[44] It is straightforward to check that transfer between the contravariant and covariant forms of the same tensor may be readily achieved using the metric tensor g: $T_{\alpha\beta} = g_{\alpha\gamma} T^{\gamma\delta} g_{\delta\beta}$, $T^{\alpha\beta} = g^{\alpha\gamma} T_{\gamma\delta} g^{\delta\beta}$.

[45] In some texts, the formulas preserving their form at a transform are called "covariant", creating a possibility for confusion with the covariant vectors and tensors. On the other hand, calling such *formulas* "invariant" would not distinguish them properly from invariant *quantities*, such as the scalar products of 4-vectors.

Of course, such a form-invariance of a relation does not mean that all component *values* of the 4-vectors participating in it are the same in both frames. For example, let us have some static charge density ρ in frame 0; then Eq. (97b), applied to the contravariant form of the 4-vector (111), reads

$$j'^{\alpha} = \frac{\partial x'^{\alpha}}{\partial x^{\beta}} j^{\beta}, \qquad \text{with } j^{\beta} = \{\rho c, 0, 0, 0\}. \tag{9.113}$$

Using the particular form (98) of the reciprocal Lorentz matrix for the coordinate choice shown in Fig. 1, we see that this relation yields

Lorentz transforms of ρ and **j**

$$\boxed{\rho' = \gamma\rho, \qquad j'_x = -\gamma\beta\rho c = -\gamma v\rho, \qquad j'_y = j'_z = 0.} \tag{9.114}$$

Since the charge velocity, as observed from frame 0', is (–**v**), the non-relativistic results would be $\rho' = \rho$, $\mathbf{j}' = -\mathbf{v}\rho$. The additional γ factor in the relativistic results is caused by the length contraction: $dx' = dx/\gamma$, so to keep the total charge $dQ = \rho d^3r = \rho dxdydz$ inside the elementary volume $d^3r = dxdydz$ intact, ρ (and hence j_x) should increase proportionally.

Next, at the end of Chapter 6 we have seen that Maxwell equations for the electromagnetic potentials ϕ and **A** may be represented in similar forms (6.118), under the Lorenz (again, not "Lorentz", please!) gauge condition (6.117). For free space, this condition takes the form

$$\nabla \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t} = 0. \tag{9.115}$$

This expression gives us a hint of how to form the 4-vector of electromagnetic potentials:[46]

4-vector of potentials

$$\boxed{A^{\alpha} \equiv \left\{\frac{\phi}{c}, \mathbf{A}\right\}, \qquad A_{\alpha} \equiv \left\{\frac{\phi}{c}, -\mathbf{A}\right\};} \tag{9.116}$$

indeed, this vector satisfies Eq. (115) in its 4-form:

Lorenz gauge: 4-form

$$\boxed{\partial^{\alpha} A_{\alpha} = \partial_{\alpha} A^{\alpha} = 0.} \tag{9.117}$$

Since this scalar product is Lorentz-invariant, and the derivatives (104)-(105) are legitimate 4-vectors, this implies that the 4-vector (116) is also legitimate, i.e. obeys the Lorentz transform formulas (97), (99). Even more convincing evidence of this fact may be obtained from the Maxwell equations (6.118) for the potentials. In free space, they may be rewritten as

$$\left[\frac{\partial^2}{\partial(ct)^2} - \nabla^2\right]\frac{\phi}{c} = \frac{\rho c}{\varepsilon_0 c^2} \equiv \mu_0(\rho c), \qquad \left[\frac{\partial^2}{\partial(ct)^2} - \nabla^2\right]\mathbf{A} = \mu_0 \mathbf{j}. \tag{9.118}$$

Using the definition (116), these equations may be merged to one:[47]

Maxwell equation for 4-potential

$$\boxed{\Box A^{\alpha} = \mu_0 j^{\alpha},} \tag{9.119}$$

where $\Box$ is the *d'Alembert operator*,[48] which may be represented as either of two scalar products:

[46] In the Gaussian units, the scalar potential should not be divided by c in this relation.

[47] In the Gaussian units, the coefficient μ_0 in Eq. (119) should be replaced, as usual, with $4\pi/c$.

D'Alembert operator

$$\Box \equiv \frac{\partial^2}{\partial(ct)^2} - \nabla^2 = \partial^\beta \partial_\beta = \partial_\beta \partial^\beta, \tag{9.120}$$

and hence is Lorentz-invariant. Because of that, and the fact that the Lorentz transform changes both 4-vectors A^α and j^α in a similar way, Eq. (119) does not depend on the reference frame choice. Thus we have arrived at a key point of this chapter: we see that the Maxwell equations are indeed form-invariant with respect to the Lorentz transform. As a by-product, the 4-vector form (119) of these equations (for potentials) is extremely simple – and beautiful!

However, as we have seen in Chapter 7, for many applications the Maxwell equations for the field vectors are more convenient; so let us represent them in the 4-form as well. For that, we may express all Cartesian components of the usual (3D) field vector vectors (6.7),

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}, \qquad \mathbf{B} = \nabla \times \mathbf{A}, \tag{9.121}$$

via those of the potential 4-vector A^α. For example,

$$E_x = -\frac{\partial \phi}{\partial x} - \frac{\partial A_x}{\partial t} \equiv -c\left(\frac{\partial}{\partial x}\frac{\phi}{c} + \frac{\partial A_x}{\partial(ct)}\right) \equiv -c\left(\partial^0 A^1 - \partial^1 A^0\right), \tag{9.122}$$

$$B_x = \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \equiv -\left(\partial^2 A^3 - \partial^3 A^2\right). \tag{9.123}$$

Completing similar calculations for other field components (or just generating them by appropriate index shifts), we find that the following antisymmetric, contravariant *field-strength tensor*,

$$\boxed{F^{\alpha\beta} \equiv \partial^\alpha A^\beta - \partial^\beta A^\alpha,} \tag{9.124}$$

may be expressed via the field components as follows:[49]

Field-strength tensors

$$\boxed{F^{\alpha\beta} = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix},} \tag{9.125a}$$

so the covariant form of the tensor is

$$\boxed{F_{\alpha\beta} \equiv g_{\alpha\gamma} F^{\gamma\delta} g_{\delta\beta} = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & -B_z & B_y \\ -E_y/c & B_z & 0 & -B_x \\ -E_z/c & -B_y & B_x & 0 \end{pmatrix}.} \tag{9.125b}$$

[48] Named after Jean-Baptiste le Rond d'Alembert (1717-1783), who has made several pioneering contributions to the general theory of waves – see, e.g., CM Chapter 6. (Some older textbooks use notation $\Box^2$ for this operator.)

[49] In Gaussian units, this formula, as well as Eq. (131) for $G^{\alpha\beta}$, do not have the factor c in all the denominators.

If Eq. (124) looks a bit too bulky, please note that as a reward, the pair of *inhomogeneous* Maxwell equations, i.e. two equations of the system (6.99), which in free space ($\mathbf{D} = \varepsilon_0\mathbf{E}$, $\mathbf{B} = \mu_0\mathbf{H}$) may be rewritten as

$$\nabla \cdot \frac{\mathbf{E}}{c} = \mu_0 c\rho, \qquad \nabla \times \mathbf{B} - \frac{\partial}{\partial(ct)}\frac{\mathbf{E}}{c} = \mu_0 \mathbf{j}, \tag{9.126}$$

may now be expressed in a very simple (and manifestly form-invariant) way,

Maxwell equation for tensor F

$$\boxed{\partial_\alpha F^{\alpha\beta} = \mu_0 j^\beta,} \tag{9.127}$$

which is comparable with Eq. (119) in its simplicity – and beauty. Somewhat counter-intuitively, the pair of *homogeneous* Maxwell equations of the system (6.99),

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0, \qquad \nabla \cdot \mathbf{B} = 0, \tag{9.128}$$

look, in the 4-vector notation, a bit more complicated:[50]

$$\boxed{\partial_\alpha F_{\beta\gamma} + \partial_\beta F_{\gamma\alpha} + \partial_\gamma F_{\alpha\beta} = 0.} \tag{9.129}$$

Note, however, that Eqs. (128) may be also represented in a much simpler 4-form,

$$\partial_\alpha G^{\alpha\beta} = 0, \tag{9.130}$$

using the so-called *dual tensor*

$$G^{\alpha\beta} = \begin{pmatrix} 0 & B_x & B_y & B_z \\ -B_x & 0 & -E_z/c & E_y/c \\ -B_y & E_z/c & 0 & -E_x/c \\ -B_z & -E_y/c & E_x/c & 0 \end{pmatrix}, \tag{9.131}$$

which may be obtained from $F^{\alpha\beta}$, given by Eq. (125a), by the following replacements:

$$\frac{\mathbf{E}}{c} \to -\mathbf{B}, \quad \mathbf{B} \to \frac{\mathbf{E}}{c}. \tag{9.132}$$

Besides the proof of the form-invariance of the Maxwell equations with respect to the Lorentz transform, the 4-vector formalism allows us to achieve our initial goal: to find out how the electric and magnetic field components change at the transfer between two (inertial!) reference frames. For that, let us apply to the tensor $F^{\alpha\beta}$ the reciprocal Lorentz transform described by the second of Eqs. (109). Generally, it gives, for each field component, a sum of 16 terms, but since (for our choice of coordinates, shown in Fig. 1) there are many zeros in the Lorentz transform matrix, and the diagonal components of $F^{\gamma\delta}$ equal zero as well, the calculations are rather doable. Let us calculate, for example, $E'_x \equiv -cF'^{01}$. The only non-zero terms on the right-hand side are

$$E'_x = -cF'^{01} = -c\left(\frac{\partial x'^0}{\partial x^1}\frac{\partial x'^1}{\partial x^0}F^{10} + \frac{\partial x'^0}{\partial x^0}\frac{\partial x'^1}{\partial x^1}F^{01}\right) \equiv -c\gamma^2\left(\beta^2 - 1\right)\frac{E_x}{c} \equiv E_x. \tag{9.133}$$

[50] To be fair, note that just as Eq. (127), Eq. (129) is also a set of four scalar equations – in the latter case with the indices α, β, and γ taking any three *different* values of the set {0, 1, 2, 3}.

Repeating the calculation for the other five components of the fields, we get very important relations

$$\begin{aligned} E'_x &= E_x, & B'_x &= B_x, \\ E'_y &= \gamma\left(E_y - vB_z\right), & B'_y &= \gamma\left(B_y + vE_z/c^2\right), \\ E'_z &= \gamma\left(E_z + vB_y\right), & B'_z &= \gamma\left(B_z - vE_y/c^2\right), \end{aligned} \quad (9.134)$$

whose more compact "semi-vector" form is

Lorentz transform of field components

$$\boxed{\begin{aligned} E'_{||} &= E_{||}, & B'_{||} &= B_{||}, \\ \mathbf{E}'_{\perp} &= \gamma\left(\mathbf{E} + \mathbf{v}\times\mathbf{B}\right)_{\perp}, & \mathbf{B}'_{\perp} &= \gamma\left(\mathbf{B} - \mathbf{v}\times\mathbf{E}/c^2\right)_{\perp}, \end{aligned}} \quad (9.135)$$

where the indices $||$ and $\perp$ stand, respectively, for the field components parallel and normal to the relative velocity $\mathbf{v}$ of the two reference frames. In the non-relativistic limit, the Lorentz factor γ tends to 1, and Eqs. (135) acquire an even simpler form

$$\mathbf{E}' \to \mathbf{E} + \mathbf{v}\times\mathbf{B}, \qquad \mathbf{B}' \to \mathbf{B} - \frac{1}{c^2}\mathbf{v}\times\mathbf{E}. \quad (9.136)$$

Thus we see that the electric and magnetic fields are transformed to each other even in the first order of the v/c ratio. For example, if we fly across the field lines of a uniform, static, purely electric field $\mathbf{E}$ (e.g., the one in a plane capacitor) we will see not only the electric field's renormalization (in the second order of the v/c ratio), but also a non-zero dc magnetic field $\mathbf{B}'$ perpendicular to both the vector $\mathbf{E}$ and the vector $\mathbf{v}$, i.e. to the direction of our motion. This is of course what might be expected from the relativity principle: from the point of view of the moving observer (which is as legitimate as that of a stationary observer), the surface charges of the capacitor's plates, that create the field $\mathbf{E}$, move back creating the dc currents (114), which induce the magnetic field $\mathbf{B}'$. Similarly, motion across a magnetic field creates, from the point of view of the moving observer, an electric field.

This fact is very important conceptually. One may say there is no such thing in Mother Nature as an electric field (or a magnetic field) all by itself. Not only can the electric field induce the magnetic field (and vice versa) in dynamics, but even in an apparently static configuration, what exactly we measure depends on our speed relative to the field sources – justifying once again the term *electromagnetism* for the field of physics we are studying in this course.

Another simple but very important application of Eqs. (134)-(135) is the calculation of the fields created by a charged particle moving in free space by inertia, i.e. along a straight line with constant velocity $\mathbf{u}$, at the *impact parameter*[51] (the closest distance) b from the observer. Selecting the reference frame 0' to move with the particle in its origin, and the reference frame 0 to reside in the "lab" in which the fields $\mathbf{E}$ and $\mathbf{B}$ are measured, we can use the above formulas with $\mathbf{v} = \mathbf{u}$. In this case, the fields $\mathbf{E}'$ and $\mathbf{B}'$ may be calculated from, respectively, electro- and magnetostatics:

$$\mathbf{E}' = \frac{q}{4\pi\varepsilon_0}\frac{\mathbf{r}'}{r'^3}, \qquad \mathbf{B}' = 0, \quad (9.137)$$

because in frame 0', the particle does not move. Selecting the coordinate axes so that at the measurement point, $x = 0$, $y = b$, $z = 0$ (Fig. 11a), for this point we may write $x' = -ut'$, $y' = b$, $z' = 0$, so $r' = (u^2t'^2 + b^2)^{1/2}$, and the Cartesian components of the fields (137) are:

[51] This term is very popular in the theory of particle scattering – see, e.g., CM Sec. 3.7.

$$E'_x = -\frac{q}{4\pi\varepsilon_0}\frac{ut'}{\left(u^2t'^2+b^2\right)^{3/2}}, \qquad E'_y = \frac{q}{4\pi\varepsilon_0}\frac{b}{\left(u^2t'^2+b^2\right)^{3/2}}, \qquad E'_z = 0,$$
$$B'_x = B'_y = B'_z = 0. \tag{9.138}$$

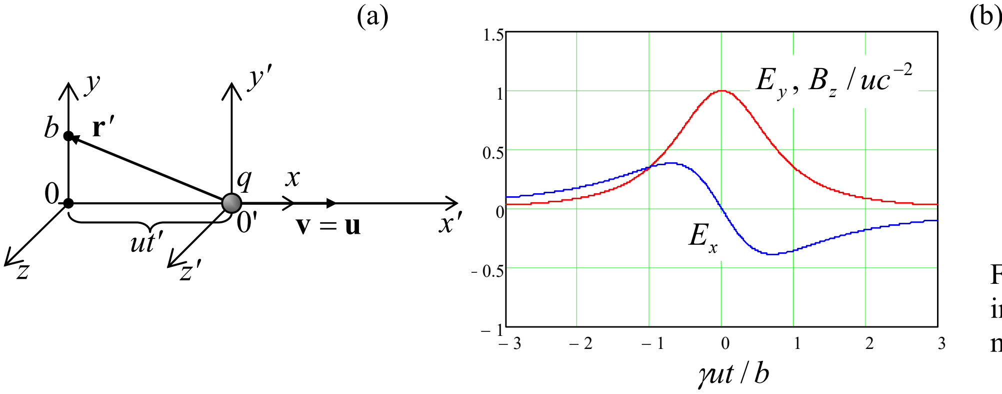


Fig. 9.11. The field pulses induced by a uniformly moving charge.

Now using the last of Eqs. (19b) with $x = 0$, giving $t' = \gamma t$, and the relations reciprocal to Eqs. (134) for the field transform (they are similar to the direct transform but with v replaced with $-v = -u$), in the lab frame we get

$$E_x = E'_x = -\frac{q}{4\pi\varepsilon_0}\frac{u\gamma t}{\left(u^2\gamma^2t^2+b^2\right)^{3/2}}, \qquad E_y = \gamma E'_y = \frac{q}{4\pi\varepsilon_0}\frac{\gamma b}{\left(u^2\gamma^2t^2+b^2\right)^{3/2}}, \qquad E_z = 0, \tag{9.139}$$

$$B_x = 0, \qquad B_y = 0, \qquad B_z = \frac{\gamma u}{c^2}E'_y = \frac{u}{c^2}\frac{q}{4\pi\varepsilon_0}\frac{\gamma b}{\left(u^2\gamma^2t^2+b^2\right)^{3/2}} \equiv \frac{u}{c^2}E_y. \tag{9.140}$$

These results,[52] plotted in Fig. 11b in the units of $\gamma q^2/4\pi\varepsilon_0 b^2$, reveal two major effects. First, the charge passage by the observer generates not only an electric field pulse but also a magnetic field pulse. This is natural, because, as was repeatedly discussed in Chapter 5, any charge motion is essentially an electric current.[53] Second, Eqs. (139)-(140) show that the pulse duration scale is

$$\Delta t = \frac{b}{\gamma u} \equiv \frac{b}{u}\left(1-\frac{u^2}{c^2}\right)^{1/2}, \tag{9.141}$$

i.e. shrinks to virtually zero as the charge's velocity u approaches the speed of light. This is of course a direct corollary of the relativistic length contraction. Indeed, in the frame 0' moving with the charge, the longitudinal spread of its electric field at distance b from the motion line is of the order of $\Delta x' = b$. When observed from the lab frame 0, this interval, in accordance with Eq. (20), shrinks to $\Delta x = \Delta x'/\gamma = b/\gamma$, and hence so does the pulse duration scale $\Delta t = \Delta x/u = b/\gamma u$.

[52] In the next chapter, we will re-derive them in a different way.

[53] It is straightforward to use Eqs. (140) and the linear superposition principle to calculate, for example, the magnetic field of a string of charges moving along the same line and separated by equal distances $\Delta x = a$ (so the average current, as measured in frame 0, is qu/a), and to show that the time-average of the magnetic field is given by the familiar Eq. (5.20) of magnetostatics, with b instead of ρ.

9.6. Relativistic particles in electric and magnetic fields

Now let us analyze the dynamics of charged particles in electric and magnetic fields. Inspired by "our" success in forming the 4-vector (75) of energy-momentum, with the contravariant form

$$p^\alpha = \left\{\frac{\mathscr{E}}{c}, \mathbf{p}\right\} = \gamma\{mc, \mathbf{p}\} = m\frac{dx^\alpha}{d\tau} \equiv mu^\alpha, \tag{9.142}$$

where u^α is the contravariant form of the 4-velocity (63) of the particle,

$$u^\alpha \equiv \frac{dx^\alpha}{d\tau}, \qquad u_\alpha \equiv \frac{dx_\alpha}{d\tau}, \tag{9.143}$$

we may notice that the non-relativistic equation of motion, resulting from the Lorentz-force formula (5.10) for the three spatial components of p^α, for a charged particle's motion in an electromagnetic field,

Particle's equation of motion

$$\boxed{\frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{u} \times \mathbf{B}),} \tag{9.144}$$

is fully consistent with the following 4-vector equality (which is evidently form-invariant with respect to the Lorentz transform):

Particle's dynamics: 4-form

$$\boxed{\frac{dp^\alpha}{d\tau} = qF^{\alpha\beta}u_\beta .} \tag{9.145}$$

For example, according to Eq. (125), the $\alpha = 1$ component of this equation reads

$$\frac{dp^1}{d\tau} = qF^{1\beta}u_\beta = q\left[\frac{E_x}{c}\gamma c + 0\cdot(-\gamma u_x) + (-B_z)(-\gamma u_y) + B_y(-\gamma u_z)\right] = q\gamma[\mathbf{E} + \mathbf{u}\times\mathbf{B}]_x, \tag{9.146}$$

and similarly for two other spatial components ($\alpha = 2$ and $\alpha = 3$). It may look that these expressions differ from the 2nd Newton law (144) by an extra factor of γ. However, plugging into Eq. (146) the definition of the proper time interval, $d\tau = dt/\gamma$, and canceling γ in both parts, we recover Eq. (144) exactly – for *any* velocity of the particle! The only caveat is that if u is comparable with c, the vector $\mathbf{p}$ in Eq. (144) has to be understood as the relativistic momentum (70), proportional to the velocity-dependent mass $M = \gamma m \geq m$ rather than to the rest mass m.

The only remaining general task is to examine the meaning of the 0th component of Eq. (145). Let us spell it out:

$$\frac{dp^0}{d\tau} = qF^{0\beta}u_\beta = q\left[0\cdot\gamma c + \left(-\frac{E_x}{c}\right)(-\gamma u_x) + \left(-\frac{E_y}{c}\right)(-\gamma u_y) + \left(-\frac{E_z}{c}\right)(-\gamma u_z)\right] = q\gamma\frac{\mathbf{E}}{c}\cdot\mathbf{u}. \tag{9.147}$$

Recalling that $p^0 = \mathscr{E}/c$, and using the basic relation $d\tau = dt/\gamma$ again, we see that Eq. (147) looks exactly like the non-relativistic relation for the kinetic energy change (what is sometimes called the *work-energy principle*, in our case for the Lorentz force only[54]):

[54] See, e.g., CM Eq. (1.20) divided by dt, and with $d\mathbf{p}/dt = \mathbf{F} = q\mathbf{E}$. (As a reminder, the magnetic field cannot affect the particle's energy, because the magnetic component of the Lorentz force is perpendicular to its velocity.)

$$\boxed{\frac{d\mathscr{E}}{dt} = q\mathbf{E}\cdot\mathbf{u},} \qquad (9.148)$$

Particle's energy: evolution

besides that in the relativistic case, the energy has to be taken in the general form (73).

Without question, the 4-component equation (145) of the relativistic dynamics is absolutely beautiful in its simplicity. However, for the solution of particular problems, Eqs. (144) and (148) are frequently more convenient. As an illustration of this point, let us now use these equations to explore relativistic effects at charged particle motion in uniform, time-independent electric and magnetic fields. In doing that, we will, for the time being, neglect the contributions into the field by the particle itself.[55]

(i) Uniform magnetic field. Let the magnetic field be constant and uniform in the "lab" reference frame 0 that is used for measurements. Then in this frame, Eqs. (144) and (148) yield

$$\frac{d\mathbf{p}}{dt} = q\mathbf{u}\times\mathbf{B}, \qquad \frac{d\mathscr{E}}{dt} = 0. \qquad (9.149)$$

From the second equation, $\mathscr{E}$ = const, we get u = const, $\beta \equiv u/c$ = const, $\gamma \equiv (1-\beta^2)^{-1/2}$ = const, and $M \equiv \gamma m$ = const, so the first of Eqs. (149) may be rewritten as

$$\frac{d\mathbf{u}}{dt} = \mathbf{u}\times\boldsymbol{\omega}_c, \qquad (9.150)$$

where $\boldsymbol{\omega}_c$ is the vector directed along the magnetic field $\mathbf{B}$, with the magnitude equal to the following *cyclotron frequency* (sometimes called "gyrofrequency"):

$$\boxed{\omega_c \equiv \frac{qB}{M} = \frac{qB}{\gamma m} = \frac{qc^2B}{\mathscr{E}}.} \qquad (9.151)$$

Cyclotron frequency

If the particle's initial velocity $\mathbf{u}_0$ is perpendicular to the magnetic field, Eq. (150) describes its circular motion, with a constant speed $u = u_0$, in a plane normal to $\mathbf{B}$, with the angular velocity (151). In the non-relativistic limit $u \ll c$, when $\gamma \to 1$, i.e. $M \to m$, the cyclotron frequency ω_c equals qB/m, i.e. is independent of the speed. However, as the kinetic energy of the particle is increased to become comparable with its rest energy mc^2, the frequency decreases, and in the ultra-relativistic limit,

$$\omega_c \approx qc\frac{B}{p} \ll \frac{qB}{m}, \qquad \text{for } u \approx c. \qquad (9.152)$$

The cyclotron motion's radius may be calculated as $R = u/\omega_c$; in the non-relativistic limit, it is proportional to the particle's speed, i.e. to the square root of its kinetic energy. However, as Eq. (151) shows, in the general case the radius is proportional to the particle's relativistic momentum rather than its speed:

$$\boxed{R = \frac{u}{\omega_c} = \frac{Mu}{qB} = \frac{m\gamma u}{qB} = \frac{1}{q}\frac{p}{B},} \qquad (9.153)$$

Cyclotron radius

so in the ultra-relativistic limit, when $p \approx \mathscr{E}/c$, R is proportional to the kinetic energy.

[55] As was emphasized earlier in this course, in statics this contribution is formally infinite and has to be ignored. In dynamics, this is generally not true; these *self-action effects* (which are, in most cases, negligible) will be discussed in the next chapter.

These dependencies of ω_c and R on energy are the major factors in the design of circular accelerators of charged particles. In the simplest of these machines (the *cyclotron*, invented in 1929 by Ernest Orlando Lawrence), the frequency ω of the accelerating ac electric field is constant, so even if it is tuned to the ω_c of the initially injected particles, the drop of the cyclotron frequency with energy eventually violates this tuning. Due to this reason, the largest achievable particle's speed is limited to just ~0.1 c (for protons, corresponding to the kinetic energy of just ~15 MeV). This problem may be addressed in several ways. In particular, in *synchrotrons* (such as Fermilab's Tevatron and the CERN's Large Hadron Collider, LHC[56]) the magnetic field is gradually increased in time to compensate for the momentum increase ($B \propto p$), so both R (148) and ω_c (147) stay constant, enabling proton acceleration to energies as high as ~ 7 TeV, i.e. ~2,000 mc^2.[57]

Returning to our initial problem, if the particle's initial velocity has a component $u_{||}$ along the magnetic field, then it is conserved in time, so the trajectory is a spiral around the magnetic field lines. As Eqs. (149) show, in this case, Eq. (150) remains valid but in Eqs. (151) and (153) the full speed and momentum have to be replaced with magnitudes of their (also time-conserved) components, $u_\perp$ and $p_\perp$, normal to **B**, while the Lorentz factor γ in those formulas still includes the full speed of the particle.

Finally, in the special case when the particle's initial velocity is directed *exactly* along the magnetic field's direction, it continues to move straight along the vector **B**. In this case, the cyclotron frequency still has the non-zero value (151) but does not correspond to any real motion, because $R = 0$.

(ii) Uniform electric field. This problem is (technically) more complex than the previous one because in the electric field, the particle's energy changes. Directing the z-axis along the field **E**, from Eq. (144) we get

$$\frac{dp_z}{dt} = qE, \qquad \frac{d\mathbf{p}_\perp}{dt} = 0. \tag{9.154}$$

If E does not change in time, the first integration of these equations is elementary,

$$p_z(t) = p_z(0) + qEt, \qquad \mathbf{p}_\perp(t) = \text{const} = \mathbf{p}_\perp(0), \tag{9.155}$$

but the further integration requires care because the effective mass $M = \gamma m$ of the particle depends on its full speed u, with

$$u^2 = u_z^2 + u_\perp^2, \tag{9.156}$$

making the two motions, along and across the field, mutually dependent.

If the initial velocity is perpendicular to the field **E**, i.e. if $p_z(0) = 0$, $p_\perp(0) = p(0) \equiv p_0$, the easiest way to proceed is to calculate the kinetic energy first:

$$\mathscr{E}^2 = \left(mc^2\right)^2 + c^2 p^2(t) \equiv \mathscr{E}_0^2 + c^2\left(qEt\right)^2, \qquad \text{where } \mathscr{E}_0 \equiv \left[(mc^2)^2 + c^2 p_0^2\right]^{1/2}. \tag{9.157}$$

On the other hand, we can calculate the same energy by integrating Eq. (148),

[56] See https://home.cern/topics/large-hadron-collider.

[57] I am sorry I have no more time/space to discuss particle accelerator physics, and have to refer the interested reader to special literature, for example, either S. Lee, *Accelerator Physics*, 2nd ed., World Scientific, 2004, or E. Wilson, *An Introduction to Particle Accelerators*, Oxford U. Press, 2001.

$$\frac{d\mathscr{E}}{dt} = q\mathbf{E}\cdot\mathbf{u} \equiv qE\frac{dz}{dt}, \tag{9.158}$$

over time, with a simple result:

$$\mathscr{E} = \mathscr{E}_0 + qEz(t), \tag{9.159}$$

where (just for the notation simplicity) I took $z(0) = 0$. Requiring Eq. (159) to give the same $\mathscr{E}^2$ as Eq. (157), we get a quadratic equation for the function $z(t)$,

$$\mathscr{E}_0^2 + c^2(qEt)^2 = [\mathscr{E}_0 + qEz(t)]^2, \tag{9.160}$$

whose solution (with the sign before the square root corresponding to $E > 0$, i.e. to $z \geq 0$) is

$$z(t) = \frac{\mathscr{E}_0}{qE}\left\{\left[1 + \left(\frac{cqEt}{\mathscr{E}_0}\right)^2\right]^{1/2} - 1\right\}. \tag{9.161}$$

Now let us find the particle's trajectory. Directing the x-axis so that the initial velocity vector (and hence the velocity vector at any further instant) is within the [x, z] plane, i.e. that $y(t) = 0$ identically, we may use Eqs. (155) to calculate the trajectory's slope, at its arbitrary point, as

$$\frac{dz}{dx} \equiv \frac{dz/dt}{dx/dt} \equiv \frac{Mu_z}{Mu_x} \equiv \frac{p_z}{p_x} = \frac{qEt}{p_0}. \tag{9.162}$$

Now let us use Eq. (160) to express the numerator of this fraction, qEt, as a function of z:

$$qEt = \frac{1}{c}\left[(\mathscr{E}_0 + qEz)^2 - \mathscr{E}_0^2\right]^{1/2}. \tag{9.163}$$

Plugging this expression into Eq. (162), we get

$$\frac{dz}{dx} = \frac{1}{cp_0}\left[(\mathscr{E}_0 + qEz)^2 - \mathscr{E}_0^2\right]^{1/2}. \tag{9.164}$$

This differential equation may be readily integrated separating the variables z and x, and using the substitution $\xi \equiv \cosh^{-1}(qEz/\mathscr{E}_0 + 1)$. Selecting the origin of axis x at the initial point, so $x(0) = 0$, we finally get the trajectory:

$$z = \frac{\mathscr{E}_0}{qE}\left(\cosh\frac{qEx}{cp_0} - 1\right). \tag{9.165}$$

This curve is usually called the *catenary*, but sometimes the "chainette" – because it (with the proper constant replacement) describes, in particular, the stationary shape of a heavy uniform chain in a uniform gravity field directed along the z-axis. At the initial part of the trajectory, where $qEx \ll cp_0(0)$, this expression may be approximated with the first non-zero term of its Taylor expansion in small x, giving the following parabola:

$$z = \frac{\mathscr{E}_0 qE}{2}\left(\frac{x}{cp_0}\right)^2, \tag{9.166}$$

so if the initial velocity of the particle is much lower than c (i.e. $p_0 \approx mu_0$, $\mathscr{E}_0 \approx mc^2$), we get the very familiar non-relativistic formula:

$$z = \frac{qE}{2mu_0^2}x^2 \equiv \frac{a}{2}t^2, \qquad \text{with } a = \frac{F}{m} = \frac{qE}{m}. \tag{9.167}$$

The generalization of this solution to the case of an arbitrary direction of the particle's initial velocity is left for the reader's exercise.

<u>(iii) Crossed uniform magnetic and electric fields</u> ($\mathbf{E} \perp \mathbf{B}$). In view of the somewhat bulky solution of the previous problem (i.e. the particular case of the current problem for $\mathbf{B} = 0$), one might think that this problem, with $\mathbf{B} \neq 0$, should be forbiddingly complex for an analytical solution. Counter-intuitively, this is not the case, due to the help from the field transform relations (135). Let us consider two possible cases.

Case 1: $E/c < B$. Let us consider an inertial reference frame 0' moving (relatively the "lab" reference frame 0 in that the fields $\mathbf{E}$ and $\mathbf{B}$ are measured) with the following velocity:

$$\mathbf{v} = \frac{\mathbf{E}\times\mathbf{B}}{B^2}, \tag{9.168}$$

and hence the speed $v = c(E/c)/B < c$. Selecting the coordinate axes as shown in Fig. 12, so

$$E_x = 0, \quad E_y = E, \quad E_z = 0; \qquad B_x = 0, \quad B_y = 0, \quad B_z = B, \tag{9.169}$$

we see that the Cartesian components of this velocity are $v_x = v$, $v_y = v_z = 0$.

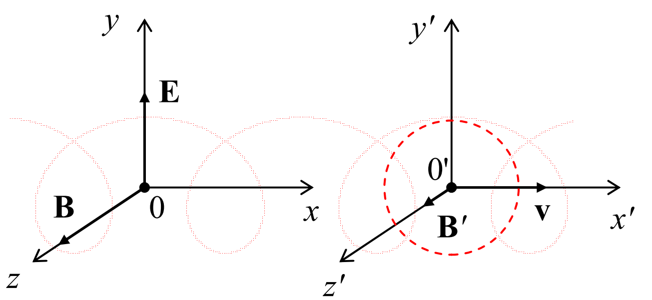


Fig. 9.12. Particle's trajectory in crossed electric and magnetic fields (at $E/c < B$).

Since this choice of the coordinates complies with the one used to derive Eqs. (134), we can readily use that simple form of the Lorentz transform to calculate the field components in the moving reference frame:

$$E'_x = 0, \qquad E'_y = \gamma(E - vB) \equiv \gamma\left(E - \frac{E}{B}B\right) \equiv 0, \qquad E'_z = 0, \tag{9.170}$$

$$B'_x = 0, \qquad B'_y = 0, \qquad B'_z = \gamma\left(B - \frac{vE}{c^2}\right) \equiv \gamma B\left(1 - \frac{vE}{Bc^2}\right) \equiv \gamma B\left(1 - \frac{v^2}{c^2}\right) \equiv \frac{B}{\gamma} \leq B, \tag{9.171}$$

where the Lorentz parameter $\gamma \equiv (1 - v^2/c^2)^{-1/2}$ corresponds to the velocity (168) rather than that of the particle. These relations show that in this special reference frame, the particle only "sees" the re-normalized uniform magnetic field $B' \leq B$, parallel to the initial field, i.e. normal to the velocity (168). Using the result of the above case (i), we see that in this frame the particle moves along either a circle or a spiral winding about the direction of the magnetic field, with the angular velocity (151):

$$\omega'_c = \frac{qB'}{\mathcal{E}'/c^2}, \tag{9.172}$$

and the radius (153):

$$R' = \frac{p'_\perp}{qB'}. \tag{9.173}$$

Hence in the lab frame, the particle performs this orbital/spiral motion plus a "drift" with the constant velocity **v** (Fig. 12). As a result, the lab-frame trajectory of the particle (or rather its projection onto the plane normal to the magnetic field) is a *trochoid*-like curve[58] that, depending on the initial velocity, may be either *prolate* (self-crossing), as in Fig. 12, or *curtate* (drift-stretched so much that it is not self-crossing).

Such looped motion of electrons is used, in particular, in *magnetrons* – very popular generators of microwave radiation. In such a device (Fig. 13), the magnetic field, usually created by specially-shaped permanent magnets, is nearly uniform (in the region of electron motion) and directed along the magnetron's axis (in Fig. 13, normal to the plane of the drawing), while the electric field of magnitude $E \ll cB$, created by the dc voltage applied between the anode and the cathode, is virtually radial.

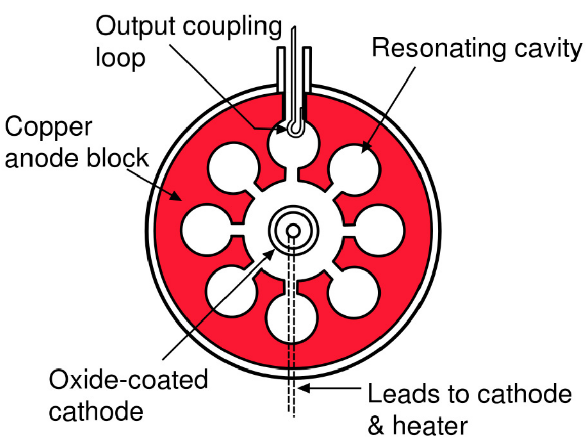


Fig. 9.13. Schematic cross-section of a typical magnetron. (Figure adapted from https://en.wikipedia.org/wiki/Cavity_magnetron under the Free GNU Documentation License.)

As a result, the above simple theory is only approximately valid, and the electron trajectories are close to *epicycloids* rather than trochoids. The applied electric field is adjusted so that these looped trajectories pass close to the anode's surface, and hence to the gap openings of the cylindrical microwave cavities drilled in the anode's bulk. The fundamental mode of such a cavity is quasi-lumped, with the cylindrical walls working mostly as inductances, and the gap openings as capacitances, with the microwave electric field concentrated in these openings. This is why the mode is strongly coupled to the electrons "licking" the anode's surface, and their interaction creates large positive feedback (equivalent to negative damping), which results in intensive microwave self-oscillations at the cavities' own frequency.[59] The oscillation energy, of course, is taken from the dc-field-accelerated electrons; due to this energy loss, the looped trajectory of each electron gradually moves closer to the anode and finally

[58] As a reminder, a trochoid may be described as the trajectory of a point on a rigid disk rolled along a straight line. It's canonical parametric representation is $x = \Theta + a\cos\Theta$, $y = a\sin\Theta$. (For $a > 1$, the trochoid is *prolate*, if $a < 1$, it is *curtate*, and if $a = 1$, it is called the *cycloid*.) Note, however, that for our problem, the trajectory in the lab frame is exactly trochoidal only in the non-relativistic limit $v \ll c$ (i.e. $E/c \ll B$).

[59] See, e.g., CM Sec. 5.4.

lands on its surface. The wide use of such generators (in particular, in microwave ovens, which operate in a narrow frequency band around 2.45 GHz, allocated for these devices to avoid their interference with wireless communication systems) is due to their simplicity and high (up to 65%) efficiency of the dc-to-rf energy transfer.

Case 2: $E/c > B$. In this case, the speed given by Eq. (168) would be above the speed of light, so let us introduce a reference frame moving with a different velocity,

$$\mathbf{v} = \frac{\mathbf{E}\times\mathbf{B}}{(E/c)^2}, \tag{9.174}$$

whose direction is the same as before (Fig. 12), and magnitude $v = c\times B/(E/c)$ is again below c. A calculation absolutely similar to the one performed above for Case 1, yields

$$E'_x = 0, \qquad E'_y = \gamma(E - vB) = \gamma E\left(1 - \frac{vB}{E}\right) = \gamma E\left(1 - \frac{v^2}{c^2}\right) = \frac{E}{\gamma} \leq E, \qquad E'_z = 0, \tag{9.175}$$

$$B'_x = 0, \qquad B'_y = 0, \qquad B'_z = \gamma\left(B - \frac{vE}{c^2}\right) = \gamma\left(B - \frac{EB}{E}\right) = 0. \tag{9.176}$$

so in the moving frame the particle "sees" only the electric field $E' \leq E$. According to the solution of our previous problem (ii), the trajectory of the particle in the moving frame is the catenary (165), so in the lab frame it has an "open", hyperbolic character as well.

To conclude this section, let me note that if the electric and magnetic fields are nonuniform, the particle motion may be much more complex, and in most cases, the integration of the system of equations (144) and (148) may be carried out only numerically. However, if the field's nonuniformity is small, approximate analytical methods may be very effective. For example, if $\mathbf{E} = 0$, and the magnetic field has a small *transverse* gradient ∇B in a direction normal to the vector $\mathbf{B}$ itself, such that

$$\eta \equiv \frac{|\nabla B|}{B} << \frac{1}{R}, \tag{9.177}$$

where R is the cyclotron radius (153), then it is straightforward to use Eq. (150) to show[60] that the cyclotron orbit drifts perpendicular to both $\mathbf{B}$ and ∇B, with the drift speed

$$v_\mathrm{d} \approx \frac{\eta}{\omega_\mathrm{c}}\left(\frac{1}{2}u_\perp^2 + u_{||}^2\right) << u. \tag{9.178}$$

The physics of this drift is rather simple: according to Eq. (153), the instant curvature of the cyclotron orbit is proportional to the local value of the field. Hence if the field is nonuniform, the trajectory bends slightly more on its parts passing through a stronger field, thus acquiring a shape close to a curate trochoid.

For experimental physics and engineering practice, the effects of *longitudinal* gradients of magnetic field on the charged particle motion are much more important, but it is more convenient for me to postpone their discussion until we have developed a little bit more analytical tools in the next section.

---

[60] See, e.g., Sec. 12.4 in J. Jackson, *Classical Electrodynamics*, 3rd ed., Wiley, 1999.

9.7. Analytical mechanics of charged particles

The general Eq. (145) gives a full description of relativistic particle dynamics in electric and magnetic fields, just as the 2nd Newton law (1) does it in the non-relativistic limit. However, we know that in the latter case, the Lagrange formalism of analytical mechanics allows an easier solution of many problems.[61] We can expect that to be true in relativistic mechanics as well, so let us expand the analysis of Sec. 3 (which was valid only for free particles) to particles in the field.

For a free particle, our main result was Eq. (68), which may be rewritten as

$$\gamma \mathscr{L} = -mc^2, \tag{9.179}$$

with $\gamma \equiv (1 - u^2/c^2)^{-1/2}$, showing that the product on the left-hand side is Lorentz-invariant. How can the electromagnetic field affect this relation? In non-relativistic electrostatics, we could write

$$\mathscr{L} = T - U = T - q\phi. \tag{9.180}$$

However, in relativity, the scalar potential ϕ is just one component of the potential 4-vector (116). The only way to get from this full 4-vector a Lorentz-invariant contribution to $\gamma\mathscr{L}$, which would be also proportional to the first power of the particle's velocity (to account for the magnetic component of the Lorentz force), is evidently

$$\gamma \mathscr{L} = -mc^2 + \text{const} \times u^\alpha A_\alpha, \tag{9.181}$$

where u^α is the 4-velocity (63). To comply with Eq. (180) at $u << c$, the constant factor should be equal to $(-q)$, so Eq. (181) becomes

$$\gamma \mathscr{L} = -mc^2 - qu^\alpha A_\alpha, \tag{9.182}$$

and with the account of Eqs. (63) and the second of Eqs. (116), we get very important equality

Particle's Lagrangian function

$$\boxed{\mathscr{L} = -\frac{mc^2}{\gamma} - q\phi + q\mathbf{u}\cdot\mathbf{A},} \tag{9.183}$$

whose Cartesian form is

$$\mathscr{L} = -mc^2\left(1 - \frac{u_x^2 + u_y^2 + u_z^2}{c^2}\right)^{1/2} - q\phi + q\left(u_x A_x + u_y A_y + u_z A_z\right). \tag{9.184}$$

Let us see whether this relation (which admittedly was derived by an educated guess rather than by a strict derivation) passes a natural sanity check. For the case of an unconstrained motion of a particle, we can select its three Cartesian coordinates r_j (j = 1, 2, 3) as the generalized coordinates, and its linear velocity components u_j as the corresponding generalized velocities. In this case, the Lagrange equations of motion are

$$\frac{d}{dt}\frac{\partial \mathscr{L}}{\partial u_j} - \frac{\partial \mathscr{L}}{\partial r_j} = 0. \tag{9.185}$$

For example, for $r_1 = x$, Eq. (184) yields

$$\frac{\partial \mathscr{L}}{\partial u_x} = \frac{mu_x}{\left(1 - u^2/c^2\right)^{1/2}} + qA_x \equiv p_x + qA_x, \qquad \frac{\partial \mathscr{L}}{\partial x} = -q\frac{\partial \phi}{\partial x} + q\mathbf{u}\cdot\frac{\partial \mathbf{A}}{\partial x}, \tag{9.186}$$

[61] See, e.g., CM Sec. 2.2 and on.

so Eq. (185) takes the form

$$\frac{dp_x}{dt} = -q\frac{\partial \phi}{\partial x} + q\mathbf{u}\cdot\frac{\partial \mathbf{A}}{\partial x} - q\frac{dA_x}{dt}. \tag{9.187}$$

In the equations of motion, the field values have to be taken at the instant position of the particle, so the last (full) derivative has components due to both the actual field's change (at a fixed point of space) and the particle's motion. Such addition is described by the so-called *convective derivative*[62]

Convective derivative

$$\boxed{\frac{d}{dt} = \frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla .} \tag{9.188}$$

Spelling out both scalar products, we may group the terms remaining after cancellations as follows:

$$\frac{dp_x}{dt} = q\left[\left(-\frac{\partial \phi}{\partial x} - \frac{\partial A_x}{\partial t}\right) + u_y\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) - u_z\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right)\right]. \tag{9.189}$$

But taking into account the relations (121) between the electric and magnetic fields and potentials, this expression is nothing more than

$$\frac{dp_x}{dt} = q\left(E_x + u_y B_z - u_z B_y\right) = q\left(\mathbf{E} + \mathbf{u}\times\mathbf{B}\right)_x, \tag{9.190}$$

i.e. the x-component of Eq. (144). Since other Cartesian coordinates participate in Eq. (184) similarly, it is evident that the Lagrangian equations of motion along other coordinates yield other components of the same vector equation of motion.

So, Eq. (183) does indeed give the correct Lagrangian function, and we can use it for further analysis, in particular to discuss the first of Eqs. (186). This relation shows that in the electromagnetic field, the generalized momentum corresponding to the particle's coordinate x is *not* $p_x = m\gamma u_x$, but[63]

$$P_x \equiv \frac{\partial \mathscr{L}}{\partial u_x} = p_x + qA_x . \tag{9.191}$$

Thus, as was already discussed (at that point, without proof) in Sec. 6.4, the particle's motion in a magnetic field may be is described by two different linear momentum vectors: the *kinetic momentum* **p** defined by Eq. (70), and the *canonical* (or "conjugate") *momentum*[64]

Particle's canonical momentum

$$\boxed{\mathbf{P} = \mathbf{p} + q\mathbf{A} .} \tag{9.192}$$

In order to facilitate discussion of this notion, let us generalize Eq. (72) for the Hamiltonian function $\mathscr{H}$ of a free particle to the case of a particle in the field:

$$\mathscr{H} \equiv \mathbf{P}\cdot\mathbf{u} - \mathscr{L} = (\mathbf{p} + q\mathbf{A})\cdot\mathbf{u} - \left(-\frac{mc^2}{\gamma} + q\mathbf{u}\cdot\mathbf{A} - q\phi\right) \equiv \mathbf{p}\cdot\mathbf{u} + \frac{mc^2}{\gamma} + q\phi . \tag{9.193}$$

[62] Alternatively called the "Lagrangian derivative"; for its (rather simple) derivation see, e.g., CM Sec. 8.3.
[63] With regrets, I have to use for the generalized momentum the same (very common) notation as was used earlier in the course for the electric polarization – which will not be discussed here and in the balance of these notes.
[64] In the Gaussian units, Eq. (192) has the form $\mathbf{P} = \mathbf{p} + q\mathbf{A}/c$.

Merging the first two terms of the last expression exactly as it was done in Eq. (72), we get an extremely simple result,

$$\mathscr{H} = \gamma mc^2 + q\phi\,, \tag{9.194a}$$

which may be spelled out as

$$\mathscr{H} = \left[1 + \left(\frac{p}{mc}\right)^2\right]^{1/2} mc^2 + q\phi, \qquad \text{i.e. } (\mathscr{H} - q\phi)^2 = (mc^2)^2 + c^2 p^2. \tag{9.194b}$$

These expressions may leave the reader wondering: where is the vector potential **A** here – and the magnetic field effects it has to describe? The resolution of this puzzle is easy: as we know from analytical mechanics,[65] for most applications, for example for an alternative derivation of the equations of motion, $\mathscr{H}$ has to be represented as a function of the particle's generalized coordinates (in the case of unconstrained motion, these may be the Cartesian components of the vector **r** that serves as an argument for the potentials **A** and ϕ), and the generalized momenta, i.e. the components of the vector **P** – generally, plus time. For that, the kinematic momentum **p** in Eq. (194b) has to be expressed via these variables. This may be done using Eq. (192), giving us the following generalization of Eq. (78):[66]

Particle's Hamiltonian function

$$\boxed{(\mathscr{H} - q\phi)^2 = (mc^2)^2 + c^2(\mathbf{P} - q\mathbf{A})^2.} \tag{9.195}$$

It is straightforward to verify that the Hamilton equations of motion for three Cartesian coordinates of the particle, obtained in a regular way from this $\mathscr{H}$, may be merged into the same vector equation (144). In the non-relativistic limit, performing the expansion of Eqs. (194b) into the Taylor series in p^2, and limiting it to two leading terms, we get the following generalization of Eq. (74):

$$\mathscr{H} \approx mc^2 + \frac{p^2}{2m} + q\phi, \qquad \text{i.e. } \mathscr{H} - mc^2 \approx \frac{1}{2m}(\mathbf{P} - q\mathbf{A})^2 + U, \qquad \text{with } U = q\phi. \tag{9.196}$$

These expressions for $\mathscr{H}$, and Eq. (183) for $\mathscr{L}$, give a clear view of the electromagnetic field effects' description in analytical mechanics. The electric part $q\mathbf{E}$ of the total Lorentz force can perform mechanical work on the particle, i.e. change its kinetic energy – see Eq. (148) and its discussion. As a result, the scalar potential ϕ, whose gradient gives a contribution to **E**, may be directly associated with the potential energy $U = q\phi$ of the particle. On the contrary, the magnetic component $q\mathbf{u}\times\mathbf{B}$ of the Lorentz force is always perpendicular to the particle's velocity **u**, and cannot perform a non-zero work on it, and as a result, cannot be described by a contribution to U. However, if **A** did not participate in the functions $\mathscr{L}$ and/or $\mathscr{H}$ at all, the analytical mechanics would be unable to describe effects of the magnetic field $\mathbf{B} = \nabla\times\mathbf{A}$ on the particle's motion. The relations (183) and (195)-(196) show the wonderful way in which physics (with some help from Mother Nature herself :-) solves this problem: the vector potential gives such contributions to the functions $\mathscr{L}$ and $\mathscr{H}$ that cannot be uniquely attributed to either kinetic or potential energy, but ensure both the Lagrange and Hamilton formalisms yield the correct equation of motion (144) – including the magnetic field effects.

[65] See, e.g., CM Sec. 10.1.

[66] Alternatively, this relation may be obtained from the expression for the Lorentz-invariant norm, $p^\alpha p_\alpha = (mc)^2$, of the 4-momentum (75), $p^\alpha = \{\mathscr{E}/c, \mathbf{p}\} = \{(\mathscr{H} - q\phi)/c, \mathbf{P} - q\mathbf{A}\}$.

I believe I still owe the reader some discussion of the physical sense of the canonical momentum **P**. For that, let us consider a charged particle moving near a region of localized magnetic field **B**(**r**,*t*), but not entering this region (see Fig. 14), so on its trajectory $\nabla\times\mathbf{A} \equiv \mathbf{B} = 0$.

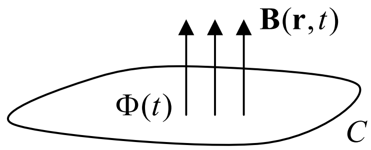


Fig. 9.14. Particle's motion around a localized magnetic field with a time-dependent flux.

If there is no electrostatic field affecting the particle (i.e. no other electric charges nearby), we may select such a local gauge that $\phi(\mathbf{r}, t) = 0$ and $\mathbf{A} = \mathbf{A}(t)$, so Eq. (144) is reduced to

$$\frac{d\mathbf{p}}{dt} = q\mathbf{E} = -q\frac{d\mathbf{A}}{dt}, \tag{9.197}$$

and Eq. (192) immediately gives

$$\frac{d\mathbf{P}}{dt} \equiv \frac{d\mathbf{p}}{dt} + q\frac{d\mathbf{A}}{dt} = 0. \tag{9.198}$$

Hence, even if the magnetic field is changed in time, so that the induced electric field **E** does accelerate the particle, its canonical momentum does not change. Hence **P** is a variable more stable to magnetic field changes than its kinetic counterpart **p**. This conclusion may be criticized because it relies on a specific gauge, and generally $\mathbf{P} \equiv \mathbf{p} + q\mathbf{A}$ is not gauge–invariant, because the vector potential **A** is not.[67] However, as was already discussed in Sec. 5.3, the integral $\oint \mathbf{A}\cdot d\mathbf{r}$ over a closed contour *is* gauge-invariant and is equal to the magnetic flux Φ through the area limited by the contour – see Eq. (5.65). So, integrating Eq. (197) over a closed trajectory of a particle (Fig. 14), and over the time of one orbit, we get

$$\Delta\oint_C \mathbf{p}\cdot d\mathbf{r} = -q\Delta\Phi, \qquad \text{so that} \quad \Delta\oint_C \mathbf{P}\cdot d\mathbf{r} = 0, \tag{9.199}$$

where ΔΦ is the change of flux during that time. This gauge-invariant result confirms the above conclusion about the stability of the canonical momentum to magnetic field variations.

Generally, Eq. (199) is invalid if a particle moves inside a magnetic field and/or changes its trajectory at the field variation. However, if the field is *almost* uniform, i.e. its gradient is small in the sense of Eq. (177), this result is (approximately) applicable. Indeed, analytical mechanics[68] tells us that for any canonical coordinate-momentum pair $\{q_j, p_j\}$, the corresponding *action variable*,

$$J_j \equiv \frac{1}{2\pi}\oint p_j dq_j, \tag{9.200}$$

remains virtually constant at slow variations of motion conditions. According to Eq. (191), for a particle in a magnetic field, the generalized momentum corresponding to the Cartesian coordinate r_j is P_j rather than p_j. Thus forming the net action variable $J \equiv J_x + J_y + J_z$, we may write

[67] In contrast, the kinetic momentum $\mathbf{p} = M\mathbf{u}$ is evidently gauge- (though not Lorentz-) invariant.
[68] See, e.g., CM Sec. 10.2.

$$2\pi J = \oint \mathbf{P} \cdot d\mathbf{r} = \oint \mathbf{p} \cdot d\mathbf{r} + q\Phi = \text{const}\,. \tag{9.201}$$

Let us apply this relation to the motion of a non-relativistic particle in an almost uniform magnetic field, with a relatively small longitudinal velocity, $u_{||}/u_{\perp} \to 0$ – see Fig. 15.

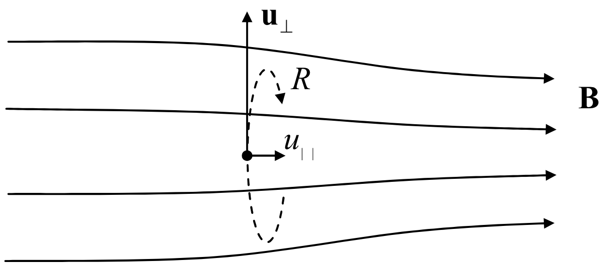


Fig. 9.15. Particle in a magnetic field with a small longitudinal gradient $\nabla B \,||\, \mathbf{B}$.

In this case, Φ in Eq. (201) is the flux encircled by the particle's cyclotron orbit, $\Phi = -\pi R^2 B$, where R is its radius given by Eq. (153), and the negative sign accounts for the fact that in our case, the "correct" direction of the normal vector $\mathbf{n}$ in the definition of flux, $\Phi = \int \mathbf{B}\cdot\mathbf{n}d^2r$, is antiparallel to the vector $\mathbf{B}$. At $u << c$, the kinetic momentum is just $p_\perp = mu_\perp$, while Eq. (153) yields

$$mu_\perp = qBR\,. \tag{9.202}$$

Plugging these relations into Eq. (201), we get

$$2\pi J = mu_\perp 2\pi R - q\pi R^2 B = m\frac{qRB}{m}2\pi R - q\pi R^2 B \equiv (2-1)q\pi R^2 B \equiv -q\Phi\,. \tag{9.203}$$

This means that even if the circular orbit slowly moves through the magnetic field, the flux encircled by the cyclotron orbit should remain virtually constant. One manifestation of this effect is the result already mentioned at the end of Sec. 6: if a small gradient of the magnetic field is perpendicular to the field itself, then the particle orbit's drift direction is perpendicular to ∇B, so Φ stays constant.

Now let us analyze the case of a small longitudinal gradient, $\nabla B \,||\, \mathbf{B}$ (Fig. 15). If a small initial longitudinal velocity $u_{||}$ is directed toward the higher field region, the cyclotron orbit has to gradually shrink to keep Φ constant. Rewriting Eq. (202) as

$$mu_\perp = q\frac{\pi R^2 B}{\pi R} = q\frac{|\Phi|}{\pi R}\,, \tag{9.204}$$

we see that this reduction of R (at constant Φ) increases the orbiting speed $u_\perp$. But since the magnetic field cannot perform any work on the particle, its kinetic energy,

$$\mathscr{E} = \frac{m}{2}\left(u_{||}^2 + u_\perp^2\right), \tag{9.205}$$

should stay constant, so the longitudinal velocity $u_{||}$ has to decrease. Hence eventually the orbit's drift has to stop, and then it has to start moving back toward the region of lower fields, being essentially repulsed from the high-field region. This effect is very important, in particular, for plasma confinement systems. In the simplest of such systems, two coaxial magnetic coils, inducing magnetic fields of the same direction (Fig. 16), naturally form a "magnetic bottle", which traps charged particles injected, with sufficiently low longitudinal velocities, into the region between the coils. More complex systems of this type, but working on the same basic principle, are the most essential components of the persisting large-scale efforts to achieve controllable nuclear fusion.[69]

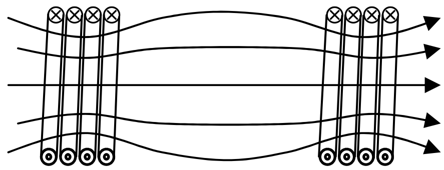


Fig. 9.16. A simple magnetic bottle (schematically).

Returning to the constancy of the magnetic flux encircled by free particles, it reminds us of the Meissner-Ochsenfeld effect, which was discussed in Sec. 6.4, and gives a motivation for a brief revisit of the electrodynamics of superconductivity. As was emphasized in that section, superconductivity is a substantially quantum phenomenon; nevertheless, the classical notion of the conjugate momentum **P** helps to understand its theoretical description. Indeed, the general rule of quantization of physical systems[70] is that each canonical pair $\{q_j, p_j\}$ of a generalized coordinate q_j and the corresponding generalized momentum p_j is described by quantum-mechanical operators that obey the following commutation relation:

$$\left[\hat{q}_j, \hat{p}_{j'}\right] = i\hbar\delta_{jj'} . \tag{9.206}$$

According to Eq. (191), for the Cartesian coordinates r_j of a particle in the magnetic field, the corresponding generalized momenta are P_j, so their operators should obey the similar commutation relations:

$$\left[\hat{r}_j, \hat{P}_{j'}\right] = i\hbar\delta_{jj'} . \tag{9.207}$$

In the coordinate representation of quantum mechanics, the canonical operators of the Cartesian components of the linear momentum are described by the corresponding components of the vector operator $-i\hbar\nabla$. As a result, ignoring the rest energy mc^2 (which gives an inconsequential phase factor $\exp\{-imc^2t/\hbar\}$ in the wavefunction), we can use Eq. (196) to rewrite the usual non-relativistic Schrödinger equation,

$$i\hbar\frac{\partial\psi}{\partial t} = \hat{\mathscr{H}}\psi , \tag{9.208}$$

as follows:

$$i\hbar\frac{\partial\psi}{\partial t} = \left(\frac{\hat{p}^2}{2m} + U\right)\psi \equiv \left[\frac{1}{2m}\left(-i\hbar\nabla - q\mathbf{A}\right)^2 + q\phi\right]\psi . \tag{9.209}$$

Thus, I believe I have finally delivered on my promise to justify the replacement (6.50), which had been used in Secs. 6.4 and 6.5 to discuss the electrodynamics of superconductors, including the Meissner-Ochsenfeld effect. The Schrödinger equation (209) may be also used as the basis for the quantum-mechanical description of other magnetic field phenomena, including the so-called Aharonov-Bohm and quantum Hall effects – see, e.g., QM Secs. 3.1-3.2.

[69] For further reading on this technology, the reader may be referred, for example, to the simple monograph by F. Chen, *Introduction to Plasma Physics and Controllable Fusion*, vol. 1, 2nd ed., Springer, 1984, and/or the graduate-level theoretical treatment by R. Hazeltine and J. Meiss, *Plasma Confinement*, Dover, 2003.

[70] See, e.g., CM Sec. 10.1.

9.8. Analytical mechanics of the electromagnetic field

We have just seen that the analytical mechanics of a *particle* in an electromagnetic field may be used to get some important results. The same is true for the analytical mechanics of the electromagnetic *field* as such, and the *field-particle system* as a whole. For such space-distributed systems as fields, governed by local dynamics laws (in our case, the Maxwell equations), we need to apply analytical mechanics to the *local densities* $\mathcal{l}$ and $\mathcal{h}$ of the Lagrangian and Hamiltonian functions, defined by relations

$$\mathscr{L} = \int \mathcal{l}\, d^3 r, \quad \mathscr{H} = \int \mathcal{h}\, d^3 r. \tag{9.210}$$

Let us start, as usual, from the Lagrange formalism. Some clues on the possible structure of the Lagrangian function density $\mathcal{l}$ may be obtained from that of the particle-field interaction description in this formalism, discussed in the last section. As we have seen, for the case of a single particle, the interaction is described by the last two terms of Eq. (183):

$$\mathscr{L}_{\text{int}} = -q\phi - q\mathbf{u}\cdot\mathbf{A}. \tag{9.211}$$

Obviously, if the charge q is continuously distributed over some volume, we may represent this $\mathscr{L}_{\text{int}}$ as a volume integral of the following Lagrangian function density:

$$\boxed{\mathcal{l}_{\text{int}} = -\rho\phi + \mathbf{j}\cdot\mathbf{A} \equiv -j_\alpha A^\alpha.} \tag{9.212}$$

Interaction Lagrangian density

Notice that this density (in contrast to $\mathscr{L}_{\text{int}}$ itself!) is Lorentz-invariant. (This is due to the contraction of the longitudinal coordinate, and hence volume, at the Lorentz transform.) Hence we may expect the density of the *field's* part of the Lagrangian to be Lorentz-invariant as well. Moreover, given the local structure of the Maxwell equations (containing only the first spatial and temporal derivatives of the fields), $\mathcal{l}_{\text{field}}$ should be a function of the potential's 4-vector and its 4-derivative:

$$\mathcal{l}_{\text{field}} = \mathcal{l}_{\text{field}}\left(A^\alpha, \partial_\alpha A^\beta\right). \tag{9.213}$$

Also, the density should be selected in such a way that the 4-vector analog of the Lagrangian equation of motion,

$$\partial_\alpha \frac{\partial \mathcal{l}_{\text{field}}}{\partial\left(\partial_\alpha A^\beta\right)} - \frac{\partial \mathcal{l}_{\text{field}}}{\partial A^\beta} = 0, \tag{9.214}$$

gave us the correct inhomogeneous Maxwell equations (127).[71] The field part $\mathcal{l}_{\text{field}}$ of the total Lagrangian density $\mathcal{l}$ should be a scalar and a quadratic form of the field strengths, i.e. of the tensor $F^{\alpha\beta}$, so the natural choice is

$$\mathcal{l}_{\text{field}} = \text{const} \times F_{\alpha\beta}F^{\alpha\beta}. \tag{9.215}$$

with the implied summation over both indices. Indeed, adding to this expression the interaction Lagrangian (212),

$$\mathcal{l} = \mathcal{l}_{\text{field}} + \mathcal{l}_{\text{int}} = \text{const} \times F_{\alpha\beta}F^{\alpha\beta} - j_\alpha A^\alpha, \tag{9.216}$$

[71] Here the implicit summation over the index α plays a role similar to the convective derivative (188) in replacing the full derivative over time, in a way that reflects the symmetry of time and space in special relativity. I do not want to spend more time justifying Eq. (214), because of the reasons that will be clear imminently.

and performing the differentiations, we see that Eqs. (214)-(215) indeed yield Eqs. (127), provided that the constant factor equals $(-1/4\mu_0)$.[72] So, the field's Lagrangian density is

Field's Lagrangian density

$$\boxed{\ell_{\text{field}} = -\frac{1}{4\mu_0} F_{\alpha\beta} F^{\alpha\beta} = \frac{1}{2\mu_0}\left(\frac{E^2}{c^2} - B^2\right) \equiv \frac{\varepsilon_0}{2} E^2 - \frac{B^2}{2\mu_0} \equiv u_e - u_m,} \tag{9.217}$$

where u_e is the electric field energy density (1.65), and u_m is the magnetic field energy density (5.57). Let me hope the reader agrees that Eq. (217) is a wonderful result because the Lagrangian function has a structure absolutely similar to the well-known expression $\mathscr{L} = T - U$ of classical mechanics. So, for the field alone, the "potential" and "kinetic" energies are separable again.[73]

Now let us explore whether we can calculate the 4-form of the field's Hamiltonian function $\mathscr{H}$. In the generic analytical mechanics,

$$\mathscr{H} = \sum_j \frac{\partial \mathscr{L}}{\partial \dot{q}_j} \dot{q}_j - \mathscr{L}. \tag{9.218}$$

However, just as for the Lagrangian function, for a field we should find the spatial density $\hbar$ of the Hamiltonian, defined by the second of Eqs. (210), for which the natural 4-form of Eq. (218) is

$$\hbar^{\alpha\beta} = \frac{\partial \ell}{\partial(\partial_\alpha A^\gamma)} \partial^\beta A^\gamma - g^{\alpha\beta} \ell. \tag{9.219}$$

Calculated for the field alone, i.e. using Eq. (217) for ℓ, this definition yields

$$\hbar^{\alpha\beta}_{\text{field}} = \theta^{\alpha\beta} - \tau_D^{\alpha\beta}, \tag{9.220}$$

where the tensor

Symmetric energy-momentum tensor

$$\boxed{\theta^{\alpha\beta} \equiv \frac{1}{\mu_0}\left(g^{\alpha\gamma} F_{\gamma\delta} F^{\delta\beta} + \frac{1}{4} g^{\alpha\beta} F_{\gamma\delta} F^{\gamma\delta}\right),} \tag{9.221}$$

is gauge-invariant, while the remaining term,

$$\tau_D^{\alpha\beta} \equiv \frac{1}{\mu_0} g^{\alpha\gamma} F_{\gamma\delta} \partial^\delta A^\beta, \tag{9.222}$$

is not, so it cannot correspond to any measurable variables. Fortunately, it is straightforward to verify that the last tensor may be represented in the form

$$\tau_D^{\alpha\beta} = \frac{1}{\mu_0} \partial_\gamma \left(F^{\gamma\alpha} A^\beta\right), \tag{9.223}$$

and as a result, obeys the following relations:

$$\partial_\alpha \tau_D^{\alpha\beta} = 0, \qquad \int \tau_D^{0\beta} d^3 r = 0, \tag{9.224}$$

[72] In the Gaussian units, this coefficient is $(-1/16\pi)$.

[73] Since the Lagrange equations of motion are homogeneous, the simultaneous change of the signs of T and U does not change them. Thus, it is not important which of the two energy densities, u_e or u_m, we count as the potential energy, and which as the kinetic energy. (Actually, such duality of the two energy components is typical for all analytical mechanics – see, e.g., the discussion of this issue in CM Sec. 2.2.)

so it does not interfere with the conservation properties of the gauge-invariant, symmetric *energy-momentum tensor* (also called the *symmetric stress tensor*) $\theta^{\alpha\beta}$, to be discussed below.

Let us use Eqs. (125) to express the elements of the latter tensor via the electric and magnetic fields. For $\alpha = \beta = 0$, we get

$$\theta^{00} = \frac{\varepsilon_0}{2}E^2 + \frac{B^2}{2\mu_0} = u_\mathrm{e} + u_\mathrm{m} \equiv u\,, \qquad (9.225)$$

i.e. the expression for the total energy density u – see Eq. (6.113). The other 3 elements of the same row/column turn out to be just the Cartesian components of the Poynting vector (6.114), divided by c:

$$\theta^{j0} = \frac{1}{\mu_0}\left(\frac{\mathbf{E}}{c}\times\mathbf{B}\right)_j = \left(\frac{\mathbf{E}}{c}\times\mathbf{H}\right)_j \equiv \frac{S_j}{c}, \qquad \text{for } j = 1, 2, 3\,. \qquad (9.226)$$

The remaining 9 elements $\theta_{jj'}$ of the tensor, with $j,\ j' = 1, 2, 3$, are usually represented as

$$\theta^{jj'} = -\tau^{(\mathrm{M})}_{jj'}\,, \qquad (9.227)$$

where $\tau^{(\mathrm{M})}$ is the so-called *Maxwell stress tensor*:

$$\boxed{\tau^{(\mathrm{M})}_{jj'} = \varepsilon_0\left(E_jE_{j'} - \frac{\delta_{jj'}}{2}E^2\right) + \frac{1}{\mu_0}\left(B_jB_{j'} - \frac{\delta_{jj'}}{2}B^2\right),} \qquad (9.228)$$

Maxwell stress tensor

so the whole symmetric energy-momentum tensor (221) may be conveniently represented in the following symbolic way:

$$\theta^{\alpha\beta} = \left(\begin{array}{c|c} u & \leftarrow \quad \mathbf{S}/c \quad \rightarrow \\ \hline \uparrow & \\ \mathbf{S}/c & -\tau^{(\mathrm{M})}_{jj'} \\ \downarrow & \end{array}\right). \qquad (9.229)$$

The physical meaning of this tensor may be revealed in the following way. Considering Eq. (221) as the *definition* of the tensor $\theta^{\alpha\beta}$,[74] and using the 4-vector form of Maxwell equations given by Eqs. (127) and (129), it is straightforward to verify an extremely simple result for the 4-derivative of the symmetric tensor:

$$\partial_\alpha\theta^{\alpha\beta} = -F^{\beta\gamma}j_\gamma\,. \qquad (9.230)$$

This expression is valid in the presence of electromagnetic field sources, e.g., for any system of charged particles and the fields they have created. Of these four equations (for four values of the index β), the temporal one (with $\beta = 0$) may be simply expressed via the energy density (225) and the Poynting vector (226):

$$\frac{\partial u}{\partial t} + \nabla\cdot\mathbf{S} = -\mathbf{j}\cdot\mathbf{E}\,, \qquad (9.231)$$

while three spatial equations (with $\beta = j = 1, 2, 3$) may be represented in the form

[74] In this way, we are using Eq. (219) just as a useful guess, which has led us to the definition of $\theta^{\alpha\beta}$, and may leave its strict justification for more in-depth field theory courses.

$$\frac{\partial}{\partial t}\frac{S_j}{c^2} - \sum_{j'=1}^{3}\frac{\partial}{\partial r_{j'}}\tau_{jj'}^{(\mathrm{M})} = -(\rho\mathbf{E} + \mathbf{j}\times\mathbf{B})_j . \qquad (9.232)$$

If integrated over a volume V limited by surface S, with the account of the divergence theorem, Eq. (231) returns us to the Poynting theorem (6.111):

$$\int_V \left(\frac{\partial u}{\partial t} + \mathbf{j}\cdot\mathbf{E}\right) d^3r + \oint_S S_n d^2r = 0, \qquad (9.233)$$

while Eq. (232) yields[75]

$$\int_V \left[\frac{\partial}{\partial t}\frac{\mathbf{S}}{c^2} + \mathbf{f}\right]_j d^3r = \sum_{j'=1}^{3}\oint_S \tau_{jj'}^{(\mathrm{M})} dA_{j'}, \quad \text{with } \mathbf{f} \equiv \rho\mathbf{E} + \mathbf{j}\times\mathbf{B}, \qquad (9.234)$$

where $dA_j = n_j dA = n_j d^2r$ is the j^{th} component of the elementary area vector $d\mathbf{A} = \mathbf{n}dA = \mathbf{n}d^2\mathrm{r}$ that is normal to the volume's surface, and directed out of the volume – see Fig. 17.[76]

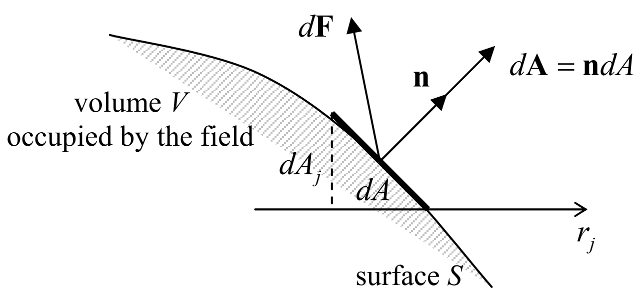


Fig. 9.17. The force $d\mathbf{F}$ exerted on a boundary element $d\mathbf{A}$ of the volume V occupied by the field.

Since, according to Eq. (5.10), the vector $\mathbf{f}$ in Eq. (234) is nothing other than the density of volume-distributed Lorentz forces exerted by the field on the charged particles, we can use the 2nd Newton law, in its relativistic form (144), to rewrite Eq. (234), for a stationary volume V, as

Field momentum's dynamics

$$\boxed{\frac{d}{dt}\left[\int_V \frac{\mathbf{S}}{c^2} d^3r + \mathbf{p}_{\text{part}}\right] = \mathbf{F},} \qquad (9.235)$$

where $\mathbf{p}_{\text{part}}$ is the total mechanical (relativistic) momentum of all particles in the volume V, and the vector $\mathbf{F}$ is defined by its Cartesian components:

Force via the Maxwell tensor

$$\boxed{F_j = \sum_{j'=1}^{3}\oint_S \tau_{jj'}^{(\mathrm{M})} dA_{j'} .} \qquad (9.236)$$

Relations (235)-(236) are our main new results. The first of them shows that the vector

[75] Just like the Poynting theorem (233), Eq. (234) may be obtained directly from the Maxwell equations, without resorting to the 4-vector formalism – see, e.g., Sec. 8.2.2 in D. Griffiths, *Introduction to Electrodynamics*, 3rd ed., Prentice-Hall, 1999. However, the derivation discussed above is superior because it shows the wonderful unity between the laws of conservation of energy and momentum.

[76] The same notions are used in the mechanical stress theory – see, e.g., CM Sec. 7.2.

$$\mathbf{g} \equiv \frac{\mathbf{S}}{c^2}, \tag{9.237}$$

already discussed in Sec. 6.8 without derivation, may be indeed interpreted as the density of momentum of the electromagnetic field (per unit volume). This classical relation is consistent with the quantum-mechanical picture of photons as ultra-relativistic particles, with a momentum of magnitude $\mathscr{E}/c$, because then the flux of the momentum carried by photons through a unit normal area per unit time may be represented either as S_n/c or as g_nc. It also allows us to revisit the Poynting vector paradox that was discussed in Sec. 6.8 – see Fig. 611 and its discussion. As was emphasized in that discussion, in this case, the vector $\mathbf{S} = \mathbf{E}\times\mathbf{H}$ does not correspond to any measurable energy flow. However, the corresponding momentum of the field, equal to the integral of the density (237) over a volume of interest,[77] is not only real but may be measured by the recoil impulse it gives to the field sources – say, to the magnetic coil inducing the field $\mathbf{H}$, or to the capacitor plates creating the field $\mathbf{E}$.

Now let us turn to our second result, Eq. (236). It tells us that the 3×3-element Maxwell stress tensor complies with the general definition of the stress tensor[78] characterizing the forces exerted on the boundaries of a volume, in our current case the volume occupied by the electromagnetic field (Fig. 17). Let us use this important result to analyze two simple examples of static fields.

(i) *Electrostatic field's* effect on a perfect conductor. Since Eq. (235) has been derived for a free space region, we have to select volume V outside the conductor, but we may align one of its faces with the conductor's surface (Fig. 18).

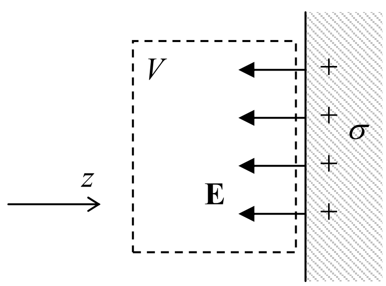


Fig. 9.18. The electrostatic field near a conductor's surface.

From Chapter 2, we know that the electrostatic field just outside the conductor's surface has to be normal to it. Selecting the z-axis in this direction, we have $E_x = E_y = 0$, $E_z = \pm E$, so only diagonal elements of the tensor (228) are not equal to zero:

$$\tau_{xx}^{(\mathrm{M})} = \tau_{yy}^{(\mathrm{M})} = -\frac{\varepsilon_0}{2}E^2, \qquad \tau_{zz}^{(\mathrm{M})} = \frac{\varepsilon_0}{2}E^2, \tag{9.238}$$

Since the elementary surface area vector has just one non-zero component, dA_z, according to Eq. (236), only the last component (that is positive regardless of the sign of E) gives a contribution to the surface force $\mathbf{F}$. We see that the force exerted *by the conductor* (and eventually by the external forces that hold the conductor in its equilibrium position) on the field is normal to the conductor and directed out of the field volume: $dF_z \geq 0$. Hence, by the 3rd Newton law, the force exerted *by the field* on the conductor's surface is directed toward the field-filled space:

[77] It is sometimes called *hidden momentum*.

[78] See, e.g., CM Sec. 7.2.

Electric field's pull

$$dF_{\text{surface}} = -dF_z = -\frac{\varepsilon_0}{2}E^2 dA\,. \tag{9.239}$$

This important result could be obtained by simpler means as well. (Actually, this was the task of one of the exercise problems assigned in Chapter 2.) For example, one could argue, quite convincingly, that the local relation between the force and the field should not depend on the global configuration creating the field, and thus consider the simplest configuration, a planar capacitor (see, e.g. Fig. 2.3) with surfaces of both plates charged by equal and opposite charges of density $\sigma = \pm\varepsilon_0 E$. According to the Coulomb law, the charges should attract each other, pulling each plate toward the field region, so the Maxwell-tensor result gives the correct direction of the force. Now the force's magnitude given by Eq. (239) may be verified either by the direct integration of the Coulomb law or by the following simple reasoning. In the plane capacitor, the inner field $E_z = \sigma/\varepsilon_0$ is equally contributed by two surface charges; hence the field created by the negative charge of the counterpart plate (not shown in Fig. 18) is $E_- = -\sigma/2\varepsilon_0$, and the force it exerts of the elementary surface charge $dQ = \sigma dA$ of the positively charged plate is $dF_{\text{surface}} = EdQ = -\sigma^2 dA/2\varepsilon_0 = \varepsilon_0 E^2 dA/2$, in accordance with Eq. (239).[79]

Quantitatively, even for such a high electric field as $E = 10^5$ V/m (close to the electric breakdown's threshold in the air at a frequency of 10 GHz[80]), the "negative pressure" (dF/dA) given by Eq. (239) is of the order of 0.05 Pa (N/m^2), i.e. many orders below the ambient atmospheric pressure of 1 bar $\approx 10^5$ Pa. Still, this negative pressure may be substantial (well above 1 bar) in some cases, for example in good dielectrics (such as the high-quality SiO_2 grown at high temperature, which is broadly used in integrated circuits), which can withstand electric fields up to $\sim 10^9$ V/m.

(ii) *Static magnetic field's* effect on its source[81] – say a solenoid's wall or a superconductor's surface (Fig. 19). With the Cartesian coordinates' choice shown in that figure, we have $B_x = B$, $B_y = B_z = 0$, so the Maxwell stress tensor (228) is diagonal again:

$$\tau_{xx}^{(\text{M})} = \frac{1}{2\mu_0}B^2, \qquad \tau_{yy}^{(\text{M})} = \tau_{zz}^{(\text{M})} = -\frac{1}{2\mu_0}B^2\,. \tag{9.240}$$

However, since for this geometry, only dA_z differs from 0 in Eq. (236), the sign of the resulting force is opposite to that in electrostatics: $dF_z \leq 0$, and the force exerted by the magnetic field upon the conductor's surface,

Magnetic field's push

$$dF_{\text{surface}} = -dF_z = \frac{1}{2\mu_0}B^2 dA\,, \tag{9.241}$$

[79] By the way, repeating these arguments for a plane capacitor filled with a linear dielectric, we may readily see that Eq. (239) may be generalized for this case by replacing ε_0 with ε. A similar replacement ($\mu_0 \to \mu$) is valid for Eq. (241) in a linear magnetic medium.

[80] Note that the breakdown field E_t in is a strong function of frequency. In the ambient air, it drops from its dc value of $\sim 3\times10^6$ V/m to $\sim 1.5\times10^5$ V/m at microwave frequencies and then rises to as much as $\sim 6\times10^9$ V/m at optical frequencies. The reason of the rise is that at very high frequencies, the amplitude of the field-induced oscillations of the rare free electrons becomes much smaller than their mean free path, inhibiting the bulk impact-ionization of neutral atoms. (Because of this reason, E_t also depends on the air's pressure.)

[81] The causal relation is not important here. Especially in the case of a superconductor, the magnetic field may be induced by another source, with the surface supercurrent **j** just shielding the superconductor's bulk from its penetration – see Sec. 6.

corresponds to positive pressure. For good laboratory magnets ($B \sim 10$ T), this pressure is of the order of 4×10^7 Pa $\approx$ 400 bars, i.e. is very substantial, so the magnets require solid mechanical design.

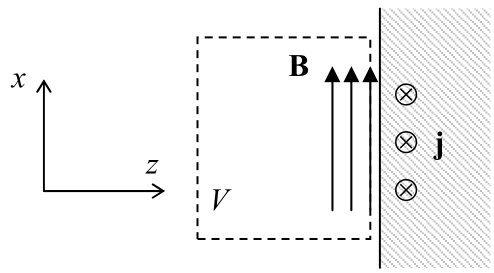


Fig. 9.19. The magnetostatic field near a current-carrying surface.

The direction of the force (241) could be also readily predicted using elementary magnetostatics arguments. Indeed, we can imagine the magnetic field volume limited by another, parallel wall with the opposite direction of surface current. According to the starting point of magnetostatics, Eq. (5.1), such surface currents of opposite directions have to repulse each other – doing that via the magnetic field.

Another explanation of the fundamental sign difference between the electric and magnetic field pressures may be provided using the electric circuit language. As we know from Chapter 2, the potential energy of the electric field stored in a capacitor may be represented in two equivalent forms,

$$U_\mathrm{e} = \frac{CV^2}{2} = \frac{Q^2}{2C}. \tag{9.242}$$

Similarly, the magnetic field energy of an inductive coil is

$$U_\mathrm{m} = \frac{LI^2}{2} = \frac{\Phi^2}{2L}. \tag{9.243}$$

If we do not want to consider the work of external sources at a virtual change of the system dimensions, we should use the last forms of these relations, i.e. consider a galvanically detached capacitor (Q = const) and an externally-shorted inductance (Φ = const).[82] Now if we let the electric field forces (239) drag the capacitor's plates in the direction they "want", i.e. toward each other, this would lead to a *reduction* of the capacitor thickness, and hence to an *increase* of its capacitance C, and hence to a *decrease* of U_e. Similarly, for a solenoid, allowing the positive pressure (241) to move its walls from each other would lead to an *increase* of the solenoid's volume, and hence of its inductance L, so the potential energy U_m would be also *reduced* – as it should be. It is remarkable (actually, beautiful!) how the local field formulas (239) and (241) "know" about these global circumstances.

Finally, let us see whether the major results (237) and (241) obtained in this section, match each other. For that, let us return to the normal incidence of a plane, monochromatic wave from the free space upon the plane surface of a perfect conductor (see, e.g., Fig. 7.8 and its discussion), and use those results to calculate the time average of the pressure dF_surface/dA imposed by the wave on the surface. At elastic reflection from the conductor's surface, the electromagnetic field's momentum retains its amplitude but reverses its sign, so the average momentum transferred to a unit area of the surface in a unit time (i.e. the average pressure) is

[82] Of course, this condition may hold "forever" only for solenoids with superconducting wiring, but even in normal-metal solenoids with practicable inductances, the flux relaxation constants L/R may be rather large (practically, up to a few minutes), quite sufficient to carry out the force measurement.

$$\overline{\frac{dF_{\text{surface}}}{dA}} = 2c g_{\text{incident}} = 2c\frac{\overline{S_{\text{incident}}}}{c^2} = 2c\frac{\overline{EH}}{c^2} = E_\omega H_\omega^*, \tag{9.244}$$

where E_ω and H_ω are complex amplitudes of the incident wave. Using the relation (7.7) between these amplitudes (for $\varepsilon = \varepsilon_0$ and $\mu = \mu_0$ giving $E_\omega = cB_\omega$), we get

$$\overline{\frac{dF_{\text{surface}}}{dA}} = \frac{1}{c}cB_\omega\frac{B_\omega^*}{\mu_0} \equiv \frac{|B_\omega|^2}{\mu_0}. \tag{9.245}$$

On the other hand, as was discussed in Sec. 7.3, at the surface of a perfect mirror the electric field vanishes while the magnetic field doubles, so we can use Eq. (241) with $B \to B(t) = 2\text{Re}[B_\omega \exp\{-i\omega t\}]$. Averaging the pressure given by Eq. (241) over time, we get

$$\overline{\frac{dF_{\text{surface}}}{dA}} = \frac{1}{2\mu_0}\overline{\left(2\,\text{Re}\left[B_\omega e^{-i\omega t}\right]\right)^2} = \frac{|B_\omega|^2}{\mu_0}, \tag{9.246}$$

i.e. the same result as Eq. (245).

For physics intuition development, it is useful to evaluate the electromagnetic radiation pressure. Even for a relatively high wave intensity S_n of 1 kW/m^2 (close to that of the direct sunlight at the Earth's surface), the pressure $2cg_n = 2S_n/c$ is somewhat below 10^{-5} Pa ~ 10^{-10} bar. Still, this extremely small effect was experimentally observed (by P. Lebedev) as early as 1899, giving one more confirmation of Maxwell's theory. Currently, there are ongoing attempts to use the pressure of the Sun's light for propelling small spacecraft, e.g., the *LightSail 2* satellite with a 32-m^2 sail, launched in 2019.

9.9. Exercise problems

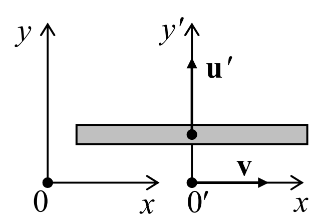
9.1. Use the pre-relativistic picture of light propagation with velocity c in a Sun-bound aether to derive Eq. (4).

9.2. Show that two successive Lorentz space/time transforms, with velocities u' and v in the same direction, are equivalent to the single transform with the velocity u given by Eq. (25).

9.3. $N + 1$ reference frames numbered by index n (taking values 0, 1, …, N) move in the same direction as a particle. Express the particle's velocity in the frame number 0 via its velocity u_N in the frame number N and the velocities v_n of the frame number n relative to the frame number $(n - 1)$.

9.4. A spaceship moving with a constant velocity v directly from the Earth sends back brief flashes of light with a period Δt_s – as measured by the spaceship's clock. Calculate the period with that an Earth-based observer may receive these signals – as measured by their clock.

9.5. From the point of view of observers in a reference frame 0′, a straight thin rod, parallel to the x'-axis, is moving without rotation with a constant velocity $\mathbf{u}'$ directed along the y'-axis. The reference frame 0′ is itself moving relative to another ("lab") reference frame 0 with a constant velocity $\mathbf{v}$ along the x-axis, also without rotation – see the figure on the right.



Calculate:

(i) the direction of the rod's velocity, and
(ii) the orientation of the rod on the [x, y] plane,

– both as observed from the lab reference frame. Is the velocity, in this frame, perpendicular to the rod?

9.6. Starting from the rest at $t = 0$, a spaceship moves directly from the Earth, with a constant acceleration as measured in its *instantaneous rest frame*. Find its displacement $x(t)$ from the Earth, as measured from the Earth's reference frame, and interpret the result.

Hint: The instantaneous rest frame of a moving particle is the inertial reference frame that, at the considered moment of time, has the same velocity as the particle.

9.7. Analyze the twin paradox for the simplest case of 1D travel with a piecewise-constant acceleration.

Hint: You may use an intermediate result of the solution of the previous problem.

9.8. Suggest a natural definition of the 4-vector of acceleration (commonly called the *4-acceleration*) of a point and calculate its components for of a relativistic point moving with velocity $\mathbf{u} = \mathbf{u}(t)$.

9.9. Calculate the first relativistic correction to the frequency of a harmonic oscillator as a function of its amplitude.

9.10. An atom with an initial rest mass m has been excited to an internal state with an additional energy $\Delta\mathscr{E}$, while still being at rest. Next, it returns to its initial state, emitting a photon. Calculate the photon's frequency, taking into account the relativistic recoil of the atom.

Hint: In this problem, and also in Problems 13-15 below, you may treat photons as classical ultra-relativistic point particles with zero rest mass, energy $\mathscr{E} = \hbar\omega$, and momentum $\mathbf{p} = \hbar\mathbf{k}$.

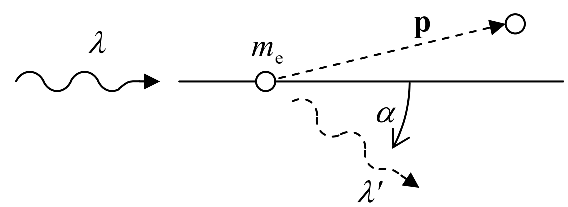
9.11. A particle of mass m, initially at rest, decays into two particles with rest masses m_1 and m_2. Calculate the total energy of the first product particle, in the c.o.m. reference frame.

9.12. A relativistic particle with a rest mass m, moving with velocity u, decays into two particles with zero rest mass.

(i) Calculate the smallest possible angle between the decay product velocities (in the lab frame, in that the velocity u is measured).
(ii) What is the largest possible energy of one product particle?

9.13. A relativistic particle flying in free space with velocity $\mathbf{u}$ decays into two photons.[83] Calculate the angular dependence of the photon detection probability, as measured in the lab frame.

[83] Such a decay may happen, for example, with a neutral pion.



9.14. A photon with wavelength λ is scattered by an electron, initially at rest. Calculate the wavelength λ' of the scattered photon as a function of the scattering angle α – see the figure on the right.[84]

9.15. Calculate the threshold energy of a γ-photon for the reaction

$$\gamma + \mathrm{p} \rightarrow \mathrm{p} + \pi^0,$$

if the proton was initially at rest.

Hint: For protons, $m_p c^2 \approx 938$ MeV, while for neutral pions, $m_\pi c^2 \approx 135$ MeV.

9.16. Calculate the largest possible velocity of the electrons emitted by (initially, resting) neutrons at their *β-decays*:

$$\mathrm{n} \rightarrow \mathrm{p} + \mathrm{e} + \bar{\nu}_e .$$

Hint: Electron neutrinos ν_e and antineutrinos $\bar{\nu}_e$ are virtually massless (on the energy scale of this problem); the rest energies $\mathscr{E} \equiv mc^2$ of the other involved particles are as follows: 939.565 MeV for the neutron, 938.272 MeV for the proton, 0.511 MeV for the electron.

9.17. A relativistic particle with a rest mass m and an energy $\mathscr{E}$ collides with a similar particle, initially at rest in the laboratory reference frame. Calculate:

(i) the final velocity of the center of mass of the system, in the lab frame,

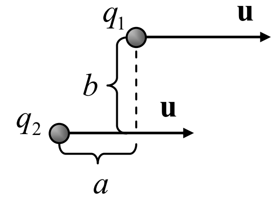
(ii) the total energy of the system, in the center-of-mass frame, and

(iii) the final velocities of both particles (in the lab frame), provided that they move along the same direction.

9.18. A "primed" reference frame moves, relative to the "lab" frame, with a reduced velocity $\boldsymbol{\beta} \equiv \mathbf{v}/c = \mathbf{n}_x\beta$. Use Eq. (109) to express the elements T'^{00} and T'^{0j} (with j = 1, 2, 3) of an arbitrary contravariant 4-tensor $T^{\gamma\delta}$ via its elements in the lab frame.

9.19. Prove that quantities $E^2 - c^2B^2$ and $\mathbf{E}\cdot\mathbf{B}$ are Lorentz-invariant.

9.20. Consider the situation when static fields **E** and **B** are uniform but arbitrary (both in magnitude and in direction). What should be the velocity of an inertial reference frame to have the vectors **E**' and **B**', observed from that frame, parallel? Is this solution unique?



9.21. Two charged particles moving with equal constant velocities **u** are offset by a constant vector $\mathbf{R} = \{a, b\}$ (see the figure on the right), as measured in the lab frame. Calculate the force of interaction between the particles – also in the lab frame.

[84] This is the famous *Compton scattering* effect, whose discovery in 1923 was one of the major motivations for the development of quantum mechanics – see, e.g., QM Sec. 1.1.

9.22. Each of two thin, long, parallel particle beams of the same velocity **u**, separated by distance d, carries electric charges with a constant density λ per unit length, as measured in the reference frame moving with the particles.

(i) Calculate the distribution of the electric and magnetic fields in the system (outside the beams), as measured in the lab reference frame.

(ii) Calculate the interaction force between the beams (per particle) and the resulting acceleration, both in the lab reference frame and in the frame moving with the particles.

(iii) Compare the results and give a brief discussion of their relation.

9.23.

(i) Spell out the Lorentz transform of the Cartesian components of the scalar potential and the vector potential of an arbitrary electromagnetic field.

(ii) Use this general result to calculate the potentials of the field created by a point charge q moving with a constant velocity **u**, as measured in the lab reference frame.

9.24. Calculate the scalar and vector potentials created by a time-independent electric dipole **p**, as measured in a reference frame that moves relative to the dipole with a constant velocity **v**, with the shortest distance ("impact parameter") equal to b.

9.25. Solve the previous problem, in the limit $v \ll c$, for a time-independent magnetic dipole **m**.

9.26. Review the solution of Problem 23 (on the hypothetical magnetic monopole passing through a superconducting ring) for the case when this particle moves with an arbitrary constant velocity.

9.27. Re-derive Eq. (161) for the simplest case **p**(0) = 0, by using the 4-vector form (145) of the equation of motion and the notion of rapidity $\varphi \equiv \tanh^{-1}\beta$ that was briefly discussed in Sec. 2.

9.28.* Calculate the trajectory of a relativistic particle in a uniform electrostatic field **E**, for an arbitrary direction of its initial velocity **u**(0), by using two different ways – at least one of them different from the approach described in Sec. 6 for the case **u**(0) $\perp$ **E**.

9.29. A charged relativistic particle the rest mass m performs planar cyclotron rotation, with velocity u, in a uniform external magnetic field of magnitude B. How much would the velocity and the orbit's radius change at a slow change of the field to a new magnitude B'?

9.30.* Analyze the motion of a relativistic particle in uniform, mutually perpendicular fields **E** and **B**, for the particular case when E is *exactly* equal to cB.

9.31. Find the law of motion of a relativistic particle in uniform static fields **E** and **B** parallel to each other.

9.32. An external Lorentz force **F** is exerted on a relativistic particle with an electric charge q and a rest mass m, moving with velocity **u**, as observed from some inertial "lab" frame. Calculate its acceleration as observed from that frame.

9.33. Neglecting relativistic kinetic effects, calculate the lowest voltage V that has to be applied between the anode and cathode of a magnetron (see Fig. 13 and its discussion) to enable electrons to reach the anode, at negligible electron-electron interactions (including the space-charge effects) and collisions with the residual gas molecules. You may:

(i) model the cathode and anode as two coaxial round cylinders, of radii R_1 and R_2, respectively;

(ii) assume that the magnetic field **B** is uniform and directed along their common axis; and

(iii) neglect the initial velocity of the electrons emitted by the cathode.

After the solution, estimate the validity of the last assumption and of the non-relativistic approximation, for reasonable values of parameters.

9.34. A charged relativistic particle has been injected into a region with a uniform electric field whose magnitude oscillates in time with frequency ω. Calculate the time dependence of the particle's velocity, as observed from the lab reference frame.

9.35.* A linearly-polarized plane electromagnetic wave of frequency ω is incident on an otherwise free relativistic particle with electric charge q. Analyze the dynamics of the particle's momentum and compare the result with those of the previous problem and Problem 7.5.

9.36. Analyze the motion of a non-relativistic particle in a region where the electric and magnetic fields are both uniform and constant in time, but not necessarily parallel or perpendicular to each other.

9.37. A static distribution of electric charge in otherwise free space has created a time-independent distribution $\mathbf{E}(\mathbf{r})$ of the electric field. Use two different approaches to express the field energy density u' and the Poynting vector $\mathbf{S}'$, as observed from a reference frame moving with a constant velocity $\mathbf{v}$, via the Cartesian components of the vector $\mathbf{E}$. In particular, is $\mathbf{S}'$ equal to $(-\mathbf{v}u')$?

9.38. A traveling plane wave of frequency ω and intensity S is normally incident on a perfect mirror moving with velocity v in the same direction as the wave.

(i) Calculate the reflected wave's frequency, and

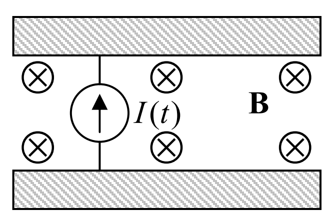
(ii) use the Lorentz transform of the fields to calculate the reflected wave's intensity

– both as observed from the lab reference frame.

9.39. Perform the second task of the previous problem by using general relations between the wave's energy, power, and momentum.

Hint: As a byproduct, this approach should also give you the pressure exerted by the wave on the moving mirror.

9.40. For the simple model of capacitor charging by a lumped source of current $I(t)$, shown in the figure on the right, prove that the momentum given by a constant, uniform external magnetic field **B** to the current-carrying conductor is equal and opposite to the momentum of the electromagnetic field that this current builds up in the capacitor. (You may assume that the capacitor is planar and very broad, and hence neglect the fringe field effects.)



9.41. Consider an electromagnetic plane wave packet propagating in free space, with its electric field represented as the Fourier integral

$$\mathbf{E}(\mathbf{r},t) = \operatorname{Re}\int_{-\infty}^{+\infty} \mathbf{E}_k e^{i\Psi_k} dk, \qquad \text{with } \Psi_k \equiv kz - \omega_k t, \quad \text{and } \omega_k \equiv c|k|.$$

Express the full linear momentum (per unit area of wave's front) of the packet via the complex amplitudes $\mathbf{E}_k$ of its Fourier components. Does the momentum depend on time? (In contrast with Problem 7.8, the wave packet is not necessarily narrow.)

9.42. Calculate the forces exerted on well-conducting walls of a waveguide with a rectangular ($a\times b$) cross-section, by a wave propagating along it in the fundamental (H_{10}) mode. Give an interpretation of the results.