2 Relativity and electromagnetism

2.1 The relativity principle

Physical phenomena are conventionally described relative to some *frame of reference* which allows us to define fundamental quantities such as position and time. Of course, there are very many different ways of choosing a reference frame, but it generally convenient to restrict our choice to the set of rigid inertial frames. A classical rigid reference frame is the imagined extension of a rigid body. For instance, the Earth determines a rigid frame throughout all space, consisting of all those points which remain rigidly at rest relative to the Earth and each other. We can associate an orthogonal Cartesian coordinate system S with such a frame, by choosing three mutually orthogonal planes within it and measuring x, y, and z as distances from these planes. A time coordinate must also be defined in order that the system can be used to specify events. A rigid frame, endowed with such properties, is called a *Cartesian frame*. The description given above presupposes that the underlying geometry of space is Euclidian, which is reasonable provided that gravitational effects are negligible (we shall assume that this is the case). An *inertial* frame is a Cartesian frame in which free particles move without acceleration, in accordance with Newton's first law of motion. There are an infinite number of different inertial frames, each moving with some constant velocity with respect to a given inertial frame.

The key to understanding special relativity is Einstein's *relativity principle*, which states that

> All inertial frames are totally equivalent for the performance of all physical experiments.

In other words, it is impossible to perform a physical experiment which differentiates in any fundamental sense between different inertial frames. By definition, Newton's laws of motion take the same form in all inertial frames. Einstein generalized this result in his special theory of relativity by asserting that *all* laws of physics take the same form in all inertial frames.

Consider a wave-like disturbance. In general, such a disturbance propagates at a fixed velocity with respect to the medium in which the disturbance takes place. For instance, sound waves (at S.T.P.) propagate at 343 meters per second with respect to air. So, in the inertial frame in which air is stationary sound waves appear to propagate at 343 meters per second. Sound waves appear to propagate at a different velocity in some other inertial frame which is moving with respect to the first frame. However, this does not violate the relativity principle, since if the air were stationary in the second frame then sound waves would appear to propagate at 343 meters per second in this frame as well. In other words, exactly the same experiment (*e.g.*, the determination of the speed of sound relative to stationary air) performed in two different inertial frames of reference yields exactly the same result, in accordance with the relativity principle.

Consider, now, a wave-like disturbance which is self-regenerating and does not require a medium through which to propagate. The most well known example of such a disturbance is a light wave. Another example is a gravity wave. According to electromagnetic theory the speed of propagation of a light wave through a vacuum is

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}} = 2.99729 \times 10^8 \text{ meters per second,} \tag{2.1}$$

where ϵ_0 and μ_0 are physical constants which can be evaluated by performing two simple experiments which involve measuring the force of attraction between two fixed changes and two fixed parallel current carrying wires. According to the relativity principle these experiments must yield the same values for ϵ_0 and μ_0 in all inertial frames. Thus, the speed of light must be the same in all inertial frames. In fact, any disturbance which does not require a medium to propagate through must appear to travel at the same velocity in all inertial frames, otherwise we could differentiate inertial frames using the apparent propagation speed of the disturbance, which would violate the relativity principle.

2.2 The Lorentz transform

Consider two Cartesian frames $S(x, y, z, t)$ and $S'(x', y', z', t')$ in the *standard configuration* in which S' moves in the x-direction of S with uniform velocity v and the corresponding axes of S and S' remain parallel throughout the motion,

having coincided at $t = t' = 0$. It is assumed that the same units of distance and time are adopted in both frames. Suppose that an *event* (*e.g.,* the flashing of a light-bulb, or the collision of two point particles) has coordinates $(x,\, y,\, z,\, t)$ relative to S and $(x',\, y',\, z',\, t')$ relative to S'. The "common sense" relationship between these two sets of coordinates is given by the Galilean transformation:

$$x' = x - vt, \tag{2.2a}$$

$$y' = y, \tag{2.2b}$$

$$z' = z, \tag{2.2c}$$

$$t' = t. \tag{2.2d}$$

This transformation is tried and tested and provides a very accurate description of our everyday experience. Nevertheless, it must be wrong! Consider a light wave which propagates along the x-axis in S with velocity c. According to the Galilean transformation the apparent speed of propagation in S' is $c - v$, which violates the relativity principle. Can we construct a new transformation which makes the velocity of light invariant between different inertial frames, in accordance with the relativity principle, but reduces to the Galilean transformation at low velocities, in accordance with our everyday experience?

Consider an event P and a neighbouring event Q whose coordinates differ from those of P by dx, dy, dz, dt in S and by dx', dy', dz', dt' in S'. Suppose that at the event P a flash of light is emitted and that Q is an event in which some particle in space is illuminated by the flash. In accordance with the laws of light-propagation, and the invariance of the velocity of light between different inertial frames, an observer in S will find that

$$dx^2 + dy^2 + dz^2 - c^2 dt^2 = 0 \tag{2.3}$$

for $dt > 0$, and an observer in S' will find that

$$dx'^2 + dy'^2 + dz'^2 - c^2 dt'^2 = 0 \tag{2.4}$$

for $dt' > 0$. Any event near P whose coordinates satisfy *either* (2.3) *or* (2.4) is illuminated by the flash from P and therefore its coordinates must satisfy *both*

(2.3) *and* (2.4). Now, no matter what form the transformation between coordinates in the two inertial frames takes, the transformation between differentials at any fixed event P is linear and homogeneous. In other words, if

$$x' = F(x, y, z, t), \tag{2.5}$$

where F is a general function, then

$$dx' = \frac{\partial F}{\partial x}\,dx + \frac{\partial F}{\partial y}\,dy + \frac{\partial F}{\partial z}\,dz + \frac{\partial F}{\partial t}\,dt. \tag{2.6}$$

It follows that

$$\begin{aligned} dx'^2 + dy'^2 + dz'^2 - c^2 dt'^2 &= a\,dx^2 + b\,dy^2 + c\,dz^2 + d\,dt^2 + g\,dx\,dt + h\,dy\,dt \\ &\quad + k\,dz\,dt + l\,dy\,dz + m\,dx\,dz + n\,dx\,dy, \end{aligned} \tag{2.7}$$

where a, b, c, *etc.* are functions of x, y, z, and t. We know that the right-hand side of the above expression vanishes for all real values of the differentials which satisfy Eq. (2.3). It follows that the right-hand side is a multiple of the quadratic in Eq. (2.3); *i.e.*,

$$dx^2 + dy^2 + dz^2 - c^2 dt^2 = K(dx'^2 + dy'^2 + dz'^2 - c^2 dt'^2), \tag{2.8}$$

where K is a function of x, y, z, and t. [We can prove this by substituting into Eq. (2.7) the following obvious zeros of the quadratic in Eq. (2.3): $(\pm 1, 0, 0, 1)$, $(0, \pm 1, 0, 1)$, $(0, 0, \pm 1, 1)$, $(0, 1/\sqrt{2}, 1/\sqrt{2}, 1)$, $(1/\sqrt{2}, 0, 1/\sqrt{2}, 1)$, $(1/\sqrt{2}, 1/\sqrt{2}, 0, 1)$: and solving the resulting conditions on the coefficients.] Note that K at P is also independent of the choice of standard coordinates in S and S'. Since the frames are Euclidian, the values of $dx^2 + dy^2 + dz^2$ and $dx'^2 + dy'^2 + dz'^2$ relevant to P and Q are independent of the choice of axes. Furthermore, the values of dt^2 and dt'^2 are independent of the choice of the origins of time. Thus, without affecting the value of K at P we can choose coordinates such that $P = (0, 0, 0, 0)$ in both S and S'. Since the orientations of the axes in S and S' are, at present, arbitrary, and since inertial frames are isotropic, the relation of S and S' relative to each other, to the event P, and to the locus of possible events Q is now completely symmetric. Thus, we can write

$$dx'^2 + dy'^2 + dz'^2 - c^2 dt'^2 = K(dx^2 + dy^2 + dz^2 - c^2 dt^2), \tag{2.9}$$

in addition to Eq. (2.8). It follows that $K = \pm 1$. $K = -1$ can be dismissed immediately, since the intervals $dx^2+dy^2+dz^2-c^2dt^2$ and $dx'^2+dy'^2+dz'^2-c^2dt'^2$ must coincide exactly when there is no motion of S' relative to S. Thus,

$$dx'^2 + dy'^2 + dz'^2 - c^2dt'^2 = dx^2 + dy^2 + dz^2 - c^2dt^2. \tag{2.10}$$

Equation (2.10) implies that the transformation equations between primed and unprimed coordinates must be *linear*. The proof of this statement is postponed until later.

The linearity of the transformation allows the coordinate axes in the two frames to be orientated so as to give the *standard configuration* mentioned earlier. Consider a fixed plane in S with the equation $lx + my + nz + p = 0$. In S' this becomes, say, $l(a_1x'+b_1y'+c_1z'+d_1t'+e_1)+m(a_2x'+\cdots)+n(a_3x'+\cdots)+p = 0$, which represents a moving plane unless $ld_1 + md_2 + nd_3 = 0$. That is, unless the normal vector to the plane (l, m, n) in S is perpendicular to the vector (d_1, d_2, d_3). All such planes intersect in lines which are fixed in both S and S', and which are parallel to the vector (d_1, d_2, d_3) in S. These lines must correspond to the direction of relative motion of the frames. By symmetry, two such frames which are orthogonal in S must also be orthogonal in S'. This allows the choice of two common coordinate planes.

Under a linear transformation the finite coordinate differences satisfy the same transformation equations as the differentials. It follows from Eq. (2.10), assuming that the events $(0, 0, 0, 0)$ coincide in both frames, that for any event with coordinates (x, y, z, t) in S and (x', y', z', t') in S' the following relation holds:

$$x^2 + y^2 + z^2 - c^2t^2 = x'^2 + y'^2 + z'^2 - c^2t'^2. \tag{2.11}$$

By hypothesis, the coordinate planes $y = 0$ and $y' = 0$ coincide permanently. Thus, $y = 0$ must imply $y' = 0$, which suggests that

$$y' = Ay, \tag{2.12}$$

where A is a constant. We can reverse the directions of the x- and z-axes in S and S', which has the effect of interchanging the roles of these frames. This procedure does not affect Eq. (2.12), but by symmetry we also have

$$y = Ay'. \tag{2.13}$$

It is clear that $A = \pm 1$. The negative sign can again be dismissed, since $y = y'$ when there is no motion between S and S'. The argument for z is similar. Thus, we have

$$y' = y, \tag{2.14a}$$

$$z' = z, \tag{2.14b}$$

as in the Galilean transformation.

Equations (2.11) and (2.14) yield

$$x^2 - c^2t^2 = x'^2 - c^2t'^2. \tag{2.15}$$

Since, $x' = 0$ must imply $x = vt$, we can write

$$x' = B(x - vt), \tag{2.16}$$

where B is a constant (possibly depending on v). It follows from the previous two equations that

$$t' = Cx + Dt, \tag{2.17}$$

where C and D are constants (possibly depending on v). Substituting Eqs. (2.16) and (2.17) into Eq. (2.15) and comparing the coefficients of x^2, xt, and t^2, we obtain

$$B = D = \frac{1}{\pm(1 - v^2/c^2)^{1/2}}, \tag{2.18a}$$

$$C = \frac{-v/c^2}{\pm(1 - v^2/c^2)^{1/2}}. \tag{2.18b}$$

We must choose the positive sign in order to ensure that $x' \to x$ as $v/c \to 0$. Thus, collecting our results, the transformation between coordinates in S and S' is given by

$$x' = \frac{x - vt}{(1 - v^2/c^2)^{1/2}}, \tag{2.19a}$$

$$y' = y, \tag{2.19b}$$

$$z' = z, \tag{2.19c}$$

$$t' = \frac{t - vx/c^2}{(1 - v^2/c^2)^{1/2}}, \tag{2.19d}$$

This is the famous *Lorentz transform.* It ensures that the velocity of light is invariant between different inertial frames, and also reduces to the more familiar Galilean transform in the limit $v/c \ll 1$. We can solve Eqs. (2.19) for x, y, z, and t to obtain the *inverse Lorentz transform*:

$$x = \frac{x' + vt'}{(1 - v^2/c^2)^{1/2}}, \tag{2.20a}$$

$$y = y', \tag{2.20b}$$

$$z = z', \tag{2.20c}$$

$$t = \frac{t' + vx'/c^2}{(1 - v^2/c^2)^{1/2}}. \tag{2.20d}$$

Clearly, the inverse transform is equivalent to a Lorentz transform in which the velocity of the moving frame is $-v$ along the x-axis instead of $+v$.

2.3 Transformation of velocities

Consider two frames S and S' in the standard configuration. Let $\boldsymbol{u}$ be the velocity of a particle in S. What is the particle velocity in S'? The components of the velocity are

$$u_1 = \frac{dx}{dt}, \tag{2.21a}$$

$$u_2 = \frac{dy}{dt}, \tag{2.21b}$$

$$u_3 = \frac{dz}{dt}, \tag{2.21c}$$

and, similarly, the components of $\boldsymbol{u}'$ are

$$u_1' = \frac{dx'}{dt'}, \tag{2.22a}$$

$$u_2' = \frac{dy'}{dt'}, \tag{2.22b}$$

$$u_3' = \frac{dz'}{dt'}. \tag{2.22c}$$

Now we can write Eqs. (2.19) in the form $dx' = \gamma(dx - vdt)$, $dy' = dy$, $dz' = dz$, and $dt' = \gamma(dt - vdx/c^2)$, where

$$\gamma = \frac{1}{(1 - v^2/c^2)^{1/2}} \tag{2.23}$$

is the well known *Lorentz factor*. If we substitute these differentials into Eqs. (2.22) and make use of Eqs. (2.21), we obtain the transformation formulae

$$u_1' = \frac{u_1 - v}{1 - u_1 v/c^2}, \tag{2.24a}$$

$$u_2' = \frac{u_2}{\gamma(1 - u_1 v/c^2)}, \tag{2.24b}$$

$$u_3' = \frac{u_3}{\gamma(1 - u_1 v/c^2)}. \tag{2.24c}$$

As in the transformation of coordinates, we can obtain the inverse transform by interchanging primed and unprimed symbols and replacing $+v$ with $-v$. Thus,

$$u_1 = \frac{u_1' + v}{1 + u_1' v/c^2}, \tag{2.25a}$$

$$u_2 = \frac{u_2'}{\gamma(1 + u_1' v/c^2)}, \tag{2.25b}$$

$$u_3 = \frac{u_3'}{\gamma(1 + u_1' v/c^2)}. \tag{2.25c}$$

Equations (2.25) can be regarded as giving the resultant, $\boldsymbol{u} = (u_1, u_2, u_3)$, of two velocities, $\boldsymbol{v} = (v, 0, 0)$ and $\boldsymbol{u}' = (u_1', u_2', u_3')$, and are therefore usually referred to as the relativistic *velocity addition formulae*. The following relation between the magnitudes $u = ({u_1}^2 + {u_2}^2 + {u_3}^2)^{1/2}$ and $u' = ({u_1'}^2 + {u_2'}^2 + {u_3'}^2)^{1/2}$ of the velocities is easily demonstrated:

$$c^2 - u^2 = \frac{c^2(c^2 - u'^2)(c^2 - v^2)}{(c^2 + u_1' v)^2}. \tag{2.26}$$

If $u' < c$ and $v < c$ the right-hand side is positive, implying that $u < c$. In other words, the resultant of two subluminal velocities is another subluminal velocity. It is evident that a particle can never attain the velocity of light relative to a given inertial frame, no matter how many subluminal velocity increments it is given. It follows that no inertial frame can appear to propagate with a superluminal velocity with respect to any other inertial frame (since we can track the origin of a given inertial frame using a particle which remains at rest at the origin in that frame).

According to Eq. (2.26), if $u' = c$ then $u = c$ no matter what value v takes; *i.e.,* the velocity of light is invariant between different inertial frames. Note that the Lorentz transform only allows *one* such invariant velocity (*i.e.,* the velocity c which appears in Eqs. (2.19)). Einstein's relativity principle tells us that any disturbance which propagates through a vacuum must appear to propagate at the same velocity in all inertial frames. It is now evident that *all* such disturbances must propagate at the velocity c. It follows immediately that all electromagnetic waves must propagate through the vacuum with this velocity, irrespective of their wavelength. In other words, it is impossible for there to be any dispersion of electromagnetic waves propagating through a vacuum. Furthermore, gravity waves must also propagate with the velocity c. It is convenient to label c as "the velocity of light" since electromagnetic radiation is, by far, the most well known and easily measurable type of disturbance which can propagate through a vacuum.

The Lorentz transformation implies that not only the velocities of material particles but the velocities of propagation of all physical effects are limited by c in deterministic physics. Consider a general process by which an event P *causes* an event Q at a velocity $U > c$ in some frame S. In other words, *information* about the event P appears to propagate to the event Q with a superluminal velocity. Let us choose coordinates such that these two events occur on the x-axis with (finite) time and distance separations $\Delta t > 0$ and $\Delta x > 0$, respectively. The time separation in some other inertial frame S' is given by (see Eq. (2.19d))

$$\Delta t' = \gamma(\Delta t - v\Delta x/c^2) = \gamma\Delta t(1 - vU/c^2). \tag{2.27}$$

Thus, for sufficiently large $v < c$ we obtain $\Delta t' < 0$; *i.e.,* there exist inertial frames in which cause and effect appear to be reversed. Of course, this is impossible in

deterministic physics. It follows, therefore, that information can never appear to propagate with a superluminal velocity in any inertial frame, otherwise causality would be violated.

2.4 Tensors

It is now convenient to briefly review the mathematics of tensors. Tensors are of primary importance in connection with coordinate transforms. They serve to isolate intrinsic geometric and physical properties from those that merely depend on coordinates.

A tensor of rank r in an n-dimensional space possesses n^r components which are, in general, functions of position in that space. A tensor of rank zero has one component A and is called a *scalar*. A tensor of rank one has n components $(A_1, A_2, \cdots, A_n)$ and is called a *vector*. A tensor of rank two has n^2 components, which can be exhibited in matrix format. Unfortunately, there is no convenient way of exhibiting a higher rank tensor. Consequently, tensors are usually represented by a typical component; *e.g.,* we talk of the tensor A_{ijk} (rank 3) or the tensor A_{ijkl} (rank 4), *etc.* The suffixes $i, j, k, \cdots$ are always understood to range from 1 to n.

For reasons which will become apparent later on, we shall represent tensor components using both superscripts and subscripts. Thus, a typical tensor might look like A^{ij} (rank 2), or B^i_j (rank 2), *etc.* It is convenient to adopt the Einstein summation convention. Namely, if any suffix appears twice in a given term, once as a subscript and once as a superscript, a summation over that suffix (from 1 to n) is implied.

To distinguish between various coordinate systems we shall use primed and multiply primed suffixes. A first system of coordinates $(x^1, x^2, \cdots, x^n)$ can then be denoted by x^i, a second system $(x^{1'}, x^{2'}, \cdots, x^{n'})$ by $x^{i'}$, *etc.* Similarly the general components of a tensor in various coordinate systems are distinguished by their suffixes. Thus, the components of some third rank tensor are denoted A_{ijk} in the x^i system, by $A_{i'j'k'}$ in the $x^{i'}$ system, *etc.*

When making a coordinate transformation from one set of coordinates x^i to

another $x^{i'}$, it is assumed that the transformation in non-singular. In other words, the equations which express the $x^{i'}$ in terms of the x^i can be inverted to express the x^i in terms of the $x^{i'}$. It is also assumed that the functions specifying a transformation are differentiable. It is convenient to write

$$\frac{\partial x^{i'}}{\partial x^i} = p_i^{i'}, \tag{2.28a}$$

$$\frac{\partial x^i}{\partial x^{i'}} = p_{i'}^i. \tag{2.28b}$$

Note that

$$p_{i'}^i p_{i''}^{i'} = p_{i''}^i, \tag{2.29a}$$

$$p_{i'}^i p_j^{i'} = \delta_j^i \tag{2.29b}$$

by the chain rule, where δ_j^i (the *Kronecker delta*) equals 1 or 0 when $i = j$ or $i \neq j$, respectively.

The formal definition of a tensor is as follows:
(i) An entity having components $A_{ij\cdots k}$ in the x^i system and $A_{i'j'\cdots k'}$ in the $x^{i'}$ system is said to behave as a *covariant tensor* under the transformation $x^i \to x^{i'}$ if

$$A_{i'j'\cdots k'} = A_{ij\cdots k}\, p_{i'}^i p_{j'}^j \cdots p_{k'}^k. \tag{2.30}$$

(ii) Similarly, $A^{ij\cdots k}$ is said to behave as a *contravariant tensor* under $x^i \to x^{i'}$ if

$$A^{i'j'\cdots k'} = A^{ij\cdots k} p_i^{i'}\, p_j^{j'} \cdots p_k^{k'}. \tag{2.31}$$

(iii) Finally, $A_{k\cdots l}^{i\cdots j}$ is said to behave as a *mixed tensor* (contravariant in $i \cdots j$ and covariant in $k \cdots l$) under $x^i \to x^{i'}$ if

$$A_{k'\cdots l'}^{i'\cdots j'} = A_{k\cdots l}^{i\cdots j}\, p_i^{i'} \cdots p_j^{j'}\, p_{k'}^k \cdots p_{l'}^l. \tag{2.32}$$

When an entity is described as a tensor it is generally understood that it behaves as a tensor under *all* non-singular differentiable transformations of the

relevant coordinates. An entity which only behaves as a tensor under a certain subgroup of non-singular differentiable coordinate transformations is called a *qualified tensor*, because its name is conventionally qualified by an adjective recalling the subgroup in question. For instance, an entity which only exhibits tensor behaviour under Lorentz transformations is called a Lorentz tensor or, more commonly, a 4-tensor.

When applied to a tensor of rank zero (a scalar), the above definitions imply that $A^{i} = A$. Thus, a scalar is a function of position only, and is independent of the coordinate system. A scalar is often termed an *invariant*.

The main theorem of tensor calculus is as follows:

> If two tensors of the same type are equal in one coordinate system, then they are equal in all coordinate systems.

The simplest example of a contravariant vector (tensor of rank one) is provided by the differentials of the coordinates, dx^{i}, since

$$dx^{i'} = \frac{\partial x^{i'}}{\partial x^{i}} dx^{i} = dx^{i} p_{i}^{i'}. \tag{2.33}$$

The coordinates themselves do not behave as tensors under all coordinate transformations. However, since they transform like their differentials under linear homogeneous coordinate transformations, they do behave as tensors under such transformations.

The simplest example of a covariant vector is provided by the gradient of a function of position $\phi = \phi(x^{1}, \cdots, x^{n})$. Since, if we write

$$\phi_{i} = \frac{\partial \phi}{\partial x^{i}}, \tag{2.34}$$

then we have

$$\phi_{i'} = \frac{\partial \phi}{\partial x^{i'}} = \frac{\partial \phi}{\partial x^{i}} \frac{\partial x^{i}}{\partial x^{i'}} = \phi_{i}\, p_{i'}^{i}. \tag{2.35}$$

An important example of a mixed second rank tensor is provided by the

Kronecker delta introduced previously. Since,

$$\delta^i_j \, p^{i'}_i \, p^j_{j'} = p^{i'}_j \, p^j_{j'} = \delta^{i'}_{j'}. \tag{2.36}$$

Tensors *of the same type* can be added or subtracted to form new tensors. Thus, if A_{ij} and B_{ij} are tensors, then $C_{ij} = A_{ij} \pm B_{ij}$ is a tensor of the same type. Note that the sum of tensors at different points in space is not a tensor if the p's are position dependent. However, under linear coordinate transformations the p's are constant, so the sum of tensors at different points behaves as a tensor under this particular type of coordinate transformation.

If A^{ij} and B_{ijk} are tensors, then $C^{ij}_{klm} = A^{ij} B_{klm}$ is a tensor of the type indicated by the suffixes. The process illustrated by this example is called *outer multiplication* of tensors.

Tensors can also be combined by *inner multiplication*, which implies at least one dummy suffix link. Thus, $C^j_{kl} = A^{ij} B_{ikl}$ and $C_k = A^{ij} B_{ijk}$ are tensors of the type indicated by the suffixes.

Finally, tensors can be formed by *contraction* from tensors of higher rank. Thus, if A^{ij}_{klm} is a tensor then $C^j_{kl} = A^{ij}_{ikl}$ and $C_k = A^{ij}_{kij}$ are tensors of the type indicated by the suffixes. The most important type of contraction occurs when no free suffixes remain: the result is a scalar. Thus, A^i_i is a scalar provided that A^j_i is a tensor.

Although we cannot usefully divide tensors, one by another, an entity like C^{ij} in the equation $A^j = C^{ij} B_i$, where A^i and B_i are tensors, can be formally regarded as the quotient of A^i and B_i. This gives the name to a particularly useful rule for recognizing tensors, the *quotient rule*. This rule states that *if a set of components, when combined by a given type of multiplication with all tensors of a given type yields a tensor, then the set is itself a tensor.* In other words, if the product $A^i = C^{ij} B_j$ transforms like a tensor for *all* tensors B_i then it follows that C^{ij} is a tensor.

Let

$$\frac{\partial A^{i\cdots j}_{k\cdots l}}{\partial x^m} = A^{i\cdots j}_{k\cdots l,m}. \tag{2.37}$$

Then if $A^{i\cdots j}_{k\cdots l}$ is a tensor, differentiation of the general tensor transformation (2.32) yields

$$A^{i'\cdots j'}_{k'\cdots l',m'} = A^{i\cdots j}_{k\cdots l,m}\, p^{i'}_i \cdots p^{j'}_j\, p^k_{k'} \cdots p^l_{l'}\, p^m_{m'} + P_1 + P_2 + \cdots, \tag{2.38}$$

where P_1, P_2, *etc.*, are terms involving derivatives of the p's. Clearly, $A^{i\cdots j}_{k\cdots l}$ is not a tensor under a general coordinate transformation. However, under a linear coordinate transformation (p's constant) $A^{i'\cdots j'}_{k'\cdots l',m'}$ behaves as a tensor of the type indicated by the suffixes, since the P_1, P_2, *etc.*, all vanish. Similarly, all higher partial derivatives,

$$A^{i\cdots j}_{k\cdots l,mn} = \frac{\partial A^{i\cdots j}_{k\cdots l}}{\partial x^m \partial x^n} \tag{2.39}$$

etc., also behave as tensors under linear transformations. Each partial differentiation has the effect of adding a new covariant suffix.

So far the space to which the coordinates x^i refer has been without structure. We can impose a structure on it by defining the distance between all pairs of neighbouring points by means of a *metric*

$$ds^2 = g_{ij}dx^i dx^j, \tag{2.40}$$

where the g_{ij} are functions of position. We can assume that $g_{ij} = g_{ji}$ without loss of generality. The above metric is analogous to, but more general than, the metric of Euclidian n-space, $ds^2 = (dx^1)^2 + (dx^2)^2 + \cdots + (dx^n)^2$. A space whose structure is determined by a metric of the type (2.40) is called *Riemannian*. Since ds^2 is invariant, it follows from a simple extension of the quotient rule that g_{ij} must be a tensor. It is called the *metric tensor*.

The elements of the inverse of the matrix g_{ij} are denoted by g^{ij}. These elements are uniquely defined by the equations

$$g^{ij}g_{jk} = \delta^i_k. \tag{2.41}$$

It is easily seen that the g^{ij} constitute the elements of a contravariant tensor. This tensor is said to be *conjugate* to g_{ij}. The conjugate metric tensor is symmetric (*i.e.*, $g^{ij} = g^{ji}$) just like the metric tensor itself.

The tensors g_{ij} and g^{ij} allow us to introduce the important operations of *raising* and *lowering suffixes*. These operations consist of forming inner products of a given tensor with g_{ij} or g^{ij}. For example, given a contravariant vector A^i, we define its covariant components A_i by the equation

$$A_i = g_{ij} A^j. \tag{2.42}$$

Conversely, given a covariant vector B_i, we can define its contravariant components B^i by the equations

$$B^i = g^{ij} B_j. \tag{2.43}$$

More generally, we can raise or lower any or all of the free suffixes of any given tensor. Thus, if A_{ij} is a tensor we define $A^i{}_j$ by the equation

$$A^i{}_j = g^{ip} A_{pj}. \tag{2.44}$$

Note that once the operations of raising and lowering suffixes has been defined the order of raised suffixes relative to lowered suffixes becomes significant.

By analogy with Euclidian space we define the *squared magnitude* $(A)^2$ of a vector A^i with respect to the metric $g_{ij} dx^i dx^j$ by the equation

$$(A)^2 = g_{ij} A^i A^j = A_i A^i. \tag{2.45}$$

A vector A^i termed a *null vector* if $(A)^2 = 0$. Two vectors A^i and B^i are said to be *orthogonal* if their inner product vanishes, *i.e.*, if

$$g_{ij} A^i B^j = A_i B^i = A^i B_i = 0. \tag{2.46}$$

Finally, let us consider differentiation with respect to distance s. The *tangent vector* dx^i/ds to a given curve in space is a contravariant tensor, since

$$\frac{dx^{i'}}{ds} = \frac{\partial x^{i'}}{\partial x^i} \frac{dx^i}{ds} = \frac{dx^i}{ds}\, p_i^{i'}. \tag{2.47}$$

The derivative $d(A^{i\cdots j}{}_{k\cdots l})/ds$ of some tensor with respect to distance is not, in general, a tensor, since

$$\frac{d(A^{i\cdots j}{}_{k\cdots l})}{ds} = A^{i\cdots j}{}_{k\cdots l,m} \frac{dx^m}{ds}, \tag{2.48}$$

and, as we have seen, the first factor on the right is not generally a tensor. However, under linear transformations it behaves as a tensor, so under linear transformations the derivative of a tensor with respect to distance behaves as a tensor of the same type.

2.5 Transformations

In this course we shall only concern ourselves with coordinate transformations which transform an inertial frame into another inertial frame. This limits us to four classes of transformations: displacements of the coordinate axes, rotations of the coordinate axes, parity reversals (*i.e.*, $x, y, z \rightarrow -x, -y, -z$), and Lorentz transformations. All of these transformations possess *group properties*. As a reminder, the requirements for an abstract multiplicative group are:

(i) The product of two elements is an element of the group.
(ii) The associative law $(ab)c = a(bc)$ holds.
(iii) There is a unit element e satisfying $ae = ea = a$ for all a.
(iv) Each element a possesses an inverse a^{-1} such that $a^{-1}a = aa^{-1} = e$.

Consider Lorentz transformations (in the standard configuration). It is easily demonstrated that the resultant of two successive Lorentz transformations, with velocities v_1 and v_2, respectively, is equivalent to a Lorentz transformation with velocity $v = (v_1+v_2)/(1+v_1v_2/c^2)$. Lorentz transformations obviously satisfy the associative law. The unit element of the transformation group is just a Lorentz transformation with $v = 0$. Finally, the inverse of a Lorentz transformation with velocity v is a transformation with velocity $-v$. We can use similar arguments to show that translations, rotations, parity inversions, and general Lorentz transformations (*i.e.*, transformations between frames which are not in the standard configuration) also possess group properties.

If we think carefully, we can see that the group properties of the above mentioned transformations are a direct consequence of the relativity principle. Let us again consider Lorentz transformations. Suppose that we have three inertial frames S, S', and S''. According to (i), if we can get from S to S' by a

Lorentz transformation, and from S' to S'' by a second Lorentz transformation, then it must always be possible to go directly from S to S'' by means of a third Lorentz transformation. Suppose, for the sake of argument, that we can find three frames for which this is not the case. In this situation, the frame S' could be distinguished from the frame S'' because it is possible to make a direct Lorentz transformation from S to the former frame, but not to the latter. This violates the relativity principle and, therefore, this situation can never arise. We can use a similar argument to demonstrate that a Lorentz transformation must possess an inverse. The associative law and the requirement that a unit element exists are trivially satisfied.

2.6 The physical significance of tensors

One of the central tenets of physics is that experiments should be repeatable. In other words, if somebody performs a physical experiment today and obtains a certain result, then somebody else performing the same experiment next week ought to obtain the same result, within the experimental errors. Presumably, in performing these hypothetical experiments both experimentalists find it necessary to set up a coordinate frame. Usually, these two frames do not coincide. After all, the experiments are, in general, performed in different places and at different times. Also, the two experimentalists are likely to orientate their coordinate axes differently. For instance, one experimentalist might align his x-axis with the North Star, whilst the other might align the same axis to point towards Mecca. Nevertheless, we still expect both experiments to yield the same result. What exactly do we mean by this statement? We do not mean that both experimentalists will obtain the same numbers when they measure something. For instance, the numbers used to denote the position of a point (*i.e.*, the coordinates of the point) are, in general, different in different coordinate frames. What we do expect is that any physically significant interrelation between physical quantities (*i.e.*, position, velocity, *etc.*) which appears to hold in the coordinate system of the first experimentalist will also appear to hold in the coordinate system of the second experimentalist. We usually refer to such interrelationships as "laws of physics." So, what we are really saying is that the laws of physics do not depend on our choice of coordinate system. In particular, if a law of physics is true in one

coordinate system then it is automatically true in every other coordinate system, subject to the proviso that both coordinate systems are inertial.

Recall that tensors are geometric objects which possess the property that if a certain interrelationship holds between various tensors in one particular coordinate system, then the same interrelationship holds in any other coordinate system which is related to the first system by a certain class of transformations. It follows that *the laws of physics are expressible as interrelationships between tensors.* In special relativity the laws of physics are only required to exhibit tensor behaviour under transformations between different inertial frames; *i.e.*, translations, rotations, and Lorentz transformations. This set of transformations forms a group known as the *Poincaré group.* Parity inversion is a special type of transformation, and will be dealt with later on. In general relativity the laws of physics are required to exhibit tensor behaviour under *all* non-singular coordinate transformations.

Consider Newton's first law of motion. These take the form of three differential equations,

$$m\,\frac{d^2x}{dt^2} = f_x, \tag{2.49a}$$

$$m\,\frac{d^2y}{dt^2} = f_y, \tag{2.49b}$$

$$m\,\frac{d^2z}{dt^2} = f_z, \tag{2.49c}$$

in a general inertial frame. However, we can also write them as a single vector differential equation,

$$m\,\frac{d^2\boldsymbol{r}}{dt^2} = \boldsymbol{f}. \tag{2.50}$$

What is the advantage of the vector notation? Many people would say that it is just a convenient form of shorthand. However, there is another, far more important, advantage. Before we can accept Newton's first law of physics as a proper law of physics we need to convince ourselves that it is coordinate independent; *i.e.*, that it also holds in coordinate frames which are related to the original frame via a general translation or rotation of the coordinate axes. It is indeed possible

to prove this, but the demonstration is rather tedious because a general rotation is a rather complicated transformation. A vector is a geometric object (in fact, it is a rank one tensor in three dimensional Euclidean space) whose three components transform under a general translation and rotation of the coordinate axes in an analogous manner to the difference in coordinates between two fixed points in space. This ensures that any vector equation which is true in one coordinate frame is also true in any other coordinate frame which is related to the original frame via a general rotation or translation of the axes. Thus, the main advantage of Eq. (2.50) is that it makes the coordinate independent nature of Newton's first law of motion manifestly obvious. Of course, we cannot deny that Newton's first law also looks simpler when it is expressed in terms of vectors. This is one example of a rather general feature of physical laws. Namely, *when the laws of physics are expressed in a manner which makes their invariance under various transformation groups manifest then they tend to take a particularly simple form.* In general, the larger the group of transformations the simpler the form taken by the laws of physics. One of the major goals of modern physics is to find the largest possible group of transformations under which the laws of physics are invariant, and then prove that when expressed in a manner which makes this invariance manifest these laws reduce to a single unifying principle.

We already know how to write the laws of physics in terms of vectors and vector fields. This means that these laws are automatically invariant under translations and rotations. However, according to the relativity principle, there is a third class of transformations under which the laws of physics must also be invariant; namely, Lorentz transformations. There are two ways in which we could verify that the laws of physics are Lorentz invariant. The direct method is extremely tedious, since Lorentz transformations are rather complicated. An alternative method is to write the laws of physics in terms of geometric objects which transform as tensors under translations, rotations, *and* Lorentz transformations. This method has the advantage that it makes the Lorentz invariant nature of the laws of physics obvious. We also expect that when the laws of physics are written in manifestly Lorentz invariant form then they will look even simpler than they do when written just in terms of vectors. The laws of electromagnetism provide a particularly good illustration of this effect.

2.7 Space-time

In special relativity we are only allowed to use inertial frames to assign coordinates to events. There are many different types of inertial frames. However, it is convenient to adhere to those with *standard coordinates*. That is, spatial coordinates which are right-handed rectilinear Cartesians based on a standard unit of length and time-scales based on a standard unit of time. We shall continue to assume that we are employing standard coordinates. However, from now on we shall make no assumptions, unless specifically stated, about the relative configuration of the two sets of spatial axes and the origins of time when dealing with two inertial frames. Thus, the most general transformation between two inertial frames consists of a Lorentz transformation in the standard configuration plus a translation (this includes a translation in time) and a rotation of the coordinate axes. The resulting transformation is called a *general Lorentz transformation*, as opposed to a Lorentz transformation in the standard configuration which will henceforth be termed a *standard Lorentz transformation.*

In Section 2.2 we proved quite generally that corresponding differentials in two inertial frames S and S' satisfy the relation

$$dx^2 + dy^2 + dz^2 - c^2 dt^2 = dx'^2 + dy'^2 + dz'^2 - c^2 dt'^2. \tag{2.51}$$

Thus, we expect this relation to remain invariant under a general Lorentz transformation. Since such a transformation is *linear* it follows that

$$\begin{aligned}(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 - c^2(t_2 - t_1)^2 = \\ (x_2' - x_1')^2 + (y_2' - y_1')^2 + (z_2' - z_1')^2 - c^2(t_2' - t_1')^2,\end{aligned} \tag{2.52}$$

where (x_1, y_1, z_1, t_1) and (x_2, y_2, z_2, t_2) are the coordinates of any two events in S and the primed symbols denote the corresponding coordinates in S'. It is convenient to write

$$-dx^2 - dy^2 - dz^2 + c^2 dt^2 = ds^2, \tag{2.53}$$

and

$$-(x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2 + c^2(t_2 - t_1)^2 = s^2. \tag{2.54}$$

The differential ds, or the finite number s, defined by these equations is called the *interval* between the corresponding events. Equations (2.51) and (2.52) express

the fact that *the interval between two events is invariant*, in the sense that it has the same value in all inertial frames. In other words, the interval between two events is invariant under a general Lorentz transformation.

Let us consider entities defined in terms of four variables

$$x^1 = x, \quad x^2 = y, \quad x^3 = z, \quad x^4 = ct, \tag{2.55}$$

and which transform as tensors (see Eqs. (2.30)–(2.32)) under a general Lorentz transformation. From now on such entities will be referred to as *4-tensors*.

Tensor analysis cannot proceed very far without the introduction of a non-singular tensor g_{ij}, the so-called *fundamental tensor*, which is used to define the operations of raising and lowering suffixes (see Eqs. (2.42)–(2.44)). The fundamental tensor is usually introduced using a metric $ds^2 = g_{ij}\, dx^i dx^j$, where ds^2 is a differential invariant. We have already come across such an invariant, namely

$$\begin{aligned} ds^2 &= -dx^2 - dy^2 - dz^2 + c^2 dt^2 \\ &= -(dx^1)^2 - (dx^2)^2 - (dx^3)^2 + (dx^4)^2 \\ &= g_{\mu\nu}\, dx^\mu dx^\nu, \end{aligned} \tag{2.56}$$

where μ, ν run from 1 to 4. Note that the use of Greek suffixes is conventional in 4-tensor theory. Roman suffixes are reserved for tensors in three dimensional Euclidian space, so-called 3-tensors. The 4-tensor $g_{\mu\nu}$ has the components $g_{11} = g_{22} = g_{33} = -1, g_{44} = 1$, and $g_{\mu\nu} = 0$ when $\mu \neq \nu$, in all permissible coordinate frames. From now on $g_{\mu\nu}$, as defined above, is adopted as the fundamental tensor for 4-tensors. $g_{\mu\nu}$ can be thought of as the *metric tensor* of the "space" whose points are the events (x^1, x^2, x^3, x^4). This "space" is usually referred to as *space-time*, for obvious reasons. Note that space-time cannot be regarded as a straightforward generalization of Euclidian 3-space to four dimensions, with time as the fourth dimension. The distribution of signs in the metric ensures that the time coordinate x^4 is not on the same footing as the three space coordinates. Thus, space-time has a non-isotropic nature which is quite unlike Euclidian space with its positive definite metric. According to the relativity principle, all physical laws are expressible as interrelationships between 4-tensors in space-time.

A tensor of rank one is called a *4-vector*. We shall also have occasion to use ordinary vectors in three dimensional Euclidian space. Such vectors are called *3-vectors* and are conventionally represented by boldface symbols. We shall use the Latin suffixes i, j, k, *etc.* to denote the components of a 3-vector; these suffixes are understood to range from 1 to 3. Thus, $\boldsymbol{u} = u^i = dx^i/dt$ denotes a velocity vector. For 3-vectors we shall use the notation $u^i = u_i$ interchangeably; *i.e.*, the level of the suffix has no physical significance.

When tensor transformations from one frame to another actually have to be computed, we shall usually find it possible to choose coordinates in the standard configuration, so that the standard Lorentz transform applies. Under it, any contravariant 4-vector T^μ transforms according to the same scheme as the difference in coordinates $x_2^\mu - x_1^\mu$ between two points in space-time. It follows that

$$T^{1'} = \gamma(T^1 - \beta\, T^4), \tag{2.57a}$$

$$T^{2'} = T^2, \tag{2.57b}$$

$$T^{3'} = T^3, \tag{2.57c}$$

$$T^{4'} = \gamma(T^4 - \beta\, T^1), \tag{2.57d}$$

where $\beta = v/c$. Higher rank 4-tensors transform according to the rules (2.30)–(2.32). The transformation coefficients take the form

$$p_\mu^{\mu'} = \left(\begin{array}{cccc} \gamma & 0 & 0 & -\gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma\beta & 0 & 0 & \gamma \end{array}\right) \tag{2.58a}$$

$$p_{\mu'}^{\mu} = \left(\begin{array}{cccc} \gamma & 0 & 0 & \gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \gamma\beta & 0 & 0 & \gamma \end{array}\right) \tag{2.58b}$$

Often the first three components of a 4-vector coincide with the components of a 3-vector. For example, the x^1, x^2, x^3 in $R^\mu = (x^1, x^2, x^3, x^4)$ are the components of $\boldsymbol{r}$, the position 3-vector of the point at which the event occurs. In such cases

we adopt the notation exemplified by $R^\mu = (\boldsymbol{r}, ct)$. The covariant form of such a vector is simply $R_\mu = (-\boldsymbol{r}, ct)$. The squared magnitude of the vector is $(R)^2 = R_\mu R^\mu = -r^2 + c^2 t^2$. The inner product $g_{\mu\nu} R^\mu Q^\nu = R_\mu Q^\mu$ of R^μ with a similar vector $Q^\mu = (\boldsymbol{q}, k)$ is given by $R_\mu Q^\mu = -\boldsymbol{r}\cdot\boldsymbol{q} + ct\, k$. The vectors R^μ and Q^μ are said to be *orthogonal* if $R_\mu Q^\mu = 0$.

Since a general Lorentz transformation is a *linear* transformation, the partial derivative of a 4-tensor is also a 4-tensor;

$$\frac{\partial A^{\nu\sigma}}{\partial x^\mu} = A^{\nu\sigma}{}_{,\mu}. \tag{2.59}$$

Clearly, a general 4-tensor acquires an extra covariant index after partial differentiation with respect to the contravariant coordinate x^μ. It is helpful to define a covariant derivative operator

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \left(\nabla, \frac{1}{c}\frac{\partial}{\partial t}\right), \tag{2.60}$$

where

$$\partial_\mu A^{\nu\sigma} \equiv A^{\nu\sigma}{}_{,\mu}. \tag{2.61}$$

There is a corresponding contravariant derivative operator

$$\partial^\mu \equiv \frac{\partial}{\partial x_\mu} = \left(-\nabla, \frac{1}{c}\frac{\partial}{\partial t}\right), \tag{2.62}$$

where

$$\partial^\mu A^{\nu\sigma} \equiv g^{\mu\tau} A^{\nu\sigma}{}_{,\tau}. \tag{2.63}$$

The 4-divergence of a 4-vector $A^\mu = (\boldsymbol{A}, A^0)$ is the invariant

$$\partial^\mu A_\mu = \partial_\mu A^\mu = \nabla\cdot\boldsymbol{A} + \frac{1}{c}\frac{\partial A^0}{\partial t}. \tag{2.64}$$

The four dimensional Laplacian operator, or *d'Alembertian*, is equivalent to the invariant contraction

$$\Box \equiv \partial_\mu \partial^\mu = -\nabla^2 + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}. \tag{2.65}$$

Recall that we still need to prove (from Section 2.2) that the invariance of the differential metric,

$$ds^2 = dx'^2 + dy'^2 + dz'^2 - c^2 dt'^2 = dx^2 + dy^2 + dz^2 - c^2 dt^2, \tag{2.66}$$

between two general inertial frames implies that the coordinate transformation between such frames is necessarily linear. To put it another way, we need to demonstrate that a transformation which transforms a metric $g_{\mu\nu}\, dx^\mu dx^\nu$ with constant coefficients into a metric $g_{\mu'\nu'}\, dx^{\mu'} dx^{\nu'}$ with constant coefficients must be linear. Now

$$g_{\mu\nu} = g_{\mu'\nu'}\, p^{\mu'}_\mu\, p^{\nu'}_\nu. \tag{2.67}$$

Differentiating with respect to x^σ we get

$$g_{\mu'\nu'}\, p^{\mu'}_{\mu\sigma} p^{\nu'}_\nu + g_{\mu'\nu'}\, p^{\mu'}_\mu\, p^{\nu'}_{\nu\sigma} = 0, \tag{2.68}$$

where

$$p^{\mu'}_{\mu\sigma} = \frac{\partial p^{\mu'}_\mu}{\partial x^\sigma} = \frac{\partial^2 x^{\mu'}}{\partial x^\mu \partial x^\sigma} = p^{\mu'}_{\sigma\mu} \tag{2.69}$$

etc. Interchanging the indices μ and σ yields

$$g_{\mu'\nu'}\, p^{\mu'}_{\mu\sigma} p^{\nu'}_\nu + g_{\mu'\nu'}\, p^{\mu'}_\sigma\, p^{\nu'}_{\nu\mu} = 0. \tag{2.70}$$

Interchanging the indices ν and σ gives

$$g_{\mu'\nu'}\, p^{\mu'}_\sigma\, p^{\nu'}_{\nu\mu} + g_{\mu'\nu'}\, p^{\mu'}_\mu\, p^{\nu'}_{\nu\sigma} = 0, \tag{2.71}$$

where the indices μ' and ν' have been interchanged in the first term. It follows from Eqs. (2.68), (2.70), and (2.71) that

$$g_{\mu'\nu'}\, p^{\mu'}_{\mu\sigma} p^{\nu'}_\nu = 0. \tag{2.72}$$

Multiplication by $p^\nu_{\sigma'}$ yields

$$g_{\mu'\nu'}\, p^{\mu'}_{\mu\sigma} p^{\nu'}_\nu\, p^\nu_{\sigma'} = g_{\mu'\sigma'}\, p^{\mu'}_{\mu\sigma} = 0. \tag{2.73}$$

Finally, multiplication by $g^{\nu'\sigma'}$ gives

$$g_{\mu'\sigma'} g^{\nu'\sigma'}\, p^{\mu'}_{\mu\sigma} = p^{\nu'}_{\mu\sigma} = 0. \tag{2.74}$$

This proves that the coefficients $p^{\nu'}_\mu$ are constants and, hence, that the transformation is linear.

2.8 Proper time

It is often helpful to write the invariant differential interval ds^2 in the form

$$ds^2 = c^2\, d\tau^2. \tag{2.75}$$

The quantity $d\tau$ is called the *proper time*. It follows that

$$d\tau^2 = -\frac{dx^2 + dy^2 + dz^2}{c^2} + dt^2. \tag{2.76}$$

Consider a series of events on the world-line of some material particle. If the particle has speed u then

$$d\tau^2 = dt^2 \left[-\frac{dx^2 + dy^2 + dz^2}{c^2 dt^2} + 1\right] = dt^2 \left(1 - \frac{u^2}{c^2}\right), \tag{2.77}$$

implying that

$$\frac{dt}{d\tau} = \gamma(u). \tag{2.78}$$

It is clear that $dt = d\tau$ in the particle's rest frame. Thus, $d\tau$ corresponds to the time difference between two neighbouring events on the particle's world-line, as measured by a clock attached to the particle (hence, the name "proper time"). According to Eq. (2.78), the particle's clock appears to run slow, by a factor $\gamma(u)$, in an inertial frame in which the particle is moving with velocity u. This is the celebrated *time dilation* effect.

Let us consider how a small 4-dimensional volume element in space-time transforms under a general Lorentz transformation. We have

$$d^4x' = \mathcal{J}\, d^4x, \tag{2.79}$$

where

$$\mathcal{J} = \frac{\partial(x^{1'}, x^{2'}, x^{3'}, x^{4'})}{\partial(x^1, x^2, x^3, x^4)} \tag{2.80}$$

is the Jacobian of the transformation; *i.e.*, the determinant of the transformation matrix $p_\mu^{\mu'}$. A general Lorentz transformation is made up of a standard Lorentz

transformation plus a displacement and a rotation. Thus, the transformation matrix is the *product* of that for a standard Lorentz transformation, a translation, and a rotation. It follows that the Jacobian of a general Lorentz transformation is the product of that for a standard Lorentz transformation, a translation, and a rotation. It is well known that the Jacobian of the latter two transformations is unity, since they are both volume preserving transformations which do not affect time. Likewise, it is easily seen (*e.g.*, by taking the determinant of the transformation matrix (2.58a)) that the Jacobian of a standard Lorentz transformation is also unity. It follows that

$$d^4x' = d^4x \tag{2.81}$$

for a general Lorentz transformation. In other words, a general Lorentz transformation preserves the volume of space-time. Since time is dilated by a factor γ in a moving frame, the volume of space-time can only be preserved if the volume of ordinary 3-space is reduced by the same factor. As is well known, this is achieved by *length contraction* along the direction of motion by a factor γ.

2.9 4-velocity and 4-acceleration

We have seen that the quantity dx^μ/ds transforms as a 4-vector under a general Lorentz transformation (see Eq. (2.47)). Since $ds \propto d\tau$ it follows that

$$U^\mu = \frac{dx^\mu}{d\tau} \tag{2.82}$$

also transforms as a 4-vector. This quantity is known as the *4-velocity*. Likewise, the quantity

$$A^\mu = \frac{d^2x^\mu}{d\tau^2} = \frac{dU^\mu}{d\tau} \tag{2.83}$$

is a 4-vector, and is called the *4-acceleration*.

For events along the world-line of a particle traveling with 3-velocity $\boldsymbol{u}$ we have

$$U^\mu = \frac{dx^\mu}{d\tau} = \frac{dx^\mu}{dt}\frac{dt}{d\tau} = \gamma(u)(\boldsymbol{u}, c), \tag{2.84}$$

where use has been made of Eq. (2.78). This gives the relationship between a particle's 3-velocity and its 4-velocity. The relationship between the 3-acceleration and the 4-acceleration is less straightforward. We have

$$A^\mu = \frac{dU^\mu}{d\tau} = \gamma\,\frac{dU^\mu}{dt} = \gamma\frac{d}{dt}(\gamma\boldsymbol{u}, \gamma c) = \gamma\left(\frac{d\gamma}{dt}\,\boldsymbol{u} + \gamma\boldsymbol{a}, c\,\frac{d\gamma}{dt}\right), \tag{2.85}$$

where $\boldsymbol{a} = d\boldsymbol{u}/dt$ is the 3-acceleration. In the rest frame of the particle $U^\mu = (\boldsymbol{0}, c)$ and $A^\mu = (\boldsymbol{a}, 0)$. It follows that

$$U_\mu A^\mu = 0 \tag{2.86}$$

(note that $U_\mu A^\mu$ is an invariant quantity). In other words, the 4-acceleration of a particle is always orthogonal to its 4-velocity.

2.10 The current density 4-vector

Let us now consider the laws of electromagnetism. We wish to demonstrate that these laws are compatible with the relativity principle. In order to achieve this it is necessary for us to make an *assumption* about the transformation properties of electric charge. The assumption which we shall make, which is well substantiated experimentally, is that charge, unlike mass, is invariant. That is, the charge carried by a given particle has the same measure in all inertial frames. In particular, the charge carried by a particle does not vary with the particle's velocity.

Let us suppose, following Lorentz, that all charge is made up of elementary particles, each carrying the invariant amount e. Suppose that n is the number density of such charges at some given point and time, moving with velocity $\boldsymbol{u}$, as observed in a frame S. Let n_0 be the number density of charges in the frame S_0 in which the charges are momentarily at rest. As is well known, a volume of measure V in S has measure $\gamma(u)\,V$ in S_0 (because of length contraction). Since observers in both frames must agree on how many particles are contained in the volume, and, hence, on how much charge it contains, it follows that $n = \gamma(u)\,n_0$. If $\rho = en$ and $\rho_0 = en_0$ are the charge densities in S and S_0, respectively, then

$$\rho = \gamma(u)\rho_0. \tag{2.87}$$

The quantity ρ_0 is called the *proper density* and is obviously Lorentz invariant.

Suppose that x^μ are the coordinates of the moving charge in S. The *current density 4-vector* is constructed as follows:

$$J^\mu = \rho_0 \frac{dx^\mu}{d\tau} = \rho_0 U^\mu. \tag{2.88}$$

Thus,

$$J^\mu = \rho_0 \, \gamma(u)(\boldsymbol{u}, c) = (\boldsymbol{j}, \rho c), \tag{2.89}$$

where $\boldsymbol{j} = \rho \boldsymbol{u}$ is the current density 3-vector. Clearly, charge density and current density transform as the time-like and space-like components of the same 4-vector.

Consider the invariant 4-divergence of J^μ:

$$\partial_\mu J^\mu = \nabla \cdot \boldsymbol{j} + \frac{\partial \rho}{\partial t}. \tag{2.90}$$

We know that one of the caveats of Maxwell's equations is the charge conservation law

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \boldsymbol{j} = 0. \tag{2.91}$$

It is clear that this expression can be rewritten in the manifestly Lorentz invariant form

$$\partial_\mu J^\mu = 0. \tag{2.92}$$

This equation tells us that there are no net sources or sinks of electric charge in nature; *i.e.*, electric charge is neither created nor destroyed.

2.11 The potential 4-vector

There are many ways of writing the laws of electromagnetism. However, the most obviously Lorentz invariant way is to write them in terms of the vector and scalar potentials. When written in this fashion, Maxwell's equations reduce to

$$\left(-\nabla^2 + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\phi = \frac{\rho}{\epsilon_0}, \tag{2.93a}$$

$$\left(-\nabla^2 + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\boldsymbol{A} = \mu_0 \boldsymbol{j}, \tag{2.93b}$$

where ϕ is the scalar potential and $\boldsymbol{A}$ is the vector potential. Note that the differential operator appearing in these equations is the Lorentz invariant d'Alembertian, defined in Eq. (2.65). The above pair of equations can be rewritten in the form

$$\Box\phi = \frac{\rho\, c}{c\,\epsilon_0}, \tag{2.94a}$$

$$\Box\, c\boldsymbol{A} = \frac{\boldsymbol{j}}{c\,\epsilon_0}. \tag{2.94b}$$

Maxwell's equations can be written in Lorentz invariant form provided that the entity

$$\Phi^\mu = (c\boldsymbol{A}, \phi) \tag{2.95}$$

transforms as a contravariant 4-vector. This entity is known as the *potential 4-vector*. It follows from Eqs. (2.89), (2.94), and (2.95) that

$$\Box\Phi^\mu = \frac{J^\mu}{c\,\epsilon_0}. \tag{2.96}$$

Thus, the field equations which govern classical electromagnetism can all be summed up in a single 4-vector equation.

2.12 Gauge invariance

The electric and magnetic fields are obtained from the vector and scalar potentials according to the prescription

$$\boldsymbol{E} = -\nabla\phi - \frac{\partial \boldsymbol{A}}{\partial t}, \tag{2.97a}$$

$$\boldsymbol{B} = \nabla\wedge\boldsymbol{A}. \tag{2.97b}$$

These fields are important because they determine the electromagnetic forces exerted on charged particles. Note that the above prescription does not uniquely determine the two potentials. It is possible to make the following transformation, known as a *gauge transformation*, which leaves the fields unaltered:

$$\phi \rightarrow \phi + \frac{\partial\psi}{\partial t}, \tag{2.98a}$$

$$\boldsymbol{A} \rightarrow \boldsymbol{A} - \nabla\psi, \tag{2.98b}$$

where $\psi(\boldsymbol{r}, t)$ is a general scalar field. It is necessary to adopt some form of convention, generally known as a *gauge condition*, to fully specify the two potentials. In fact, there is only one gauge condition which is consistent with Eqs. (2.93). This is the *Lorentz gauge condition*,

$$\frac{1}{c^2}\frac{\partial \phi}{\partial t} + \nabla\cdot\boldsymbol{A} = 0. \tag{2.99}$$

Note that this condition can be written in the Lorentz invariant form

$$\partial_\mu \Phi^\mu = 0. \tag{2.100}$$

This implies that if the Lorentz gauge holds in one particular inertial frame then it automatically holds in all other inertial frames. A general gauge transformation can be written

$$\Phi^\mu \rightarrow \Phi^\mu + c\,\partial^\mu \psi. \tag{2.101}$$

Note that even after the Lorentz gauge has been adopted the potentials are undetermined to a gauge transformation using a scalar field ψ which satisfies the sourceless wave equation

$$\Box\psi = 0. \tag{2.102}$$

However, if we adopt "sensible" boundary conditions in both space and time then the only solution to the above equation is $\psi = 0$.

2.13 Solution of the inhomogeneous wave equation

Equations (2.93) all have the general form

$$\Box\psi(\boldsymbol{r}, t) = g(\boldsymbol{r}, t). \tag{2.103}$$

Can we find a *unique* solution to the above equation? Let us assume that the source function $g(\boldsymbol{r}, t)$ can be expressed as a Fourier integral

$$g(\boldsymbol{r}, t) = \int_{-\infty}^{\infty} g_\omega(\boldsymbol{r})\, e^{-\mathrm{i}\,\omega t}\, d\omega. \tag{2.104}$$

The inverse transform is

$$g_\omega(\boldsymbol{r}) = \frac{1}{2\pi}\int_{-\infty}^{\infty} g(\boldsymbol{r},t)\, e^{\,\mathrm{i}\,\omega t}\, dt. \tag{2.105}$$

Similarly, we may write the general potential $\psi(\boldsymbol{r},t)$ as a Fourier integral

$$\psi(\boldsymbol{r},t) = \int_{-\infty}^{\infty} \psi_\omega(\boldsymbol{r})\, e^{-\mathrm{i}\,\omega t}\, d\omega, \tag{2.106}$$

with the corresponding inverse

$$\psi_\omega(\boldsymbol{r}) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \psi(\boldsymbol{r},t)\, e^{\,\mathrm{i}\,\omega t}\, dt. \tag{2.107}$$

Fourier transformation of Eq. (2.103) yields

$$(\nabla^2 + k^2)\psi_\omega = -g_\omega, \tag{2.108}$$

where $k = \omega/c$.

The above equation, which reduces to Poisson's equation in the limit $k \to 0$, and is called *Helmholtz's equation*, is linear, so we may attempt a Green's function method of solution. Let us try to find a function $G_\omega(\boldsymbol{r},\boldsymbol{r}')$ such that

$$(\nabla^2 + k^2)G_\omega(\boldsymbol{r},\boldsymbol{r}') = -\delta(\boldsymbol{r}-\boldsymbol{r}'). \tag{2.109}$$

The general solution is then

$$\psi_\omega(\boldsymbol{r}) = \int g_\omega(\boldsymbol{r}')\, G_\omega(\boldsymbol{r},\boldsymbol{r}')\, dV'. \tag{2.110}$$

The "sensible" spatial boundary conditions which we impose are that $G_\omega(\boldsymbol{r},\boldsymbol{r}') \to 0$ as $|\boldsymbol{r}-\boldsymbol{r}'| \to \infty$. In other words, the field goes to zero a long way from the source. Since the system we are solving is spherically symmetric about the point $\boldsymbol{r}'$ it is plausible that the Green's function itself is spherically symmetric. It follows that

$$\frac{1}{R}\frac{d^2(R\,G_\omega)}{dR^2} + k^2 G_\omega = -\delta(\boldsymbol{R}), \tag{2.111}$$

where $\boldsymbol{R} = \boldsymbol{r} - \boldsymbol{r}'$ and $R = |\boldsymbol{R}|$. The most general solution to the above equation in the region $R > 0$ is[1]

$$G_\omega(R) = \frac{A\, e^{\,\mathrm{i}\,kR} + B\, e^{-\mathrm{i}\,kR}}{4\pi\, R}. \tag{2.112}$$

We know that in the limit $k \to 0$ the Green's function for Helmholtz's equation must tend towards that for Poisson's equation, which is

$$G_\omega(R) = \frac{1}{4\pi\, R}. \tag{2.113}$$

This is only the case if $A + B = 1$.

Reconstructing $\psi(\boldsymbol{r}, t)$ from Eqs. (2.106), (2.110), and (2.112), we obtain

$$\psi(\boldsymbol{r}, t) = \frac{1}{4\pi} \int\!\!\int \frac{g_\omega(\boldsymbol{r}')}{R} \left[A\, e^{-\mathrm{i}\,\omega(t-R/c)} + B\, e^{-\mathrm{i}\,\omega(t+R/c)}\right] d\omega\, dV'. \tag{2.114}$$

It follows from Eq. (2.104) that

$$\psi(\boldsymbol{r}, t) = \frac{A}{4\pi} \int \frac{g(\boldsymbol{r}', t - R/c)}{R}\, dV' + \frac{B}{4\pi} \int \frac{g(\boldsymbol{r}', t + R/c)}{R}\, dV'. \tag{2.115}$$

Now, the real space Green's function for the inhomogeneous wave equation (2.103) satisfies

$$\Box G(\boldsymbol{r}, \boldsymbol{r}'; t, t') = \delta(\boldsymbol{r} - \boldsymbol{r}')\, \delta(t - t'). \tag{2.116}$$

Hence, the most general solution of this equation takes the form

$$\psi(\boldsymbol{r}, t) = \int\!\!\int g(\boldsymbol{r}', t')\, G(\boldsymbol{r}, \boldsymbol{r}'; t, t')\, dV' dt'. \tag{2.117}$$

Comparing Eqs. (2.115) and (2.117) we obtain

$$G(\boldsymbol{r}, \boldsymbol{r}'; t, t') = A\, G^{(+)}(\boldsymbol{r}, \boldsymbol{r}'; t, t') + B\, G^{(-)}(\boldsymbol{r}, \boldsymbol{r}'; t, t'), \tag{2.118}$$

[1] In principle, $A = A(\omega)$ and $B = B(\omega)$, with $A + B = 1$. However, later on we shall demonstrate that $B = 0$, otherwise causality is violated. It follows that $A = 1$. Thus, it is legitimate to assume, for the moment, that A and B are constants.

where

$$G^{(\pm)}(\boldsymbol{r}, \boldsymbol{r}'; t, t') = \frac{\delta(t' - [t \mp |\boldsymbol{r} - \boldsymbol{r}'|/c])}{4\pi\, |\boldsymbol{r} - \boldsymbol{r}'|}, \tag{2.119}$$

and $A + B = 1$.

The real space Green's function specifies the response of the system to a point source at position $\mathbf{r}'$ which appears momentarily at time t'. According to the *retarded Green's function* $G^{(+)}$ the response consists of a spherical wave, centred on $\boldsymbol{r}'$, which propagates forward in time. In order for the wave to reach position $\boldsymbol{r}$ at time t it must have been emitted from the source at $\boldsymbol{r}'$ at the *retarded time* $t_r = t - |\mathbf{r} - \mathbf{r}'|/c$. According to the *advanced Green's function* $G^{(-)}$ the response consists of a spherical wave, centred on $\boldsymbol{r}'$, which propagates backward in time. Clearly, the advanced potential is not consistent with our ideas about causality, which demand that an effect can never precede its cause in time. Thus, the Green's function which is consistent with our experience is

$$G(\boldsymbol{r}, \boldsymbol{r}'; t, t') = G^{(+)}(\boldsymbol{r}, \boldsymbol{r}'; t, t') = \frac{\delta(t' - [t - |\boldsymbol{r} - \boldsymbol{r}'|/c])}{4\pi\, |\boldsymbol{r} - \boldsymbol{r}'|}. \tag{2.120}$$

We are able to find solutions of the inhomogeneous wave equation (2.103) which propagate backward in time because this equation is time symmetric (*i.e.*, it is invariant under the transformation $t \to -t$).

In conclusion, the most general solution of the inhomogeneous wave equation (2.103) which satisfies sensible boundary conditions at infinity and is consistent with causality is

$$\psi(\boldsymbol{r}, t) = \int \frac{g(\boldsymbol{r}', t - |\boldsymbol{r} - \boldsymbol{r}'|/c)}{4\pi\, |\boldsymbol{r} - \boldsymbol{r}'|}\, dV'. \tag{2.121}$$

This expression is sometimes written

$$\psi(\boldsymbol{r}, t) = \int \frac{[g(\boldsymbol{r}')]}{4\pi\, |\boldsymbol{r} - \boldsymbol{r}'|}\, dV', \tag{2.122}$$

where the rectangular bracket symbol [ ] denotes that the terms inside the bracket are to be evaluated at the retarded time $t - |\boldsymbol{r} - \boldsymbol{r}'|/c$. Note, in particular, from Eq. (2.122) that if there is no source (*i.e.*, $g(\boldsymbol{r}, t) = 0$) then there is no field (*i.e.*, $\psi(\boldsymbol{r}, t) = 0$). But, is the above solution really *unique*? Unfortunately, there is a

weak link in our derivation, between Eqs. (2.110) and (2.111), where we *assume* that the Green's function for the Helmholtz equation subject to the boundary condition $G_\omega(\boldsymbol{r}, \boldsymbol{r}') \to 0$ as $|\boldsymbol{r} - \boldsymbol{r}'| \to \infty$ is *spherically symmetric*. Let us try to fix this problem.

With the benefit of hindsight, we can see that the Green's function

$$G_\omega(R) = \frac{e^{\mathrm{i}\,kR}}{4\pi\,R} \tag{2.123}$$

corresponds to the retarded solution in real space and is, therefore, the correct physical Green's function. The Green's function

$$G_\omega(R) = \frac{e^{-\mathrm{i}\,kR}}{4\pi\,R} \tag{2.124}$$

corresponds to the advanced solution in real space and must, therefore, be rejected. We can select the retarded Green's function by imposing the following boundary condition at infinity

$$\lim_{R\to\infty} R\left(\frac{\partial G}{\partial R} - \mathrm{i}\,kG\right) = 0. \tag{2.125}$$

This is called the *Sommerfeld radiation condition*; it basically ensures that sources radiate waves instead of absorbing them. But, does this boundary condition *uniquely* select the spherically symmetric Green's function (2.123) as the solution of

$$(\nabla^2 + k^2)G_\omega(R, \theta, \varphi) = -\delta(\boldsymbol{R})? \tag{2.126}$$

Here, (R, θ, φ) are spherical polar coordinates. If it does then we can be sure that Eq. (2.122) represents the *unique* solution of the wave equation (2.103) which is consistent with causality.

Let us suppose that there are two solutions of Eq. (2.126) which satisfy the boundary condition (2.125) and revert to the unique Green's function for Poisson's equation (2.113) in the limit $R \to 0$. Let us call these solutions u_1 and u_2, and let us form the difference $w = u_1 - u_2$. Consider a surface Σ_0 which is a sphere of arbitrarily small radius centred on the origin. Consider a second surface Σ_∞ which is a sphere of arbitrarily large radius centred on the origin. Let V denote

the volume enclosed by these surfaces. The difference function w satisfies the homogeneous Helmholtz equation,

$$(\nabla^2 + k^2) w = 0, \tag{2.127}$$

throughout V. According to Green's theorem

$$\int_V (w\nabla^2 w^* - w^* \nabla^2 w)\, dV = \left(\int_{\Sigma_0} + \int_{\Sigma_\infty}\right)\left(w\frac{\partial w^*}{\partial n} - w^*\frac{\partial w}{\partial n}\right) dS, \tag{2.128}$$

where $\partial/\partial n$ denotes a derivative normal to the surface in question. It is clear from Eq. (2.127) that the volume integral is zero. It is also clear that the first surface integral is zero, since both u_1 and u_2 must revert to the Green's function for Poisson's equation in the limit $R \rightarrow 0$. Thus,

$$\int_{\Sigma_\infty}\left(w\frac{\partial w^*}{\partial n} - w^*\frac{\partial w}{\partial n}\right) dS = 0. \tag{2.129}$$

Equation (2.127) can be written

$$\frac{\partial^2 (R\,w)}{\partial R^2} + \frac{D(R\,w)}{R^2} + k^2\, Rw = 0, \tag{2.130}$$

where D is the spherical harmonic operator

$$D = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}. \tag{2.131}$$

The most general solution of Eq. (2.130) takes the form (see Section 7)

$$w(R,\theta,\varphi) = \sum_{l,m=0}^{\infty}\left[C_{lm}\, h_l^{(1)}(kR) + D_{lm}\, h_l^{(2)}(kR)\right] Y_{lm}(\theta,\varphi). \tag{2.132}$$

Here, the C_{lm} and D_{lm} are arbitrary coefficients, the Y_{lm} are spherical harmonics,[2] and

$$h_l^{(1,2)}(\rho) = \sqrt{\frac{\pi}{2\rho}}\, H^{1,2}_{l+1/2}(\rho), \tag{2.133}$$

[2] J.D. Jackson, *Classical Electrodynamics*, (Wiley, 1962), p. 99

where $H_n^{1,2}$ are Hankel functions of the first and second kind.[3] It can be demonstrated that[4]

$$H_n^1(\rho) = \sqrt{\frac{2}{\pi\rho}}\, e^{\,\mathrm{i}\,(\rho-(n+1/2)\pi/2)} \sum_{m=0,1,2,\cdots} \frac{(n,m)}{(-2\mathrm{i}\,\rho)^m}, \tag{2.134a}$$

$$H_n^2(\rho) = \sqrt{\frac{2}{\pi\rho}}\, e^{-\mathrm{i}\,(\rho-(n+1/2)\pi/2)} \sum_{m=0,1,2,\cdots} \frac{(n,m)}{(+2\mathrm{i}\,\rho)^m}, \tag{2.134b}$$

where

$$(n,m) = \frac{(4n^2-1)(4n^2-9)\cdots(4n^2-\{2m-1\}^2)}{2^{2m}\,m!} \tag{2.135}$$

and $(n,0)=1$. Note that the summations in Eqs. (2.314) terminate after $n+1/2$ terms.

The large R behaviour of the $h_l^{(2)}$ is clearly inconsistent with the Sommerfeld radiation condition (2.125). It follows that all of the D_{lm} in Eq. (2.132) are zero. The most general solution can now be expressed in the form

$$w(R,\theta,\varphi) = \frac{e^{\,\mathrm{i}\,kR}}{R} \sum_{n=0}^{\infty} \frac{f_n(\theta,\varphi)}{R^n}, \tag{2.136}$$

where the $f_n(\theta,\varphi)$ are various weighted sums of the spherical harmonics. Substitution of this solution into the differential equation (2.130) yields

$$e^{\,\mathrm{i}\,kR} \sum_{n=0}^{\infty} \left(-\frac{2\mathrm{i}\,k\,n}{R^{n+1}} + \frac{n(n+1)}{R^{n+2}} + \frac{D}{R^{n+2}} \right) f_n = 0. \tag{2.137}$$

Replacing the index of summation n in the first term of the parentheses by $n+1$ we obtain

$$e^{\,\mathrm{i}\,kR} \sum_{n=0}^{\infty} \frac{-2\mathrm{i}k\,(n+1) f_{n+1} + [n(n+1)+D] f_n}{R^{n+2}} = 0, \tag{2.138}$$

[3] J.D. Jackson, *Classical Electrodynamics*, (Wiley, 1962), p. 104

[4] A. Sommerfeld, *Partial differential equations in physics*, (Academic Press, New York, 1964), p. 117

which gives us the recursion relation

$$2\mathrm{i}\,k(n+1)f_{n+1} = [n(n+1)+D]f_n. \tag{2.139}$$

It follows that if $f_0 = 0$ then all of the f_n are equal to zero.

Let us now consider the surface integral (2.129). Since we are interested in the limit $R \to \infty$ we can replace w by the first term of its expansion in (2.136), so

$$\int_{\Sigma_\infty} \left(w\frac{\partial w^*}{\partial n} - w^*\frac{\partial w}{\partial n}\right) dS = -2\mathrm{i}\,k \int |f_0|^2\, d\Omega = 0, \tag{2.140}$$

where $d\Omega$ is a unit of solid angle. It is clear that $f_0 = 0$. This implies that $f_1 = f_2 = \cdots = 0$ and, hence, that $w = 0$. Thus, there is only one solution of Eq. (2.126) which is consistent with the Sommerfeld radiation condition, and this is given by Eq. (2.123). We can now be sure that Eq. (2.122) is a *unique* solution of Eq. (2.103) subject to the boundary condition (2.125). This boundary condition basically says that infinity is an absorber of radiation but not an emitter, which seems entirely reasonable.

2.14 Retarded potentials

Equations (2.94) have the same form as the inhomogeneous wave equation (2.103), so we can immediately write the solutions to these equations as

$$\phi(\boldsymbol{r},t) = \frac{1}{4\pi\,\epsilon_0}\int \frac{[\rho(\boldsymbol{r}')]}{|\boldsymbol{r}-\boldsymbol{r}'|}\,dV', \tag{2.141a}$$

$$\boldsymbol{A}(\boldsymbol{r},t) = \frac{\mu_0}{4\pi}\int \frac{[\boldsymbol{j}(\boldsymbol{r}')]}{|\boldsymbol{r}-\boldsymbol{r}'|}\,dV'. \tag{2.141b}$$

Moreover, we can be sure that these solutions are *unique*, subject to the reasonable proviso that infinity is an absorber of radiation but not an emitter. This is a crucially important point. Whenever the above solutions are presented in physics textbooks there is a tacit assumption that they are unique. After all, if they were not unique why should we choose to study them instead of one of the other possible solutions? The uniqueness of the above solutions has a physical

interpretation. It is clear from Eqs. (2.141) that in the absence of any charges and currents there are no electromagnetic fields. In other words, if we observe an electromagnetic field we can be certain that if we were to trace it backward in time we would eventually discover that it was emitted by a charge or a current. In proving that the solutions of Maxwell's equations are unique, and then finding a solution in which all waves are emitted by sources, we have effectively ruled out the possibility that the vacuum can be "unstable" to the production of electromagnetic waves without the need for any sources.

Equations (2.141) can be combined to form the solution of the 4-vector wave equation (2.96),

$$\Phi^\mu = \frac{1}{4\pi\,\epsilon_0\,c}\int\frac{[J^\mu]}{r}\,dV. \tag{2.142}$$

Here, the components of the 4-potential are evaluated at some event P in space-time, r is the distance of the volume element dV from P, and the square brackets indicate that the 4-current is to be evaluated at the retarded time; *i.e.*, at a time r/c before P.

But, does the right-hand side of Eq. (2.142) really transform as a contravariant 4-vector? This is not a trivial question since volume integrals in 3-space are not, in general, Lorentz invariant due to the length contraction effect. However, the integral in Eq. (2.142) is not a straightforward volume integral because the integrand is evaluated at the retarded time. In fact, the integral is best regarded as an integral over events in space-time. The events which enter the integral are those which intersect a spherical light wave launched from the event P and evolved backwards in time. In other words, the events occur before the event P and have zero interval with respect to P. It is clear that observers in all inertial frames will, at least, agree on which events are to be included in the integral, since both the interval between events and the absolute order in which events occur are invariant under a general Lorentz transformation.

We shall now demonstrate that all observers obtain the same value of dV/r for each elementary contribution to the integral. Suppose that S and S' are two inertial frames in the standard configuration. Let unprimed and primed symbols denote corresponding quantities in S and S', respectively. Let us assign coordinates $(0, 0, 0, 0)$ to P and (x, y, z, ct) to the retarded event Q for which r

and dV are evaluated. Using the standard Lorentz transformation (2.19), the fact that the interval between events P and Q is zero, and the fact that both t and t' are negative, we obtain

$$r' = -ct' = -c\gamma\left(t - \frac{vx}{c^2}\right), \tag{2.143}$$

where v is the relative velocity between frames S' and S, γ is the Lorentz factor, and $r = \sqrt{x^2 + y^2 + z^2}$, *etc.* It follows that

$$r' = r\gamma\left(-\frac{ct}{r} + \frac{vx}{cr}\right) = r\gamma\left(1 + \frac{v}{c}\cos\theta\right), \tag{2.144}$$

where θ is the angle (in 3-space) subtended between the line PQ and the x-axis.

We now know the transformation for r. What about the transformation for dV? We might be tempted to set $dV' = \gamma\, dV$, according to the usual length contraction rule. However, this is wrong. The contraction by a factor γ only applies if the whole of the volume is measured at the same time, which is not the case in the present problem. Now, the dimensions of dV along the $y-$ and $z-$ axes are the same in both S and S', according to Eqs. (2.19). For the x-dimension these equations give $dx' = \gamma(dx - v\, dt)$. The extremities of dx are measured at times differing by dt, where[5]

$$dt = -\frac{dr}{c} = -\frac{dx}{c}\cos\theta. \tag{2.145}$$

Thus,

$$dx' = \left(1 + \frac{v}{c}\cos\theta\right)\gamma\, dx, \tag{2.146}$$

giving

$$dV' = \left(1 + \frac{v}{c}\cos\theta\right)\gamma\, dV. \tag{2.147}$$

It follows from Eqs. (2.144) and (2.147) that $dV'/r' = dV/r$. This result will clearly remain valid even when S and S' are not in the standard configuration.

[5]Note that $dr = dx\cos\theta$, despite the fact that $x = r\cos\theta$. This comes about because the volume element dV is aligned along a radius vector.

Thus, dV/r is an invariant and, therefore, $[J^\mu]\,dV/r$ is a contravariant 4-vector. For linear transformations, such as a general Lorentz transformation, the result of adding 4-tensors evaluated at different 4-points is itself a 4-tensor. It follows that the right-hand side of Eq. (2.142) is a contravariant 4-vector. Thus, this 4-vector equation can be properly regarded as the solution to the 4-vector wave equation (2.96).

2.15 Tensors and pseudo-tensors

The totally antisymmetric fourth rank tensor is defined

$$\epsilon^{\alpha\beta\gamma\delta} = \begin{cases} +1 & \text{for } \alpha,\beta,\gamma,\delta \text{ any even permutation of } 1,2,3,4 \\ -1 & \text{for } \alpha,\beta,\gamma,\delta \text{ any odd permutation of } 1,2,3,4 \\ \;\;0 & \text{otherwise} \end{cases} \tag{2.148}$$

The components of this tensor are invariant under a general Lorentz transformation, since

$$\epsilon^{\alpha\beta\gamma\delta}\,p_\alpha^{\alpha'}\,p_\beta^{\beta'}\,p_\gamma^{\gamma'}\,p_\delta^{\delta'} = \epsilon^{\alpha'\beta'\gamma'\delta'}\,|p_\mu^{\mu'}| = \epsilon^{\alpha'\beta'\gamma'\delta'}, \tag{2.149}$$

where $|p_\mu^{\mu'}|$ denotes the determinant of the transformation matrix, or the Jacobian of the transformation, which we have already established is unity for a general Lorentz transformation. We can also define a totally antisymmetric third rank tensor ϵ^{ijk} which stands in the same relation to 3-space as $\epsilon^{\alpha\beta\gamma\delta}$ does to space-time. It is easily demonstrated that the elements of ϵ^{ijk} are invariant under a general translation or rotation of the coordinate axes. The totally antisymmetric third rank tensor is used to define the cross product of two 3-vectors,

$$(\boldsymbol{a}\wedge\boldsymbol{b})^i = \epsilon^{ijk}\,a_j\,b_k, \tag{2.150}$$

and the curl of a 3-vector field,

$$(\nabla\wedge\boldsymbol{A})^i = \epsilon^{ijk}\,\frac{\partial A_k}{\partial x^j}. \tag{2.151}$$

The following two rules are often useful in deriving vector identities

$$\epsilon^{ijk}\epsilon_{iab} = \delta_a^j\,\delta_b^k - \delta_b^j\,\delta_a^k, \tag{2.152a}$$

$$\epsilon^{ijk}\epsilon_{ijb} = 2\,\delta^k_b. \tag{2.152b}$$

Up to now we have restricted ourselves to three basic types of coordinate transformation; namely, translations, rotations, and standard Lorentz transformations. An arbitrary combination of these three transformations constitutes a general Lorentz transformation. Let us now extend our investigations to include a fourth type of transformation known as a parity inversion; *i.e.*, $x, y, z, \rightarrow -x, -y, -z$. A reflection is a combination of a parity inversion and a rotation. As is easily demonstrated, the Jacobian of a parity inversion is -1, unlike a translation, rotation, or standard Lorentz transformation, which all possess Jacobians of $+1$.

The prototype of all 3-vectors is the difference in coordinates between two points in space, $\boldsymbol{r}$. Likewise, the prototype of all 4-vectors is the difference in coordinates between two events in space-time, $R^\mu = (\boldsymbol{r}, ct)$. It is not difficult to appreciate that both of these objects are invariant under a parity transformation (in the sense that they correspond to the same geometric object before and after the transformation). It follows that any 3- or 4-tensor which is directly related to $\boldsymbol{r}$ and R^μ, respectively, is also invariant under a parity inversion. Such tensors include the distance between two points in 3-space, the interval between two points in space-time, 3-velocity, 3-acceleration, 4-velocity, 4-acceleration, and the metric tensor. Tensors which exhibit tensor behaviour under translations, rotations, special Lorentz transformations, *and* are invariant under parity inversions, are termed *proper tensors*, or sometimes *polar tensors*. Since electric charge is clearly invariant under such transformations (*i.e.*, it is a proper scalar) it follows that 3-current and 4-current are proper vectors. It is also clear from Eq. (2.96) that the scalar potential, the vector potential, and the potential 4-vector, are proper tensors.

It follows from Eq. (2.149) that $\epsilon^{\alpha\beta\gamma\delta} \rightarrow -\epsilon^{\alpha\beta\gamma\delta}$ under a parity inversion. Tensors like this, which exhibit tensor behaviour under translations, rotations, and special Lorentz transformations, but are *not* invariant under parity inversions (in the sense that they correspond to different geometric objects before and after the transformation), are called *pseudo-tensors*, or sometimes *axial tensors*. Equations (2.150) and (2.151) imply that the cross product of two proper vectors is a pseudo-vector, and the curl of a proper vector field is a pseudo-vector field.

One particularly simple way of performing a parity transformation is to exchange positive and negative numbers on the three Cartesian axes. A proper vector is unaffected by such a procedure (*i.e.*, its magnitude and direction are the same before and after). On the other hand, a pseudo-vector ends up pointing in the opposite direction after the axes are renumbered.

What is the fundamental difference between proper tensors and pseudo-tensors? The answer is that all pseudo-tensors are defined according to a handedness convention. For instance, the cross product between two vectors is conventionally defined according to a right-hand rule. The only reason for this is that the majority of human beings are right-handed. Presumably, if the opposite were true then cross products *etc.* would be defined according to a left-hand rule and would, therefore, take minus their conventional values. The totally antisymmetric tensor is the prototype pseudo-tensor and is, of course, conventionally defined with respect to a right-handed spatial coordinate system. A parity inversion converts left into right and *vice versa* and, thereby, effectively swaps left- and right-handed conventions.

The use of conventions in physics is perfectly acceptable provided that we recognize that they are conventions and are *consistent* in their use. It follows that laws of physics cannot contain mixtures of tensors and pseudo-tensors, otherwise they would depend our choice of handedness convention.[6]

Let us now consider electric and magnetic fields. We know that

$$\boldsymbol{E} = -\nabla\phi - \frac{\partial \boldsymbol{A}}{\partial t}, \tag{2.153a}$$

$$\boldsymbol{B} = \nabla \wedge \boldsymbol{A}. \tag{2.153b}$$

We have already seen that the scalar and the vector potential are proper scalars and vectors, respectively. It follows that $\boldsymbol{E}$ is a proper vector but that $\boldsymbol{B}$ is a pseudo-vector (since it is the curl of a proper vector). In order to fully appreciate the difference between electric and magnetic fields let us consider a thought

[6]Here, we are assuming that the laws of physics do not possess an intrinsic handedness. This is certainly the case for mechanics and electromagnetism. However, the weak interaction *does* possess an intrinsic handedness; *i.e.*, it is fundamentally different in a parity inverted universe. So, the equations governing the weak interaction do actually contain mixtures of tensors and pseudo-tensors.

experiment first proposed by Richard Feynman. Suppose that we are in radio contact with a race of aliens and are trying to explain to them our system of physics. Suppose, further, that the aliens live sufficiently far away from us that there are no common objects which we both can see. The question is this: could we unambiguously explain to these aliens our concepts of electric and magnetic fields? We could certainly explain electric and magnetic lines of force. The former are the paths of charged particles (assuming that the particles are subject only to electric fields) and the latter can be mapped out using small test magnets. We could also explain how we put arrows on electric lines of force to convert them into electric field lines: the arrows run from positive charges (*i.e.*, charges with the same sign as atomic nuclei) to negative charges. This explanation is unambiguous provided that our aliens live in a matter (rather than an anti-matter) dominated part of the universe. But, could we explain how we put arrows on magnetic lines of force in order to convert them into magnetic field lines? The answer is no. By definition, magnetic field lines emerge from the north poles of permanent magnets and converge on the corresponding south poles. The definition of the north pole of a magnet is simply that it possesses the same magnetic polarity as the north pole of the Earth. This is obviously a convention. In fact, we could redefine magnetic field lines to run from the south poles to the north poles of magnets without significantly altering our laws of physics (we would just have to replace $\boldsymbol{B}$ by $-\boldsymbol{B}$ in all our equations). In a parity inverted universe a north pole becomes a south pole and *vice versa*, so it is hardly surprising that $\boldsymbol{B} \to -\boldsymbol{B}$.[7]

2.16 The electromagnetic field tensor

Let us now investigate whether we can write the components of the electric and magnetic fields as the components of some *proper* 4-tensor. There is an obvious problem here. How can we identify the components of the magnetic field, which is a pseudo-vector, with any of the components of a proper-4-tensor? The former components transform differently under parity inversion than the latter compo-

[7]Note that it would actually be possible to unambiguously communicate to our concepts of left and right to our hypothetical aliens using the fact that the weak interaction possesses an intrinsic handedness.

nents. Consider a proper-3-tensor whose covariant components are written B_{ik}, and which is antisymmetric:

$$B_{ij} = -B_{ji}. \tag{2.154}$$

This immediately implies that all of the diagonal components of the tensor are zero. In fact, there are only three independent non-zero components of such a tensor. Could we, perhaps, use these components to represent the components of a pseudo-3-vector? Let us write

$$B^i = \frac{1}{2}\,\epsilon^{ijk} B_{jk}. \tag{2.155}$$

It is clear that B^i transforms as a contravariant pseudo-3-vector. It is easily seen that

$$B^{ij} = B_{ij} = \begin{pmatrix} 0 & B_z & -B_y \\ -B_z & 0 & B_x \\ B_y & -B_x & 0 \end{pmatrix}, \tag{2.156}$$

where $B^1 = B_1 \equiv B_x$, *etc.* In this manner, we can actually write the components of a pseudo-3-vector as the components of an antisymmetric proper-3-tensor. In particular, we can write the components of the magnetic field $\boldsymbol{B}$ in terms of an antisymmetric proper magnetic field 3-tensor which we shall denote B_{ij}.

Let us now examine Eqs. (2.153) more carefully. Recall that $\Phi_\mu = (-c\boldsymbol{A}, \phi)$ and $\partial_\mu = (\nabla, c^{-1}\partial/\partial t)$. It follows that we can write Eq. (2.153a) in the form

$$E_i = -\partial_i \Phi_4 + \partial_4 \Phi_i. \tag{2.157}$$

Equation (2.153b) can be written

$$cB^i = \frac{1}{2}\,\epsilon^{ijk}\,cB_{jk} = -\epsilon^{ijk}\,\partial_j \Phi_k. \tag{2.158}$$

Let us multiply this expression by ϵ_{iab}, making use of the identity

$$\epsilon_{iab}\,\epsilon^{ijk} = \delta_a^j \delta_b^k - \delta_b^j \delta_a^k. \tag{2.159}$$

We obtain

$$\frac{c}{2}\,(B_{ab} - B_{ba}) = -\partial_a \Phi_b + \partial_b \Phi_a, \tag{2.160}$$

or

$$cB_{ij} = -\partial_i \Phi_j + \partial_j \Phi_i, \tag{2.161}$$

since $B_{ij} = -B_{ji}$.

Let us define a proper-4-tensor whose covariant components are given by

$$F_{\mu\nu} = \partial_\mu \Phi_\nu - \partial_\nu \Phi_\mu. \tag{2.162}$$

It is clear that this tensor is antisymmetric:

$$F_{\mu\nu} = -F_{\nu\mu}. \tag{2.163}$$

This implies that the tensor only possesses six independent non-zero components. Maybe it can be used to specify the components of $\boldsymbol{E}$ and $\boldsymbol{B}$?

Equations (2.157) and (2.162) yield

$$F_{4i} = \partial_4 \Phi_i - \partial_i \Phi_4 = E_i. \tag{2.164}$$

Likewise, Eqs. (2.161) and (2.162) imply that

$$F_{ij} = \partial_i \Phi_j - \partial_j \Phi_i = -cB_{ij}. \tag{2.165}$$

Thus,

$$F_{i4} = -F_{4i} = -E_i, \tag{2.166a}$$

$$F_{ij} = -F_{ji} = -cB_{ij}. \tag{2.166b}$$

In other words, the completely space-like components of the tensor specify the components of the magnetic field, whereas the hybrid space and time-like components specify the components of the electric field. The covariant components of the tensor can be written

$$F_{\mu\nu} = \left(\begin{array}{cccc} 0 & -cB_z & +cB_y & -E_x \\ +cB_z & 0 & -cB_x & -E_y \\ -cB_y & +cB_x & 0 & -E_z \\ +E_x & +E_y & +E_z & 0 \end{array} \right). \tag{2.167}$$

Not surprisingly, $F_{\mu\nu}$ is usually called the *electromagnetic field tensor*. The above expression, which appears in all standard textbooks, is very misleading. Taken at face value, it is simply wrong! We cannot form a proper-4-tensor from the components of a proper-3-vector and a pseudo-3-vector. The expression only makes sense if we interpret B_x, say, as representing the component B_{23} of the proper magnetic field 3-tensor B_{ij}

The contravariant components of the electromagnetic field tensor are given by

$$F^{i4} = -F^{4i} = +E^i, \tag{2.168a}$$

$$F^{ij} = -F^{ji} = -cB^{ij}, \tag{2.168b}$$

or

$$F^{\mu\nu} = \left(\begin{array}{cccc} 0 & -cB_z & +cB_y & +E_x \\ +cB_z & 0 & -cB_x & +E_y \\ -cB_y & +cB_x & 0 & +E_z \\ -E_x & -E_y & -E_z & 0 \end{array} \right). \tag{2.169}$$

Let us now consider two of Maxwell's equations:

$$\nabla\cdot\boldsymbol{E} = \frac{\rho}{\epsilon_0}, \tag{2.170a}$$

$$\nabla\wedge\boldsymbol{B} = \mu_0\left(\boldsymbol{j} + \epsilon_0\,\frac{\partial\boldsymbol{E}}{\partial t}\right). \tag{2.170b}$$

Recall that the 4-current is defined $J^\mu = (\boldsymbol{j}, \rho c)$. The first of these equations can be written

$$\partial_i E^i = \partial_i F^{i4} + \partial_4 F^{44} = \frac{J^4}{c\,\epsilon_0}. \tag{2.171}$$

since $F^{44} = 0$. The second of these equations takes the form

$$\epsilon^{ijk}\,\partial_j\,cB_k - \partial_4 E^i = \epsilon^{ijk}\,\partial_j(1/2\,\epsilon_{kab}\,cB^{ab}) + \partial_4 F^{4i} = \frac{J^i}{c\,\epsilon_0}. \tag{2.172}$$

Making use of Eq. (2.159), the above expression reduces to

$$\frac{1}{2}\,\partial_j(cB^{ij} - cB^{ji}) + \partial_4 F^{4i} = \partial_j F^{ji} + \partial_4 F^{4i} = \frac{J^i}{c\,\epsilon_0}. \tag{2.173}$$

Equations (2.171) and (2.173) can be combined to give

$$\partial_\mu F^{\mu\nu} = \frac{J^\nu}{c\,\epsilon_0}. \tag{2.174}$$

This equation is consistent with the equation of charge continuity, $\partial_\mu J^\mu = 0$, because of the antisymmetry of the electromagnetic field tensor.

2.17 The dual electromagnetic field tensor

We have seen that it is possible to write the components of the electric and magnetic fields as the components of a proper-4-tensor. Is it also possible to write the components of these fields as the components of some *pseudo*-4-tensor? It is obvious that we cannot identify the components of the proper-3-vector $\boldsymbol{E}$ with any of the components of a pseudo-tensor. However, we can represent the components of $\boldsymbol{E}$ in terms of those of an antisymmetric pseudo-3-tensor E_{ij} by writing

$$E^i = \frac{1}{2}\,\epsilon^{ijk} E_{jk}. \tag{2.175}$$

It is easily demonstrated that

$$E^{ij} = E_{ij} = \begin{pmatrix} 0 & E_z & -E_y \\ -E_z & 0 & E_x \\ E_y & -E_x & 0 \end{pmatrix}, \tag{2.176}$$

in a right-handed coordinate system.

Consider the *dual electromagnetic field tensor* $G^{\mu\nu}$, which is defined

$$G^{\mu\nu} = \frac{1}{2}\,\epsilon^{\mu\nu\alpha\beta} F_{\alpha\beta}. \tag{2.177}$$

This tensor is clearly an antisymmetric pseudo-4-tensor. We have

$$G^{4i} = \frac{1}{2}\,\epsilon^{4ijk}\,F_{jk} = -\frac{1}{2}\,\epsilon^{ijk4}\,F_{jk} = \frac{1}{2}\,\epsilon^{ijk}\,cB_{jk} = cB^i, \tag{2.178}$$

plus

$$G^{ij} = \frac{1}{2}\left(\epsilon^{ijk4} F_{k4} + \epsilon^{ij4k} F_{4k}\right) = \epsilon^{ijk} F_{k4}, \tag{2.179}$$

where use has been made of $F_{\mu\nu} = -F_{\nu\mu}$. The above expression yields

$$G^{ij} = -\epsilon^{ijk}\, E_k = -\frac{1}{2}\,\epsilon^{ijk}\epsilon_{kab}\, E^{ab} = -E^{ij}. \tag{2.180}$$

It follows that

$$G^{i4} = -G^{4i} = -cB^i, \tag{2.181a}$$

$$G^{ij} = -G^{ji} = -E^{ij}, \tag{2.181b}$$

or

$$G^{\mu\nu} = \begin{pmatrix} 0 & -E_z & +E_y & -cB_x \\ +E_z & 0 & -E_x & -cB_y \\ -E_y & +E_x & 0 & -cB_z \\ +cB_x & +cB_y & +cB_z & 0 \end{pmatrix}. \tag{2.182}$$

The above expression is, again, slightly misleading, since E_x stands for the component E^{23} of the pseudo-3-tensor E^{ij} and not for an element of the proper-3-vector $\boldsymbol{E}$. Of course, in this case B_x really does represent the first element of the pseudo-3-vector $\boldsymbol{B}$. Note that the elements of $G^{\mu\nu}$ are obtained from those of $F^{\mu\nu}$ by making the transformation $cB^{ij} \rightarrow E^{ij}$ and $E^i \rightarrow -cB^i$.

The covariant elements of the dual electromagnetic field tensor are given by

$$G_{i4} = -G_{4i} = +cB_i, \tag{2.183a}$$

$$G_{ij} = -G_{ji} = -E_{ij}, \tag{2.183b}$$

or

$$G_{\mu\nu} = \begin{pmatrix} 0 & -E_z & +E_y & +cB_x \\ +E_z & 0 & -E_x & +cB_y \\ -E_y & +E_x & 0 & +cB_z \\ -cB_x & -cB_y & -cB_z & 0 \end{pmatrix}. \tag{2.184}$$

The elements of $G_{\mu\nu}$ are obtained from those of $F_{\mu\nu}$ by making the transformation $cB_{ij} \rightarrow E_{ij}$ and $E_i \rightarrow -cB_i$.

Let us now consider the two Maxwell equations

$$\nabla\cdot\boldsymbol{B} = 0, \tag{2.185a}$$

$$\nabla\wedge\boldsymbol{E} = -\frac{\partial\boldsymbol{B}}{\partial t}. \tag{2.185b}$$

The first of these equations can be written

$$-\partial_i\, cB^i = \partial_i G^{i4} + \partial_4 G^{44} = 0, \tag{2.186}$$

since $G^{44} = 0$. The second equation takes the form

$$\epsilon^{ijk}\partial_j E_k = \epsilon^{ijk}\partial_j(1/2\,\epsilon_{kab}E^{ab}) = \partial_j E^{ij} = -\partial_4\, cB^i, \tag{2.187}$$

or

$$\partial_j G^{ji} + \partial_4 G^{4i} = 0. \tag{2.188}$$

Equations (2.186) and (2.188) can be combined to give

$$\partial_\mu G^{\mu\nu} = 0. \tag{2.189}$$

Thus, we conclude that Maxwell's equations for the electromagnetic fields are equivalent to the following pair of 4-tensor equations:

$$\partial_\mu F^{\mu\nu} = \frac{J^\nu}{c\,\epsilon_0}, \tag{2.190a}$$

$$\partial_\mu G^{\mu\nu} = 0. \tag{2.190b}$$

It is obvious from the form of these equations that the laws of electromagnetism are invariant under translations, rotations, special Lorentz transformations, parity inversions, or any combination of these transformations.

2.18 The transformation of electromagnetic fields

The electromagnetic field tensor transforms according to the standard rule

$$F^{\mu'\nu'} = F^{\mu\nu} p_\mu^{\mu'} p_\nu^{\nu'}. \tag{2.191}$$

This easily yields the celebrated rules for transforming electromagnetic fields:

$$E'_{\|} = E_{\|}, \tag{2.192a}$$

$$B'_{\|} = B_{\|}, \tag{2.192b}$$

$$\boldsymbol{E}'_{\perp} = \gamma(\boldsymbol{E}_{\perp} + \boldsymbol{v} \wedge \boldsymbol{B}), \tag{2.192c}$$

$$\boldsymbol{B}'_{\perp} = \gamma(\boldsymbol{B}_{\perp} - \boldsymbol{v} \wedge \boldsymbol{E}/c^2), \tag{2.192d}$$

where $\boldsymbol{v}$ is the relative velocity between the primed and unprimed frames, and the perpendicular and parallel directions are, respectively, perpendicular and parallel to $\boldsymbol{v}$.

At this stage we may conveniently note two important invariants of the electromagnetic field. They are

$$\frac{1}{2}\, F_{\mu\nu} F^{\mu\nu} = c^2 B^2 - E^2, \tag{2.193}$$

and

$$\frac{1}{4}\, G_{\mu\nu} F^{\mu\nu} = c\, \boldsymbol{E}\cdot\boldsymbol{B}. \tag{2.194}$$

The first of these quantities is a proper-scalar and the second is a pseudo-scalar.

2.19 The potential due to a moving charge

Suppose that a particle carrying a charge e moves with *uniform* velocity $\boldsymbol{u}$ through a frame S. Let us evaluate the vector potential $\boldsymbol{A}$ and the scalar potential ϕ due to this charge at a given event P in S.

Let us choose coordinates in S so that $P = (0, 0, 0, 0)$ and $\boldsymbol{u} = (u, 0, 0)$. Let S' be that frame in the standard configuration with respect to S in which the charge is (permanently) at rest, say at the point (x', y', z'). In S' the potential at P is the usual potential due to a stationary charge

$$\boldsymbol{A}' = 0, \tag{2.195a}$$

$$\phi' = \frac{e}{4\pi\epsilon_0\, r'}, \tag{2.195b}$$

where $r' = \sqrt{x'^2 + y'^2 + z'^2}$. Let us now transform these equations directly into the frame S. Since $A^\mu = (c\boldsymbol{A}, \phi)$ is a contravariant 4-vector, its components transform according to the standard rules (2.57). Thus,

$$
\begin{aligned}
cA_1 &= \gamma\left(cA_1' + \frac{u}{c}\,\phi'\right) = \frac{\gamma u\, e}{4\pi\epsilon_0\, c\, r'}, && \text{(2.196a)}\\
cA_2 &= cA_2' = 0, && \text{(2.196b)}\\
cA_3 &= cA_3' = 0, && \text{(2.196c)}\\
\phi &= \gamma\left(\phi' + \frac{u}{c}\, cA_1'\right) = \frac{\gamma\, e}{4\pi\epsilon_0\, r'}, && \text{(2.196d)}
\end{aligned}
$$

since $\beta = -u/c$ in this case. It remains to express the quantity r' in terms of quantities measured in S. The most physically meaningful way of doing this is to express r' in terms of *retarded* values in S. Consider the retarded event at the charge for which, by definition, $r' = -ct'$ and $r = -ct$. Using the standard Lorentz transformation (2.19) we find that

$$
r' = -ct' = -c\gamma(t - ux/c^2) = r\gamma(1 + u_r/c), \tag{2.197}
$$

where $u_r = ux/r = \boldsymbol{r}\cdot\boldsymbol{u}/r$ denotes the radial velocity of the change in S. We can now rewrite Eqs. (2.196) in the form

$$
\begin{aligned}
\boldsymbol{A} &= \frac{\mu_0\, e}{4\pi}\,\frac{[\boldsymbol{u}]}{[r + \boldsymbol{r}\cdot\boldsymbol{u}/c]}, && \text{(2.198a)}\\
\phi &= \frac{e}{4\pi\epsilon_0}\,\frac{1}{[r + \boldsymbol{r}\cdot\boldsymbol{u}/c]}, && \text{(2.198b)}
\end{aligned}
$$

where the square brackets, as usual, indicate that the enclosed quantities must be retarded. For a uniformly moving charge the retardation of $\boldsymbol{u}$ is, of course, superfluous. However, since

$$
\Phi^\mu = \frac{1}{4\pi\epsilon_0\, c}\int \frac{[J^\mu]}{r}\, dV, \tag{2.199}
$$

it is clear that the potentials depend only on the (retarded) velocity of the charge and not on its acceleration. Consequently, the expressions (2.198) give the correct potentials for an *arbitrarily* moving charge. They are known as the *Liénard-Wiechert potentials.*

2.20 The electromagnetic field due to a uniformly moving charge

Although the field generated by a uniformly moving charge can be calculated from the expressions (2.198) for the potentials, it is simpler to calculate it relativistically from first principles.

Let a charge e, whose position vector at time $t = 0$ is $\boldsymbol{r}$, move with uniform velocity $\boldsymbol{u}$ in a frame S whose x-axis has been chosen in the direction of $\boldsymbol{u}$. We require to find the field strengths $\boldsymbol{E}$ and $\boldsymbol{B}$ at the event $P = (0, 0, 0, 0)$. Let S' be that frame in standard configuration with S in which the charge is permanently at rest. In S' the field is given by

$$\boldsymbol{B}' = 0, \tag{2.200a}$$

$$\boldsymbol{E}' = -\frac{e}{4\pi\epsilon_0}\frac{\boldsymbol{r}'}{r'^3}. \tag{2.200b}$$

This field must now be transformed into the frame S. The direct method, using Eqs. (2.192), is somewhat simpler here, but we shall use a somewhat indirect method because of its intrinsic interest.

In order to express Eq. (2.200) in tensor form, we need the electromagnetic field tensor $F^{\mu\nu}$ on the left, and the position 4-vector $R^\mu = (\boldsymbol{r}, ct)$ and the scalar $e/(4\pi\epsilon_0\, r'^3)$ on the right. (We regard r' as an invariant for all observers.) To get a vanishing magnetic field in S' we multiply on the right by the 4-velocity $U^\mu = \gamma(u)(\boldsymbol{u}, c)$, thus tentatively arriving at the equation

$$F^{\mu\nu} = \frac{e}{4\pi\epsilon_0\, c\, r'^3}\, U^\mu R^\nu. \tag{2.201}$$

Recall that $F^{4i} = -E^i$ and $F^{ij} = -cB^{ij}$. This equation cannot be correct, because the antisymmetric tensor $F^{\mu\nu}$ can only be equated to another antisymmetric tensor. Consequently, we try the equation

$$F^{\mu\nu} = \frac{e}{4\pi\epsilon_0\, c\, r'^3}\,(U^\mu R^\nu - U^\nu R^\mu). \tag{2.202}$$

This is found to give the correct field at P in S' as long as R^μ refers to any event at the charge, no matter which. It only remains to interpret (2.202) in S. It is

convenient to choose for R^μ that event at the charge at which $t = 0$ (not the retarded event). Thus,

$$F^{jk} = -cB^{jk} = \frac{e}{4\pi\epsilon_0\, c\, r'^3}\,\gamma(u)\,(u^j r^k - u^k r^j), \tag{2.203}$$

giving

$$B_i = \frac{1}{2}\,\epsilon_{ijk} B^{jk} = -\frac{\mu_0\, e}{4\pi\, r'^3}\,\gamma(u)\,\epsilon_{ijk} u^j r^k, \tag{2.204}$$

or

$$\boldsymbol{B} = -\frac{\mu_0\, e\, \gamma}{4\pi\, r'^3}\,\boldsymbol{u}\wedge\boldsymbol{r}. \tag{2.205}$$

Likewise,

$$F^{4i} = -E^i = \frac{e\,\gamma}{4\pi\epsilon_0\, r'^3}\, r^i, \tag{2.206}$$

or

$$\boldsymbol{E} = -\frac{e\,\gamma}{4\pi\epsilon_0\, r'^3}\;\boldsymbol{r}. \tag{2.207}$$

Lastly, we must find an expression for r'^3 in terms of quantities measured in S at time $t = 0$. If t' is the corresponding time in S' at the charge, we have

$$r'^2 = r^2 + c^2 t'^2 = r^2 + \frac{\gamma^2\, u^2 x^2}{c^2} = r^2\left(1 + \frac{\gamma^2 u_r{}^2}{c^2}\right). \tag{2.208}$$

Thus,

$$\boldsymbol{E} \;=\; -\frac{e}{4\pi\epsilon_0}\frac{\gamma}{r^3(1 + u_r{}^2\,\gamma^2/c^2)^{3/2}}\,\boldsymbol{r}, \tag{2.209a}$$

$$\boldsymbol{B} \;=\; -\frac{\mu_0\, e}{4\pi}\,\frac{\gamma}{r^3(1 + u_r{}^2\,\gamma^2/c^2)^{3/2}}\,\boldsymbol{u}\wedge\boldsymbol{r} = \frac{1}{c^2}\,\boldsymbol{u}\wedge\boldsymbol{E}. \tag{2.209b}$$

Note that $\boldsymbol{E}$ acts in line with the point which the charge occupies *at the instant of measurement* despite the fact that, owing to the finite speed of propagation of all physical effects, the behaviour of the charge during a finite period before that instant can no longer affect the measurement. Note also that, unlike Eqs. (2.198), the above expressions for the fields are not valid for an arbitrarily moving charge,

not can they be made valid by merely using retarded values. For whereas acceleration does not affect the potentials, it does affect the fields, which involve the derivatives of the potential.

For low velocities, $u/c \to 0$, Eqs. (2.209) reduce to the well known Coulomb and Biot-Savart fields. However, at high velocities, $\gamma(u) \gg 1$, the fields exhibit some interesting behaviour. The peak electric field, which occurs at the point of closest approach of the charge to the observation point, becomes equal to γ times its non-relativistic value. However, the duration of appreciable field strength at the point P is decreased. A measure of the time interval over which the field is appreciable is

$$\Delta t \sim \frac{b}{\gamma c}, \tag{2.210}$$

where b is the distance of closest approach (assuming $\gamma \gg 1$). As γ increases, the peak field increases in proportion, but its duration goes in the inverse proportion. The time integral of the field is independent of γ. As $\gamma \to \infty$ the observer at P sees electric and magnetic fields which are indistinguishable from the fields of a pulse of plane polarized radiation propagating in the x-direction. The direction of polarization is along the radius vector pointing towards the particle's actual position at the time of observation.

2.21 Relativistic particle dynamics

Consider a particle which, in its instantaneous rest frame S_0, has mass m_0 and constant acceleration in the x-direction a_0. Let us transform to a frame S, in the standard configuration with respect to S_0, in which the particle's instantaneous velocity is u. What is the value of a, the particle's instantaneous x-acceleration, in S?

The easiest way in which to answer this question is to consider the acceleration 4-vector (see Eq. (2.85))

$$A^\mu = \gamma \left(\frac{d\gamma}{dt}\, \boldsymbol{u} + \gamma \boldsymbol{a}, c\, \frac{d\gamma}{dt} \right). \tag{2.211}$$

Using the standard transformation (2.57) for 4-vectors, we obtain

$$\frac{d\gamma}{dt}\,u + \gamma a = a_0, \tag{2.212a}$$

$$\frac{d\gamma}{dt} = \frac{u\,a_0}{c^2}. \tag{2.212b}$$

It follows that

$$a = \frac{a_0}{\gamma^3}. \tag{2.213}$$

The above equation can be written

$$f = m_0\gamma^3\,\frac{du}{dt}, \tag{2.214}$$

where $f = m_0 a_0$ is the constant force (in the x-direction) acting on the particle in S_0.

Equation (2.214) is equivalent to

$$f = \frac{d(mu)}{dt}, \tag{2.215}$$

where

$$m = \gamma m_0. \tag{2.216}$$

Thus, we can account for the ever decreasing acceleration of a particle subject to a constant force (see Eq. (2.213)) by supposing that the inertial mass of the particle increases with its velocity according to the rule (2.216). Henceforth, m_0 is termed the *rest mass*, and m the *inertial mass*.

The rate of increase of the particle's energy E satisfies

$$\frac{dE}{dt} = f u = m_0\gamma^3\,u\,\frac{du}{dt}. \tag{2.217}$$

This equation can be written

$$\frac{dE}{dt} = \frac{d(mc^2)}{dt}, \tag{2.218}$$

which can be integrated to yield Einstein's famous formula

$$E = mc^2. \tag{2.219}$$

The 3-momentum of a particle is defined

$$\boldsymbol{p} = m\boldsymbol{u}, \tag{2.220}$$

where $\boldsymbol{u}$ is its 3-velocity. Thus, by analogy with Eq. (2.215), Newton's law of motion can be written

$$\boldsymbol{f} = \frac{d\boldsymbol{p}}{dt}, \tag{2.221}$$

where $\boldsymbol{f}$ is the 3-force acting on the particle.

The 4-momentum of a particle is defined

$$P^\mu = m_0 U^\mu = \gamma m_0(\boldsymbol{u}, c) = (\boldsymbol{p}, E/c), \tag{2.222}$$

where U^μ is its 4-velocity. The 4-force acting on the particle obeys

$$\mathcal{F}^\mu = \frac{dP^\mu}{d\tau} = m_0 A^\mu, \tag{2.223}$$

where A^μ is its 4-acceleration. It is easily demonstrated that

$$\mathcal{F}^\mu = \gamma\left(\boldsymbol{f}, c\,\frac{dm}{dt}\right) = \gamma\left(\boldsymbol{f}, \frac{\boldsymbol{f}\cdot\boldsymbol{u}}{c}\right), \tag{2.224}$$

since

$$\frac{dE}{dt} = \boldsymbol{f}\cdot\boldsymbol{u}. \tag{2.225}$$

2.22 The force on a moving charge

The electromagnetic 3-force acting on a charge e moving with 3-velocity $\boldsymbol{u}$ is given by the well known formula

$$\boldsymbol{f} = e(\boldsymbol{E} + \boldsymbol{u}\wedge\boldsymbol{B}). \tag{2.226}$$

When written in component form this expression becomes

$$f_i = e(E_i + \epsilon_{ijk}\, u^j\, B^k), \tag{2.227}$$

or

$$f_i = e(E_i + B_{ij}\, u^j), \tag{2.228}$$

where use has been made of Eq. (2.155).

Recall that the components of the $\boldsymbol{E}$ and $\boldsymbol{B}$ fields can be written in terms of an antisymmetric electromagnetic field tensor $F_{\mu\nu}$ via

$$F_{i4} = -F_{4i} = -E_i, \tag{2.229a}$$

$$F_{ij} = -F_{ji} = -c B_{ij}. \tag{2.229b}$$

Equation (2.228) can be written

$$f_i = -\frac{e}{\gamma\, c}\,(F_{i4}\, U^4 + F_{ij}\, U^j), \tag{2.230}$$

where $U^\mu = \gamma(\boldsymbol{u}, c)$ is the particle's 4-velocity. It is easily demonstrated that

$$\frac{\boldsymbol{f}\cdot\boldsymbol{u}}{c} = \frac{e}{c}\,\boldsymbol{E}\cdot\boldsymbol{u} = \frac{e}{c}\,E_i\,u^i = \frac{e}{\gamma\, c}(F_{4i}\, U^i + F_{44}\, U^4). \tag{2.231}$$

Thus, the 4-force acting on the particle,

$$\mathcal{F}_\mu = \gamma\left(-\boldsymbol{f}, \frac{\boldsymbol{f}\cdot\boldsymbol{u}}{c}\right), \tag{2.232}$$

can be written in the form

$$\mathcal{F}_\mu = \frac{e}{c}\,F_{\mu\nu}\,U^\nu. \tag{2.233}$$

The skew symmetry of the electromagnetic field tensor ensures that

$$\mathcal{F}_\mu U^\mu = \frac{e}{c}\,F_{\mu\nu}\,U^\mu U^\nu = 0. \tag{2.234}$$

This is an important result since it ensures that electromagnetic fields do not change the rest mass of charged particles. In order to appreciate this, let us assume that the rest mass m_0 is not a constant. Since

$$\mathcal{F}_\mu = \frac{d(m_0 U_\mu)}{d\tau} = m_0\, A_\mu + \frac{dm_0}{d\tau}\, U_\mu, \tag{2.235}$$

we can use the standard results $U_\mu U^\mu = c^2$ and $A_\mu U^\mu = 0$ to give

$$\mathcal{F}_\mu U^\mu = c^2\,\frac{dm_0}{d\tau}. \tag{2.236}$$

Thus, if rest mass is to remain an invariant it is imperative that all laws of physics predict 4-forces acting on particles which are orthogonal to the particles' 4-velocities. The laws of electromagnetism pass this test.

2.23 The electromagnetic energy tensor

Consider a continuous volume distribution of charged matter in the presence of an electromagnetic field. Let there be n_0 particles per unit proper volume (unit volume determined in the local rest frame), each carrying a charge e. Consider an inertial frame in which the 3-velocity field of the particles is $\boldsymbol{u}$. The number density of the particles in this frame is $n = \gamma(u)\, n_0$. The charge density and the 3-current due to the particles are $\rho = en$ and $\boldsymbol{j} = en\, \boldsymbol{u}$, respectively. Multiplying Eq. (2.233) by the proper number density of particles n_0, we obtain an expression

$$f_\mu = \frac{1}{c} F_{\mu\nu} J^\nu \tag{2.237}$$

for the 4-force f_μ acting on unit proper volume of the distribution due to the ambient electromagnetic fields. Here, we have made use of the definition $J^\mu = e n_0 U^\mu$. It is easily demonstrated, using some of the results obtained in the previous section, that

$$f^\mu = \left(\rho\boldsymbol{E} + \boldsymbol{j}\wedge\boldsymbol{B}, \frac{\boldsymbol{E}\cdot\boldsymbol{j}}{c}\right). \tag{2.238}$$

The above expression remains valid when there are many charge species (*e.g.*, electrons and ions) possessing different number density and 3-velocity fields. The 4-vector f^μ is usually called the *Lorentz force density*.

We know that Maxwell's equations reduce to

$$\partial_\mu F^{\mu\nu} = \frac{J^\nu}{c\,\epsilon_0}, \tag{2.239a}$$

$$\partial_\mu G^{\mu\nu} = 0, \tag{2.239b}$$

where $F^{\mu\nu}$ is the electromagnetic field tensor and $G^{\mu\nu}$ is its dual. As is easily verified, Eq. (2.239b) can also be written in the form

$$\partial_\mu F_{\nu\sigma} + \partial_\nu F_{\sigma\mu} + \partial_\sigma F_{\mu\nu} = 0. \tag{2.240}$$

Equations (2.237) and (2.239a) can be combined to give

$$f_\nu = \epsilon_0\, F_{\nu\sigma}\,\partial_\mu F^{\mu\sigma}. \tag{2.241}$$

This expression can also be written

$$f_\nu = \epsilon_0(\partial_\mu(F^{\mu\sigma} F_{\nu\sigma}) - F^{\mu\sigma}\,\partial_\mu F_{\nu\sigma}). \tag{2.242}$$

Now,

$$F^{\mu\sigma}\,\partial_\mu F_{\nu\sigma} = \frac{1}{2} F^{\mu\sigma}(\partial_\mu F_{\nu\sigma} + \partial_\sigma F_{\mu\nu}), \tag{2.243}$$

where use has been made of the antisymmetry of the electromagnetic field tensor. It follows from Eq. (2.240) that

$$F^{\mu\sigma}\,\partial_\mu F_{\nu\sigma} = -\frac{1}{2} F^{\mu\sigma}\,\partial_\nu F_{\sigma\mu} = \frac{1}{4}\partial_\nu(F^{\mu\sigma} F_{\mu\sigma}). \tag{2.244}$$

Thus,

$$f_\nu = \epsilon_0 \left(\partial_\mu(F^{\mu\sigma} F_{\nu\sigma}) - \frac{1}{4}\partial_\nu(F^{\mu\sigma} F_{\mu\sigma})\right). \tag{2.245}$$

The above expression can also be written

$$f_\nu = -\partial_\mu T^\mu{}_\nu, \tag{2.246}$$

where

$$T^\mu{}_\nu = \epsilon_0 \left(F^{\mu\sigma} F_{\sigma\nu} + \frac{1}{4}\delta^\mu_\nu (F^{\rho\sigma} F_{\rho\sigma})\right) \tag{2.247}$$

is called the *electromagnetic energy tensor*. Note that $T^{\mu}{}_{\nu}$ is a proper-4-tensor. It follows from Eqs. (2.167), (2.169), and (2.193) that

$$T^{i}{}_{j} = \epsilon_0 E^i E_j + \frac{B^i B_j}{\mu_0} - \delta^i_j \frac{1}{2}\left(\epsilon_0 E^k E_k + \frac{B^k B_k}{\mu_0}\right), \tag{2.248a}$$

$$T^{i}{}_{4} = -T^{4}{}_{i} = \frac{\epsilon^{ijk} E_j B_k}{\mu_0\, c}, \tag{2.248b}$$

$$T^4_4 = \frac{1}{2}\left(\epsilon_0 E^k E_k + \frac{B^k B_k}{\mu_0}\right). \tag{2.248c}$$

Equation (2.246) can also be written

$$f^{\nu} = -\partial_{\mu} T^{\mu\nu}, \tag{2.249}$$

where $T^{\mu\nu}$ is a symmetric tensor whose elements are

$$T^{ij} = -\epsilon_0 E^i E^j - \frac{B^i B^j}{\mu_0} + \delta^{ij} \frac{1}{2}\left(\epsilon_0 E^2 + \frac{B^2}{\mu_0}\right), \tag{2.250a}$$

$$T^{i4} = T^{4i} = \frac{(\boldsymbol{E}\wedge\boldsymbol{B})^i}{\mu_0\, c}, \tag{2.250b}$$

$$T^{44} = \frac{1}{2}\left(\epsilon_0 E^2 + \frac{B^2}{\mu_0}\right). \tag{2.250c}$$

Consider the time-like component of Eq. (2.249). It follows from Eq. (2.238) that

$$\frac{\boldsymbol{E}\cdot\boldsymbol{j}}{c} = -\partial_i T^{i4} - \partial_4 T^{44}. \tag{2.251}$$

This equation can be rearranged to give

$$\frac{\partial W}{\partial t} + \nabla\cdot\boldsymbol{\epsilon} = -\boldsymbol{E}\cdot\boldsymbol{j}, \tag{2.252}$$

where $W = T^{44}$ and $\epsilon^i = c T^{i4}$, so that

$$W = \frac{\epsilon_0 E^2}{2} + \frac{B^2}{2\mu_0}, \tag{2.253}$$

and

$$\boldsymbol{\epsilon} = \frac{\boldsymbol{E} \wedge \boldsymbol{B}}{\mu_0}. \tag{2.254}$$

The right-hand side of Eq. (2.252) represents the rate per unit volume at which energy is transferred from the electromagnetic field to charged particles. It is clear, therefore, that Eq. (2.252) is an energy conservation equation for the electromagnetic field. The proper-3-scalar W can be identified as the energy density of the electromagnetic field, whereas the proper-3-vector $\boldsymbol{\epsilon}$ is the energy flux due to the electromagnetic field. The latter quantity is called the *Poynting vector.*

Consider the space-like components of Eq. (2.249). It is easily demonstrated that these reduce to

$$\frac{\partial \boldsymbol{g}}{\partial t} + \nabla \cdot \boldsymbol{P} = -\rho \boldsymbol{E} - \boldsymbol{j} \wedge \boldsymbol{B}, \tag{2.255}$$

where $P^{ij} = T^{ij}$ and $g^i = T^{4i}/c$, or

$$P^{ij} = -\epsilon_0 E^i E^j - \frac{B^i B^j}{\mu_0} + \delta^{ij} \, \frac{1}{2}\left(\epsilon_0 E^2 + \frac{B^2}{\mu_0}\right), \tag{2.256}$$

and

$$\boldsymbol{g} = \frac{\boldsymbol{\epsilon}}{c^2} = \epsilon_0 \boldsymbol{E} \wedge \boldsymbol{B}. \tag{2.257}$$

Equation (2.255) is basically a momentum conservation equation for the electromagnetic field. The right-hand side represents the rate per unit volume at which momentum is transferred from the electromagnetic field to charged particles. The symmetric proper-3-tensor P^{ij} is called the *Maxwell stress tensor*. The element P^{ij} gives the flux of electromagnetic momentum parallel to the ith axis crossing a surface normal to the jth axis. The proper-3-vector $\boldsymbol{g}$ represents the momentum density of the electromagnetic field. It is clear that the energy conservation law (2.252) and the momentum conservation law (2.255) can be combined together to give the relativistically invariant energy-momentum conservation law (2.249).

2.24 The electromagnetic field due to an accelerated charge

Let us calculate the electric and magnetic fields observed at position x^i and time t due to a charge e whose *retarded* position and time are $x^{i'}$ and t', respectively.

From now on (x^i, t) is termed the *field point* and $(x^{i'}, t')$ is termed the *source point*. It is assumed that we are given the retarded position of the charge as a function of its retarded time; *i.e.*, $x^{i'}(t')$. The retarded velocity and acceleration of the charge are

$$u^i = \frac{dx^{i'}}{dt'}, \tag{2.258}$$

and

$$\dot{u}^i = \frac{du^i}{dt'}, \tag{2.259}$$

respectively. The radius vector $\boldsymbol{r}$ is defined to extend *from* the retarded position of the charge *to* the field point, so that $r^i = x^i - x^{i'}$. (Note that this is the *opposite* convention to that adopted in Sections 2.19 and 2.20). It follows that

$$\frac{d\boldsymbol{r}}{dt'} = -\boldsymbol{u}. \tag{2.260}$$

The field and the source point variables are connected by the retardation condition

$$r(x^i, x^{i'}) = \left[(x^i - x^{i'})(x_i - x_i{}')\right]^{1/2} = c(t - t'). \tag{2.261}$$

The potentials generated by the charge are given by the Liénard-Wiechert formulae

$$\boldsymbol{A}(x^i, t) = \frac{\mu_0\, e}{4\pi}\frac{\boldsymbol{u}}{s}, \tag{2.262a}$$

$$\phi(x^i, t) = \frac{e}{4\pi\epsilon_0}\frac{1}{s}, \tag{2.262b}$$

where $s = r - \boldsymbol{r}\cdot\boldsymbol{u}/c$ is a function both of the field point and the source point variables. Recall that the Liénard-Wiechert potentials are valid for accelerating as well as uniformly moving charges.

The fields $\boldsymbol{E}$ and $\boldsymbol{B}$ are derived from the potentials in the usual manner

$$\boldsymbol{E} = -\nabla\phi - \frac{\partial \boldsymbol{A}}{\partial t}, \tag{2.263a}$$

$$\boldsymbol{B} = \nabla\wedge\boldsymbol{A}. \tag{2.263b}$$

However, the components of the gradient operator ∇ are partial derivatives at constant time t, and *not* at constant time t'. Partial differentiation with respect to the x^i compares the potentials at neighbouring points at the same time, but these potential signals originate from the charge at different retarded times. Similarly, the partial derivative with respect to t implies constant x^i, and, hence, refers to the comparison of the potentials at a given field point over an interval of time during which the retarded coordinates of the source have changed. Since we only know the time variation of the particle's retarded position with respect to t' we must transform $\partial/\partial t|_{x^i}$ and $\partial/\partial x^i|_t$ to expressions involving $\partial/\partial t'|_{x^i}$ and $\partial/\partial x^i|_{t'}$.

Now, since $x^{i'}$ is assumed to be given as a function of t', we have

$$r(x^i, x^{i'}(t')\,) \equiv r(x^i, t') = c(t - t'), \tag{2.264}$$

which is a functional relationship between x^i, t, and t'. Note that

$$\left(\frac{\partial r}{\partial t'}\right)_{x^i} = -\frac{\boldsymbol{r}\cdot\boldsymbol{u}}{r}. \tag{2.265}$$

It follows that

$$\frac{\partial r}{\partial t} = c\left(1 - \frac{\partial t'}{\partial t}\right) = \frac{\partial r}{\partial t'}\frac{\partial t'}{\partial t} = -\frac{\boldsymbol{r}\cdot\boldsymbol{u}}{r}\,\frac{\partial t'}{\partial t}, \tag{2.266}$$

where all differentiation is at constant x^i. Thus,

$$\frac{\partial t'}{\partial t} = \frac{1}{1 - \boldsymbol{r}\cdot\boldsymbol{u}/rc} = \frac{r}{s}, \tag{2.267}$$

giving

$$\frac{\partial}{\partial t} = \frac{r}{s}\frac{\partial}{\partial t'}. \tag{2.268}$$

Similarly,

$$\nabla r = -c\nabla t' = \nabla' r + \frac{\partial r}{\partial t'}\,\nabla t' = \frac{\boldsymbol{r}}{r} - \frac{\boldsymbol{r}\cdot\boldsymbol{u}}{r}\,\nabla t', \tag{2.269}$$

where ∇' denotes differentiation with respect to x^i at constant t'. It follows that

$$\nabla t' = -\frac{\boldsymbol{r}}{sc}, \tag{2.270}$$

so that

$$\nabla = \nabla' - \frac{\boldsymbol{r}}{s c}\frac{\partial}{\partial t'}. \tag{2.271}$$

Equation (2.263a) yields

$$\frac{4\pi\epsilon_0}{e}\,\boldsymbol{E} = \frac{\nabla s}{s^2} - \frac{\partial}{\partial t}\frac{\boldsymbol{u}}{s c^2}, \tag{2.272}$$

or

$$\frac{4\pi\epsilon_0}{e}\,\boldsymbol{E} = \frac{\nabla' s}{s^2} - \frac{\boldsymbol{r}}{s^3 c}\frac{\partial s}{\partial t'} - \frac{r}{s^2 c^2}\,\dot{\boldsymbol{u}} + \frac{r\,\boldsymbol{u}}{s^3 c^2}\frac{\partial s}{\partial t'}. \tag{2.273}$$

However,

$$\nabla' s = \frac{\boldsymbol{r}}{r} - \frac{\boldsymbol{u}}{c}, \tag{2.274}$$

and

$$\frac{\partial s}{\partial t'} = \frac{\partial r}{\partial t'} - \frac{\boldsymbol{r}\cdot\dot{\boldsymbol{u}}}{c} + \frac{\boldsymbol{u}\cdot\boldsymbol{u}}{c} = -\frac{\boldsymbol{r}\cdot\boldsymbol{u}}{r} - \frac{\boldsymbol{r}\cdot\dot{\boldsymbol{u}}}{c} + \frac{u^2}{c}. \tag{2.275}$$

Thus,

$$\frac{4\pi\epsilon_0}{e}\,\boldsymbol{E} = \frac{1}{s^2 r}\left(\boldsymbol{r} - \frac{r\boldsymbol{u}}{c}\right) + \frac{1}{s^3 c}\left(\boldsymbol{r} - \frac{r\,\boldsymbol{u}}{c}\right)\left(\frac{\boldsymbol{r}\cdot\boldsymbol{u}}{r} - \frac{u^2}{c} + \frac{\boldsymbol{r}\cdot\dot{\boldsymbol{u}}}{c}\right) - \frac{r}{s^2 c^2}\,\dot{\boldsymbol{u}}, \tag{2.276}$$

which reduces to

$$\frac{4\pi\epsilon_0}{e}\,\boldsymbol{E} = \frac{1}{s^3}\left(\boldsymbol{r} - \frac{r\,\boldsymbol{u}}{c}\right)\left(1 - \frac{u^2}{c^2}\right) + \frac{1}{s^3 c^2}\left(\boldsymbol{r}\wedge\left[\left(\boldsymbol{r} - \frac{r\,\boldsymbol{u}}{c}\right)\wedge\dot{\boldsymbol{u}}\right]\right). \tag{2.277}$$

Similarly,

$$\frac{4\pi}{\mu_0\, e}\,\boldsymbol{B} = \nabla\wedge\frac{\boldsymbol{u}}{s} = -\frac{\nabla' s\wedge\boldsymbol{u}}{s^2} - \frac{\boldsymbol{r}}{s c}\wedge\left(\frac{\dot{\boldsymbol{u}}}{s} - \frac{u}{s^2}\frac{\partial s}{\partial t'}\right), \tag{2.278}$$

or

$$\frac{4\pi}{\mu_0\, e}\,\boldsymbol{B} = -\frac{\boldsymbol{r}\wedge\boldsymbol{u}}{s^2 r} - \frac{\boldsymbol{r}}{s c}\wedge\left[\frac{\dot{\boldsymbol{u}}}{s} + \frac{\boldsymbol{u}}{s^2}\left(\frac{\boldsymbol{r}\cdot\boldsymbol{u}}{r} + \frac{\boldsymbol{r}\cdot\dot{\boldsymbol{u}}}{c} - \frac{u^2}{c}\right)\right], \tag{2.279}$$

which reduces to

$$\frac{4\pi}{\mu_0\, e}\,\boldsymbol{B} = \frac{\boldsymbol{u}\wedge\boldsymbol{r}}{s^3}\left(1 - \frac{u^2}{c^2}\right) + \frac{1}{s^3 c}\frac{\boldsymbol{r}}{r}\wedge\left(\boldsymbol{r}\wedge\left[\left(\boldsymbol{r} - \frac{r\,\boldsymbol{u}}{c}\right)\wedge\dot{\boldsymbol{u}}\right]\right). \tag{2.280}$$

A comparison of Eqs. (2.277) and (2.280) yields

$$\boldsymbol{B} = \frac{\boldsymbol{r}\wedge\boldsymbol{E}}{r c}. \tag{2.281}$$

Thus, the magnetic field is always perpendicular to $\boldsymbol{E}$ and the *retarded* radius vector $\boldsymbol{r}$. Note that all terms appearing in the above formulae are retarded.

The electric field is composed of two separate parts. The first term in Eq. (2.277) varies as $1/r^2$ for large distances from the charge. We can think of $\boldsymbol{r}_u = \boldsymbol{r} - r\,\boldsymbol{u}/c$ as the *virtual present radius vector*; *i.e.*, the radius vector directed from the position the charge would occupy at time t if it had continued with uniform velocity from its retarded position to the field point. In terms of $\boldsymbol{r}_u$ the $1/r^2$ field is simply

$$\boldsymbol{E}_{\rm induction} = \frac{e}{4\pi\epsilon_0}\frac{1-u^2/c^2}{s^3}\,\boldsymbol{r}_u. \tag{2.282}$$

We can rewrite the expression (2.209a) for the electric field generated by a *uniformly* moving charge in the form

$$\boldsymbol{E} = \frac{e}{4\pi\epsilon_0}\frac{1-u^2/c^2}{{r_0}^3(1-u^2/c^2+{u_r}^2/c^2)^{3/2}}\,\boldsymbol{r}_0, \tag{2.283}$$

where $\boldsymbol{r}_0$ is the radius vector directed from the *present* position of the charge at time t to the field point, and $u_r = \boldsymbol{u}\cdot\boldsymbol{r}_0/r_0$. For the case of uniform motion the relationship between the retarded radius vector $\boldsymbol{r}$ and the actual radius vector $\boldsymbol{r}_0$ is simply

$$\boldsymbol{r}_0 = \boldsymbol{r} - \frac{r}{c}\,\boldsymbol{u}. \tag{2.284}$$

It is straightforward to demonstrate that

$$s = r_0\sqrt{1-u^2/c^2+{u_r}^2/c^2} \tag{2.285}$$

in this case. Thus, the electric field generated by a uniformly moving charge can be written

$$\boldsymbol{E} = \frac{e}{4\pi\epsilon_0}\frac{1-u^2/c^2}{s^3}\,\boldsymbol{r}_0. \tag{2.286}$$

Since $\boldsymbol{r}_u = \boldsymbol{r}_0$ for the case of a uniformly moving charge, it is clear that Eq. (2.282) is equivalent to the electric field generated by a uniformly moving charge located

at the position the charge would occupy if it had continued with uniform velocity from its retarded position.

The second term in Eq. (2.277),

$$\boldsymbol{E}_{\rm radiation} = \frac{e}{4\pi\epsilon_0 c^2}\frac{\boldsymbol{r}\wedge(\boldsymbol{r}_u\wedge\dot{\boldsymbol{u}})}{s^3}, \tag{2.287}$$

is of order $1/r$ and, therefore, represents a radiation field in the sense of contributing to the energy flux over a large sphere. Similar considerations hold for the two terms of Eq. (2.280).

2.25 The Lamor formula

Let us transform to the inertial frame in which the charge is instantaneously at rest at the origin at time $t = 0$. In this frame $u \ll c$, so that $\boldsymbol{r}_u \simeq \boldsymbol{r}$ and $s \simeq r$, for events which are sufficiently close to the origin at $t = 0$ that the retarded charge still appears to travel with a velocity which is small compared to that of light. It follows from the previous section that

$$\boldsymbol{E}_{\rm rad} \simeq \frac{e}{4\pi\epsilon_0 c^2}\frac{\boldsymbol{r}\wedge(\boldsymbol{r}\wedge\dot{\boldsymbol{u}})}{r^3}, \tag{2.288a}$$

$$\boldsymbol{B}_{\rm rad} \simeq \frac{e}{4\pi\epsilon_0 c^3}\frac{\dot{\boldsymbol{u}}\wedge\boldsymbol{r}}{r^2}. \tag{2.288b}$$

Let us define spherical polar coordinates whose axis points along the direction of instantaneous acceleration of the charge. It is easily demonstrated that

$$E_\theta \simeq \frac{e}{4\pi\epsilon_0 c^2}\frac{\sin\theta}{r}\,\dot{u}, \tag{2.289a}$$

$$B_\phi \simeq \frac{e}{4\pi\epsilon_0 c^3}\frac{\sin\theta}{r}\,\dot{u}. \tag{2.289b}$$

These fields are identical to those of a radiating dipole whose axis is aligned along the direction of instantaneous acceleration. The Poynting flux is given by

$$\epsilon_r = \frac{E_\theta B_\phi}{\mu_0} = \frac{e^2}{16\pi^2\epsilon_0 c^3}\frac{\sin^2\theta}{r^2}\,\dot{u}^2. \tag{2.290}$$

We can integrate this expression to obtain the instantaneous power radiated by the charge

$$P = \frac{e^2}{6\pi\epsilon_0 c^3}\,\dot{u}^2. \tag{2.291}$$

This is known as *Lamor's formula.* Note that zero net momentum is carried off by the fields (2.289).

In order to proceed further it is necessary to prove two useful theorems. The first theorem states that if a 4-vector field T^μ satisfies

$$\partial_\mu T^\mu = 0, \tag{2.292}$$

and if the components of T^μ are non-zero only in a finite spatial region, then the integral over 3-space,

$$I = \int T^4\, d^3x, \tag{2.293}$$

is an invariant. In order to prove this theorem we need to use the 4-dimensional analog of Gauss's theorem, which states that

$$\int_V \partial_\mu T^\mu\, d^4x = \oint_S T^\mu\, dS_\mu, \tag{2.294}$$

where dS_μ is an element of the 3-dimensional surface S bounding the 4-dimensional volume V. The particular volume over which the integration is performed is indicated in Fig. 1. The surfaces A and C are chosen so that the spatial components of T^μ vanish on A and C. This is always possible because it is assumed that the region over which the components of T^μ are non-zero is of finite extent. The surface B is chosen normal to the x^4-axis whereas the surface D is chosen normal to the $x^{4'}$-axis. Here, the x^μ and the $x^{\mu'}$ are coordinates in two arbitrarily chosen inertial frames. It follows from Eq. (2.294) that

$$\int T^4\, dS_4 + \int T^{4'}\, dS_{4'} = 0. \tag{2.295}$$

Here, we have made use of the fact that $T^\mu\, dS_\mu$ is a scalar and, therefore, has the same value in all inertial frames. Since $dS_4 = -d^3x$ and $dS_{4'} = d^3x'$ it follows that $I = \int T^4\, d^3x$ is an invariant under a Lorentz transformation. Incidentally,

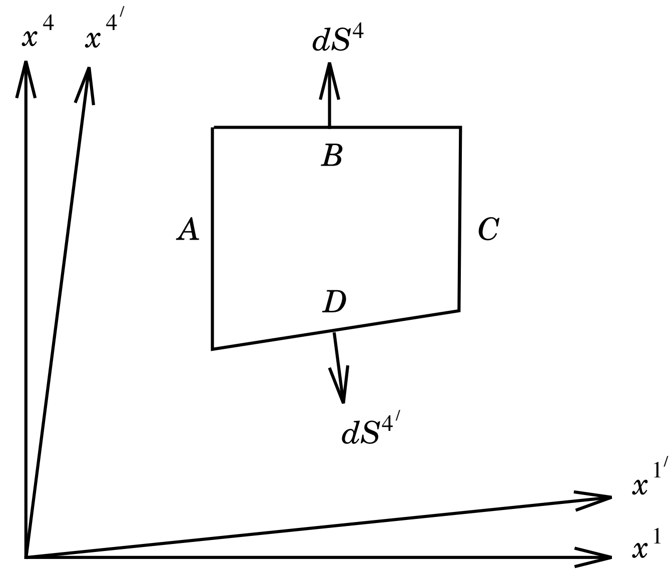

Figure 1: *The region of integration for proving the theorem associated with Eq. (2.293)*

the above argument also demonstrates that I is constant in time (just take the limit in which the two inertial frames are identical).

The second theorem is an extension of the first. Suppose that a 4-tensor field $Q^{\mu\nu}$ satisfies

$$\partial_\mu Q^{\mu\nu} = 0, \tag{2.296}$$

and has components which are only non-zero in a finite spatial region. Let A_μ be a 4-vector whose coefficients do not vary with position in space-time. It follows that $T^\mu = A_\nu Q^{\mu\nu}$ satisfies Eq. (2.292). Therefore,

$$I = \int A_\nu Q^{4\nu}\, d^3x \tag{2.297}$$

is an invariant. However, we can write

$$I = A_\mu B^\mu, \tag{2.298}$$

where

$$B^\mu = \int Q^{4\mu}\, d^3x. \tag{2.299}$$

It follows from the quotient rule that if $A_\mu B^\mu$ is an invariant for arbitrary A_μ then B^μ must transform as a constant (in time) 4-vector.

These two theorems enable us to convert differential conservation laws into integral conservation laws. For instance, in differential form the conservation of electrical charge is written

$$\partial_\mu J^\mu = 0. \tag{2.300}$$

However, from Eq. (2.293) this immediately implies that

$$Q = \frac{1}{c}\int J^4\, d^3x = \int \rho\, d^3x \tag{2.301}$$

is an invariant. In other words, the total electrical charge contained in space is both constant in time and the same in all inertial frames.

Suppose that S is the instantaneous rest frame of the charge. Let us consider the electromagnetic energy tensor $T^{\mu\nu}$ associated with all of the radiation emitted by the charge between times $t = 0$ and $t = dt$. According to Eq. (2.249) this tensor field satisfies

$$\partial_\mu T^{\mu\nu} = 0, \tag{2.302}$$

apart from a region of space of measure zero in the vicinity of the charge. Furthermore, the region of space over which $T^{\mu\nu}$ is non-zero is clearly finite, since we are only considering the fields emitted by the charge in a small time interval, and these fields propagate at a finite velocity. Thus, according to the second theorem

$$P^\mu = \frac{1}{c}\int T^{4\mu}\, d^3x \tag{2.303}$$

is a 4-vector. It follows from Section 2.23 that we can write $P^\mu = (d\boldsymbol{p}, dE/c)$, where $d\boldsymbol{p}$ and dE are the total momentum and energy carried off by the radiation emitted between times $t = 0$ and $t = dt$, respectively. As we have already mentioned, $d\boldsymbol{p} = 0$ in the instantaneous rest frame S. Transforming to an arbitrary inertial frame S' in which the instantaneous velocity of the charge is u, we obtain

$$dE^{'} = \gamma(u)\left(dE + u\, dp^1\right) = \gamma\, dE. \tag{2.304}$$

However, the time interval over which the radiation is emitted in S' is $dt' = \gamma dt$. Thus, the instantaneous power radiated by the charge,

$$P' = \frac{dE'}{dt'} = \frac{dE}{dt} = P, \tag{2.305}$$

is the same in all inertial frames.

We can make use of the fact that the power radiated by an accelerating charge is Lorentz invariant to find a relativistic generalization of the Lamor formula (2.291) which is valid in all inertial frames. We expect the power emitted by the charge to depend only on its 4-velocity and 4-acceleration. It follows that the Lamor formula can be written in Lorentz invariant form as

$$P = -\frac{e^2}{6\pi\epsilon_0 c^3}\, A_\mu A^\mu, \tag{2.306}$$

since the 4-acceleration takes the form $A^\mu = (\dot{\boldsymbol{u}}, 0)$ in the instantaneous rest frame. In a general inertial frame

$$-A_\mu A^\mu = \gamma^2 \left(\frac{d\gamma}{dt}\,\boldsymbol{u} + \gamma\,\dot{\boldsymbol{u}}\right)^2 - \gamma^2 c^2 \left(\frac{d\gamma}{dt}\right)^2, \tag{2.307}$$

where use has been made of Eq. (2.85). Furthermore, it is easily demonstrated that

$$\frac{d\gamma}{dt} = \gamma^3\,\frac{\boldsymbol{u}\cdot\dot{\boldsymbol{u}}}{c^2}. \tag{2.308}$$

It follows, after a little algebra, that the relativistic generalization of Lamor's formula takes the form

$$P = \frac{e^2}{6\pi\epsilon_0 c^3}\,\gamma^6 \left[\dot{\boldsymbol{u}}^2 - \frac{(\boldsymbol{u}\wedge\dot{\boldsymbol{u}})^2}{c^2}\right]. \tag{2.309}$$

2.26 Radiation losses in charged particle accelerators

Radiation losses often limit the maximum practical energy attainable in a charged particle accelerator. Let us investigate radiation losses in various different types of accelerator device using the relativistic Lamor formula.

For a linear accelerator the motion is one dimensional. In this case, it is easily demonstrated that

$$\frac{dp}{dt} = m_0\gamma^3\,\dot{u}, \tag{2.310}$$

where use has been made of Eq. (2.308), and $p = \gamma m_0 u$ is the particle momentum in the direction of acceleration (the x-direction, say). Here, m_0 is the particle rest mass. Thus, Eq. (2.309) yields

$$P = \frac{e^2}{6\pi\epsilon_0 {m_0}^2 c^3}\left(\frac{dp}{dt}\right)^2. \tag{2.311}$$

The rate of change of momentum is equal to the force exerted on the particle in the x-direction, which in turn equals the change in the energy, E, of the particle per unit distance. Consequently,

$$P = \frac{e^2}{6\pi\epsilon_0 {m_0}^2 c^3}\left(\frac{dE}{dx}\right)^2. \tag{2.312}$$

Thus, in a linear accelerator the radiated power depends on the external force acting on the particle, and not on the actual energy or momentum of the particle. It is obvious from the above formula that light particles such as electrons are going to radiate a lot more than heavier particles such as protons. The ratio of the power radiated to the power supplied by the external sources is

$$\frac{P}{dE/dt} = \frac{e^2}{6\pi\epsilon_0 {m_0}^2 c^3}\frac{1}{u}\frac{dE}{dx} \simeq \frac{e^2}{6\pi\epsilon_0 m_0 c^2}\frac{1}{m_0 c^2}\frac{dE}{dx}, \tag{2.313}$$

since $u \simeq c$ for a highly relativistic particle. It is clear from the above expression that the radiation losses in an electron linear accelerator are negligible unless the gain in energy is of order $m_e c^2 = 0.511$ MeV in a distance of $e^2/(6\pi\epsilon_0\, m_e c^2) = 1.28 \times 10^{-15}$ meters. That is 3×10^{14} MeV/meter. Typical energy gains are less that 10 MeV/meter. It is, therefore, obvious that radiation losses are completely negligible in linear accelerators, whether for electrons or for other heavier particles.

The situation is quite different in circular accelerator devices such as the synchrotron and the betatron. In such machines the momentum $\boldsymbol{p}$ changes rapidly

in direction as the particle rotates, but the change in energy per revolution is small. Furthermore, the direction of acceleration is always perpendicular to the direction of motion. It follows from Eq. (2.309) that

$$P = \frac{e^2}{6\pi\epsilon_0 c^3}\,\gamma^4\,\dot{u}^2 = \frac{e^2}{6\pi\epsilon_0 c^3}\,\frac{\gamma^4 u^4}{\rho^2}, \tag{2.314}$$

where ρ is the orbit radius. Here, use has been made of the standard result $\dot{u} = u^2/\rho$ for circular motion. The radiative energy loss per revolution is given by

$$\delta E = \frac{2\pi\rho}{u}P = \frac{e^2}{3\epsilon_0 c^3}\,\frac{\gamma^4 u^3}{\rho}. \tag{2.315}$$

For highly relativistic ($u \simeq c$) electrons this expression yields

$$\delta E(\text{MeV}) = 8.85\times 10^{-2}\,\frac{[E(\text{GeV})]^4}{\rho(\text{meters})}. \tag{2.316}$$

In the first electron synchrotrons, $\rho \sim 1$ meter, $E_{\rm max} \sim 0.3$ GeV. Hence, $\delta E_{\rm max} \sim$ 1 keV per revolution. This was less than, but not negligible compared to, the energy gain of a few keV per turn. For modern electron synchrotrons the limitation on the available radio-frequency power needed to overcome radiation losses becomes a major consideration, as is clear from the E^4 dependence of the radiated power per turn.

2.27 The angular distribution of radiation emitted by an accelerated charge

In order to calculate the angular distribution of the energy radiated by an accelerated charge we must think carefully about what is meant by the "rate of radiation" of the charge. This quantity is actually the amount of energy lost by the charge in a retarded time interval dt' during the emission of the signal. Thus,

$$P(t') = -\frac{dE}{dt'}, \tag{2.317}$$

where E is the energy of the charge. The Poynting vector

$$\boldsymbol{\epsilon} = \frac{\boldsymbol{E}_{\rm rad} \wedge \boldsymbol{B}_{\rm rad}}{\mu_0} = \epsilon_0 c\, E_{\rm rad}^{\,2}\, \frac{\boldsymbol{r}}{r}, \tag{2.318}$$

where use has been made of $\boldsymbol{B}_{\rm rad} = (\boldsymbol{r} \wedge \boldsymbol{E}_{\rm rad})/rc$ (see Eq. (2.281)), represents the energy flux per unit actual time, t. Thus, the energy loss rate of the charge into a given element of solid angle $d\Omega$ is

$$\frac{dP(t')}{d\Omega}\, d\Omega = -\frac{dE(\theta,\varphi)}{dt'}\, d\Omega = |\boldsymbol{\epsilon}|\, \frac{dt}{dt'}\, r^2\, d\Omega = \epsilon_0 c\, E_{\rm rad}^{\,2}\, \frac{s}{r}\, r^2\, d\Omega, \tag{2.319}$$

where use has been made of Eq. (2.267). Here, θ and φ are angular coordinates used to locate the element of solid angle. It follows from Eq. (2.287) that

$$\frac{dP(t')}{d\Omega} = \frac{e^2 r}{16\pi^2\epsilon_0 c^3} \frac{[\boldsymbol{r} \wedge (\boldsymbol{r}_u \wedge \dot{\boldsymbol{u}})]^2}{s^5}. \tag{2.320}$$

Consider the special case where the direction of acceleration coincides with the direction of motion. Let us define spherical polar coordinates whose axis points along this common direction. It is easily demonstrated that the above expression reduces to

$$\frac{dP(t')}{d\Omega} = \frac{e^2\, \dot{u}^2}{16\pi^2\epsilon_0 c^3} \frac{\sin^2\theta}{[1-(u/c)\,\cos\theta]^5} \tag{2.321}$$

in this case. In the non-relativistic limit $u/c \rightarrow 0$ the radiation pattern has the familiar $\sin^2\theta$ dependence of dipole radiation. In particular, the pattern is symmetric in the forward ($\theta < \pi/2$) and backward ($\theta > \pi/2$) directions. However, as $u/c \rightarrow 1$ the radiation pattern becomes more and more concentrated in the forward direction. The angle $\theta_{\rm max}$ for which the intensity is a maximum is

$$\theta_{\rm max} = \cos^{-1}\left[\frac{1}{3\,u/c}(\sqrt{1+15\,u^2/c^2}-1)\right]. \tag{2.322}$$

This expression yields $\theta_{\rm max} \rightarrow \pi/2$ as $u/c \rightarrow 0$ and $\theta_{\rm max} \rightarrow 1/(2\gamma)$ as $u/c \rightarrow 1$. Thus, for a highly relativistic charge the radiation is emitted in a narrow cone whose axis is aligned along the direction of motion. In this case, the angular distribution (2.321) reduces to

$$\frac{dP(t')}{d\Omega} \simeq \frac{2e^2\, \dot{u}^2}{\pi^2\epsilon_0 c^3}\, \gamma^8 \frac{(\gamma\theta)^2}{[1+(\gamma\theta)^2]^5}. \tag{2.323}$$

The total power radiated by the charge is obtained by integrating Eq. (2.321) over all solid angles. We obtain

$$P(t') = \frac{e^2\,\dot{u}^2}{8\pi\epsilon_0 c^3}\int_0^{\pi}\frac{\sin^3\theta\,d\theta}{[1-(u/c)\,\cos\theta]^5} = \frac{e^2\,\dot{u}^2}{8\pi\epsilon_0 c^3}\int_{-1}^{+1}\frac{(1-\mu^2)\,d\mu}{[1-(u/c)\,\mu]^5}. \tag{2.324}$$

It is easily verified that

$$\int_{-1}^{+1}\frac{(1-\mu^2)\,d\mu}{[1-(u/c)\,\mu]^5} = \frac{4}{3}\,\gamma^6. \tag{2.325}$$

Hence,

$$P(t') = \frac{e^2}{6\pi\epsilon_0 c^3}\,\gamma^6\,\dot{u}^2, \tag{2.326}$$

which agrees with Eq. (2.309) provided that $\boldsymbol{u}\wedge\dot{\boldsymbol{u}} = 0$.

2.28 Synchrotron radiation

Synchrotron radiation (*i.e.*, radiation emitted by a charged particle constrained to follow a circular orbit by a magnetic field) is of particular importance in astrophysics, since much of the observed radio frequency emission from supernova remnants and active galactic nuclei is thought to be of this type.

Consider a charged particle moving in a circle of radius a with constant angular velocity ω_0. Suppose that the orbit lies in the x-y plane. The radius vector pointing from the centre of the orbit to the retarded position of the charge is defined

$$\boldsymbol{\rho} = a\,(\cos\phi, \sin\phi, 0), \tag{2.327}$$

where $\phi = \omega_0\,t'$ is the angle subtended between this vector and the x-axis. The retarded velocity and acceleration of the charge take the form

$$\boldsymbol{u} = \frac{d\boldsymbol{\rho}}{dt'} = u\,(-\sin\phi, \cos\phi, 0), \tag{2.328a}$$

$$\dot{\boldsymbol{u}} = \frac{d\boldsymbol{u}}{dt'} = -\dot{u}\,(\cos\phi, \sin\phi, 0), \tag{2.328b}$$

where $u = a\,\omega_0$ and $\dot{u} = a\,{\omega_0}^2$. The observation point is chosen such that the radius vector $\boldsymbol{r}$, pointing from the retarded position of the charge to the observation point, is parallel to the y-z plane. Thus, we can write

$$\boldsymbol{r} = r\,(0, \sin\alpha, \cos\alpha), \tag{2.329}$$

where α is the angle subtended between this vector and the z-axis. As usual, we define θ as the angle subtended between the retarded radius vector $\boldsymbol{r}$ and the retarded direction of motion of the charge $\boldsymbol{u}$. It follows that

$$\cos\theta = \frac{\boldsymbol{u}\cdot\boldsymbol{r}}{u\,r} = \sin\alpha\,\cos\phi. \tag{2.330}$$

It is easily seen that

$$\dot{\boldsymbol{u}}\cdot\boldsymbol{r} = -\dot{u}r\,\sin\alpha\,\sin\phi. \tag{2.331}$$

A little vector algebra shows that

$$[\boldsymbol{r}\wedge(\boldsymbol{r}_u\wedge\dot{\boldsymbol{u}})]^2 = -(\boldsymbol{r}\cdot\dot{\boldsymbol{u}})^2\,r^2\,(1-u^2/c^2) + \dot{u}^2\,r^4\,(1-\boldsymbol{r}\cdot\boldsymbol{u}/rc)^2, \tag{2.332}$$

giving

$$[\boldsymbol{r}\wedge(\boldsymbol{r}_u\wedge\dot{\boldsymbol{u}})]^2 = \dot{u}^2\,r^4\left[\left(1-\frac{u}{c}\,\cos\theta\right)^2 - \left(1-\frac{u^2}{c^2}\right)\tan^2\phi\cos^2\theta\right]. \tag{2.333}$$

Making use of Eq. (2.320), we obtain

$$\frac{dP(t')}{d\Omega} = \frac{e^2\,\dot{u}^2}{16\pi^2\epsilon_0 c^3}\frac{[1-(u/c)\,\cos\theta)]^2-(1-u^2/c^2)\tan^2\phi\,\cos^2\theta}{[1-(u/c)\,\cos\theta]^5}. \tag{2.334}$$

It is convenient to write this result in terms of the angles α and ϕ, instead of θ and ϕ. After a little algebra we obtain

$$\frac{dP(t')}{d\Omega} = \frac{e^2\,\dot{u}^2}{16\pi^2\epsilon_0 c^3}\frac{[1-(u^2/c^2)]\cos^2\alpha + [(u/c)-\sin\alpha\,\cos\phi]^2}{[1-(u/c)\,\sin\alpha\,\cos\phi]^5}. \tag{2.335}$$

Let us consider the radiation pattern emitted in the plane of the orbit; *i.e.*, $\alpha = \pi/2$, with $\cos\phi = \cos\theta$. It is easily seen that

$$\frac{dP(t')}{d\Omega} = \frac{e^2\,\dot{u}^2}{16\pi^2\epsilon_0 c^3}\frac{[(u/c)-\cos\theta]^2}{[1-(u/c)\,\cos\theta]^5}. \tag{2.336}$$

In the non-relativistic limit the radiation pattern has a $\cos^2\theta$ dependence. Thus, the pattern is like that of dipole radiation where the axis is aligned along the instantaneous direction of acceleration. As the charge becomes more relativistic the radiation lobe in the forward direction (*i.e.*, $0 < \theta < \pi/2$) becomes more more focused and more intense. Likewise, the radiation lobe in the backward direction (*i.e.*, $\pi/2 < \theta < \pi$) becomes more diffuse. The radiation pattern has zero intensity at the angles

$$\theta_0 = \cos^{-1}(u/c). \tag{2.337}$$

These angles demark the boundaries between the two radiation lobes. In the non-relativistic limit $\theta_0 = \pm\pi/2$, so the two lobes are of equal angular extents. In the highly relativistic limit $\theta_0 \to \pm 1/\gamma$, so the forward lobe becomes highly concentrated about the forward direction ($\theta = 0$). In the latter limit Eq. (2.336) reduces to

$$\frac{dP(t')}{d\Omega} \simeq \frac{e^2\,\dot{u}^2}{2\pi^2\epsilon_0 c^3}\,\gamma^6\,\frac{[1-(\gamma\theta)^2]^2}{[1+(\gamma\theta)^2]^5}. \tag{2.338}$$

Thus, the radiation emitted by a highly relativistic charge is focused into an intense beam of angular extent $1/\gamma$ pointing in the instantaneous direction of motion. The maximum intensity of the beam scales like γ^6.

Integration of Eq. (2.335) over all solid angle (using $d\Omega = \sin\alpha\,d\alpha\,d\phi$) yields (not very easily!)

$$P(t') = \frac{e^2}{6\pi\epsilon_0 c^3}\,\gamma^4\,\dot{u}^2, \tag{2.339}$$

which agrees with Eq. (2.309) provided that $\boldsymbol{u}\cdot\dot{\boldsymbol{u}} = 0$. This expression can also be written

$$\frac{P}{m_0 c^2} = \frac{2}{3}\frac{{\omega_0}^2 r_0}{c}\,\beta^2\,\gamma^4, \tag{2.340}$$

where $r_0 = e^2/(4\pi\epsilon_0\,m_0 c^2) = 2.82\times 10^{-15}$ meters is the *classical electron radius*, m_0 is the rest mass of the charge, and $\beta = u/c$. If the circular motion takes place in an orbit of radius a perpendicular to a magnetic field $\boldsymbol{B}$, then ω_0 satisfies $\omega_0 = eB/m_0\gamma$. Thus, the radiated power is

$$\frac{P}{m_0 c^2} = \frac{2}{3}\left(\frac{eB}{m_0}\right)^2\frac{r_0}{c}\,(\beta\gamma)^2, \tag{2.341}$$

and the radiated energy ΔE per revolution is

$$\frac{\Delta E}{m_0 c^2} = \frac{4\pi}{3} \frac{r_0}{a} \beta^3 \gamma^4. \tag{2.342}$$

Let us consider the frequency distribution of the emitted radiation in the highly relativistic limit. Suppose, for the sake of simplicity, that the observation point lies in the plane of the orbit (*i.e.*, $\alpha = \pi/2$). Since the radiation emitted by the charge is beamed very strongly in the charge's instantaneous direction of motion, a fixed observer is only going to see radiation (at some later time) when this direction points almost directly towards the point of observation. This occurs once every rotation period when $\phi \simeq 0$, assuming that $\omega_0 > 0$. Note that the point of observation is located many orbit radii away from the centre of the orbit along the positive y-axis. Thus, our observer sees short periodic pulses of radiation from the charge. The repetition frequency of the pulses (in radians per second) is ω_0. Let us calculate the duration of each pulse. Since the radiation emitted by the charge is focused into a narrow beam of angular extent $\Delta\theta \sim 1/\gamma$, our observer only sees radiation from the charge when $\phi \lesssim \Delta\theta$. Thus, the observed pulse is emitted during a time interval $\Delta t' = \Delta\theta/\omega_0$. However, the pulse is received in a somewhat shorter time interval

$$\Delta t = \frac{\Delta\theta}{\omega_0} \left(1 - \frac{u}{c}\right), \tag{2.343}$$

because the charge is slightly closer to the point of observation at the end of the pulse than at the beginning. The above equation reduces to

$$\Delta t \simeq \frac{\Delta\theta}{2\omega_0 \gamma^2} \sim \frac{1}{\omega_0 \gamma^3}, \tag{2.344}$$

since $\gamma \gg 1$ and $\Delta\theta \sim 1/\gamma$. The width $\Delta\omega$ of the pulse in frequency space obeys $\Delta\omega\, \Delta t \sim 1$. Hence,

$$\Delta\omega = \gamma^3 \omega_0. \tag{2.345}$$

In other words, the emitted frequency spectrum contains harmonics of frequency up to γ^3 times that of the fundamental, ω_0.

More involved calculations[8] show that in the ultra-relativistic limit $\gamma \gg 1$ the power radiated in the lth harmonic (whose frequency is $\omega = l\omega_0$) is given by

$$P_l = 0.52 \left(\frac{e^2}{4\pi\epsilon_0 c}\right) {\omega_0}^2\, l^{1/3} \tag{2.346}$$

for $1 \ll l \ll \gamma^3$, and

$$P_l = \frac{1}{2\sqrt{\pi}} \left(\frac{e^2}{4\pi\epsilon_0 c}\right) {\omega_0}^2 \left(\frac{l}{\gamma}\right)^{1/2} \exp[(-2/3)(l/\gamma^3)] \tag{2.347}$$

for $l \gg \gamma^3$. Note that the spectrum cuts off approximately at the harmonic order γ^3, as predicted earlier. It can also be demonstrated[9] that *seven* times as much energy is radiated with a polarization parallel to the orbital plane than with a perpendicular polarization. A $P(\omega) \propto \omega^{1/3}$ power spectrum at low frequencies coupled with a high degree of polarization are the hallmarks of synchrotron radiation. In fact, these two features are used in astrophysics to identify synchrotron radiation from supernova remnants, active galactic nuclei, *etc.*

[8] L. Landau, and E. Lifshitz, *The classical theory of fields*, (Addison-Wesley, 1951), pp. 215 ff.

[9] J.D. Jackson, *Classical electrodynamics*, (Wiley, 1962), pp. 672 ff.