Section 8: Special Relativity

8.1 Galileo's Relativity

Before special relativity was formulated, the fundamental laws of physics were understood to obey Galileo's principle of relativity. The key concept is the *reference frame*, which we define as an oriented system of coordinates in three-dimensional space equipped with rulers and clocks to perform measurements of position and time. The latter permit us to define an event as an occurrence at a fixed point in space and time (x, y, z, t). Newton gave special attention to *inertial* frames where objects move with constant velocity if they are not acted on by external forces. Newton also emphasized the universal nature of time, in the sense that the clocks in all inertial frames tick at the same rate, independent of all external influences.

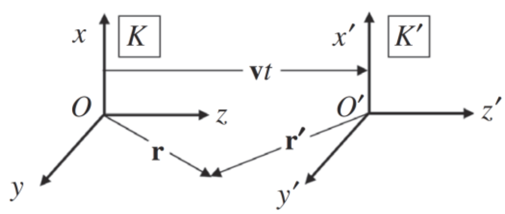

Fig. 8.1

Galileo's relativity principle states that the laws of motion are the same in all inertial frames. Consider, for example, the inertial frames K and K' shown in Figure 8.1. The frames are arranged in a "standard configuration" where their coordinate axes are similarly oriented, their origins O and O' coincide in space when $t = t' = 0$, and their relative motion occurs with constant speed $\mathbf{v}$ along the parallel axes z and z'. Often, we will say that K is the "laboratory frame". In this frame, Newton's second law reads

$$m\frac{d^2\mathbf{r}}{dt^2} = \mathbf{F}. \tag{8.1}$$

Consider now the frame K'. Newton's assumption of universal time guarantees that $t' = t$. Combining this with the vector addition indicated in Figure 8.1 gives the complete Galilean transformation as

$$\mathbf{r}' = \mathbf{r} - \mathbf{v}t \quad \text{and} \quad t' = t. \tag{8.2}$$

Because $\mathbf{v}$ is constant (by assumption), Eq. (8.2) implies that $\dfrac{d^2\mathbf{r}'}{dt'^2} = \dfrac{d^2\mathbf{r}}{dt^2}$. Therefore, if the mass does not depend on velocity, the rule $\mathbf{F}' = \mathbf{F}$ brings Newton's law into accord with Galileo's relativity principle because (8.1) transforms to

$$m\frac{d^2\mathbf{r}'}{dt'^2} = \mathbf{F}'. \tag{8.3}$$

Unlike single-particle motion, however, sound waves and water waves do not behave identically in all inertial frames. To see this explicitly, consider a scalar field which satisfies the wave equation in frame K with speed c:

$$\left[\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right] f(\mathbf{r},t) = 0. \tag{8.4}$$

The wave $f(\mathbf{r}, t)$ in frame K becomes a wave $f'(\mathbf{r}', t')$ in frame K' and its time evolution is determined by the wave operator in Eq. (8.4) transformed into primed variables. Using Eq. (8.2), the chain rule gives the derivatives we need as

$$\nabla = \sum_i \frac{\partial}{\partial x_i}\hat{\mathbf{x}}_i = \sum_i \left(\frac{\partial}{\partial x_i'}\frac{\partial x_i'}{\partial x_i} + \frac{\partial}{\partial t'}\frac{\partial t'}{\partial x_i} \right)\hat{\mathbf{x}}_i' = \sum_i \frac{\partial}{\partial x_i'}\hat{\mathbf{x}}_i' = \nabla', \tag{8.5}$$

$$\frac{\partial}{\partial t} = \sum_i \left(\frac{\partial}{\partial x_i'}\frac{\partial x_i'}{\partial t} + \frac{\partial}{\partial t'}\frac{\partial t'}{\partial t} \right) = \frac{\partial}{\partial t'} - \mathbf{v}\cdot\nabla'. \tag{8.6}$$

Using Eqs. (8.5) and (8.6) to compute the second derivatives in Eq. (8.4) gives the propagation equation in frame K' as follows:

$$\left[\nabla'^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t'^2} + \frac{2}{c^2}(\mathbf{v}\cdot\nabla')\frac{\partial}{\partial t'} - \frac{1}{c^2}(\mathbf{v}\cdot\nabla')^2 \right] f'(\mathbf{r}',t') = 0. \tag{8.7}$$

Comparing Eqs. (8.4) and (8.7) shows that the Galilean transformation (8.2) does *not* preserve the form of the wave equation (as it does Newton's second law) because classical waves propagate *relative* to any uniform motion of the host medium (water, air, etc.). For example, let $\mathbf{v} = v\hat{\mathbf{z}}$ and consider waves propagating in the $+z$-direction. If $g(s)$ is an arbitrary function of one variable, direct substitution confirms that a plane wave solution to (8.7) is

$$f(x',y',z',t') = g(z' - ct' + vt'). \tag{8.8}$$

If $v = c$, the solution (8.8) tells us that an observer at rest in frame K' sees no wave propagation at all, only a static displacement of the particles of the medium. Wave motion is *not* invariant to a Galilean transformation.

8.2 Einstein's Relativity

Einstein resolved the conceptual issues associated with the electrodynamics of moving bodies by rejecting the universal validity of Newton's laws and embracing the universal validity of Maxwell's laws. In his famous 1905 paper, he proposed the solution of this problem based on two postulates:

I. The laws of physics take the same form in every inertial frame.

II. The speed of light in vacuum is the same in every inertial frame.

Postulate I is a generalization of Galileo's relativity principle to include Maxwell's laws of electrodynamics. Postulate II explained the failure to detect relative motion between light and the host medium. These two Einstein's postulates are sufficient to construct the entire edifice of special relativity.

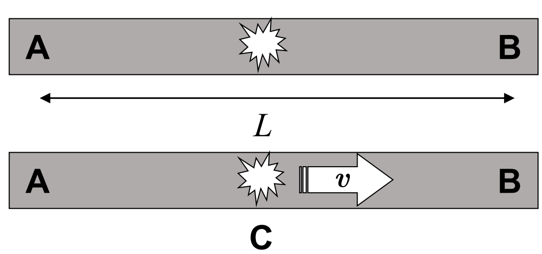

Fig. 8.2

Special relativity destroyed Newton's concept of absolute time. To make this clear, Figure 8.2 presents a thought-experiment where observers A and B wish to synchronize their clocks. They do this by agreeing to start their clocks when each detects a flash of light from a source located at the midpoint between them. In this way, each reasonably concludes that their flash observations were simultaneous events.

Now consider the same scenario from the point of view of an observer C, who sees A, B, and the light source all moving uniformly to the right with speed v. Because light travels at speed c, observer C recognizes that it reaches A sooner than it reaches B, and concludes that A starts the clock before B. If C saw A and B moving in the opposite direction, he would conclude that B starts the clock before A. Quantitatively, C computes that the two clocks are out of synchronization by an amount

$$\Delta T = \frac{L/2}{c-v} - \frac{L/2}{c+v} = \frac{L}{2}\frac{v/c}{1-v^2/c^2}. \tag{8.9}$$

This thought-experiment indicates that two inertial observers will not necessarily agree that two events are simultaneous. Therefore, Newton's concept of an absolute time breaks down when the relative velocity between frames approaches the speed of light.

8.3 Lorentz Transformation

The different perception of time by different inertial observers leads us to treat space and time on an equal footing and to locate events in a venue called *space-time*. The most general transformation law between two inertial frames K and K' in space-time is

$$x' = x'(x,y,z,t), \quad y' = y'(x,y,z,t), \quad z' = z'(x,y,z,t), \quad t' = t'(x,y,z,t). \tag{8.10}$$

The functions in Eqs. (8.10) are very restricted if we assume that the properties of space are homogeneous and do not vary from point to point or as a function of time. In particular, the infinitesimal displacement

$$dx' = \frac{\partial x'}{\partial x}dx + \frac{\partial x'}{\partial y}dy + \frac{\partial x'}{\partial z}dz + \frac{\partial x'}{\partial t}dt\ . \tag{8.11}$$

cannot be an explicit function of (x, y, z, t). This tells us that the partial derivatives in (8.11) are constants. The same is true for the three other functions in (8.11). Therefore, the transformation laws are linear functions of their arguments. It is a preview of future notation when we let r_μ (with μ = 1, 2, 3, 4) stand for x, y, z, ct and write our deduction for the transformation law to this point in the form

$$r'_\mu = \sum_{\nu=1}^{4} L_{\mu\nu} r_\nu + a_\mu\ . \tag{8.12}$$

The right side of Eq. (8.12) contains 20 parameters. This number drops to a handful if we assume that (i) the coordinate axes in frame K are aligned with their counterparts in frame K'; (ii) the origins of the two frames coincide at $t = t' = 0$; and (iii) the velocity vector which moves frame K' with respect frame K is $\mathbf{v} = v\hat{\mathbf{z}}$. This returns us to the "standard configuration" of Figure 8.1, where the most general transformation law consistent with the homogeneous and isotropic free space is

$$x' = Cx, \quad y' = Cy, \quad z' = Az + Bt, \quad t' = Dz + Et. \tag{8.13}$$

Our task is to determine the constants A, B, C, D, and E. Our first deduction is that $B = -vA$. This follows from (8.13) because the standard configuration requires $z' = 0$ to coincide with $z = vt$. By symmetry, $x' = Cx$ and $y' = Cy$ supplement (8.13) because C cannot depend on the direction of motion of one frame with respect to the other. Because of homogeneity of space, $C = 1$ is the only reasonable choice.

To find other constants, we adopt Einstein's original method and consider a point source of light which emits a spherical wave at $t = 0$ from the origin of K. Such a wave propagates radially at speed c and reaches the observation point (x, y, z) at a time t such that

$$x^2 + y^2 + z^2 - c^2t^2 = 0. \tag{8.14}$$

According to postulate II of Section 8.2, the same event viewed from frame K' of Figure 8.1 satisfies

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = 0. \tag{8.15}$$

Now, we substitute Eqs. (8.13) into (8.15) (with $B=-vA$ and $C = 1$) and insist that the result reproduces (8.14). This procedure generates three constraints on the coefficients:

$$A^2 - c^2 D^2 = 1, \tag{8.16}$$

$$E^2 - \frac{v^2}{c^2} A^2 = 1, \tag{8.17}$$

$$vA^2 + c^2 DE = 0. \tag{8.18}$$

Solving these equations, we find

$$A = E = \frac{1}{\sqrt{1 - v^2 / c^2}}, \tag{8.19}$$

$$D = -\frac{v / c^2}{\sqrt{1 - v^2 / c^2}}. \tag{8.20}$$

Therefore, the Lorentz transformation from inertial frame K to inertial frame K' is

$$x' = x, \quad y' = y, \quad z' = \frac{z - vt}{\sqrt{1 - v^2 / c^2}}, \quad t' = \frac{t - vz / c^2}{\sqrt{1 - v^2 / c^2}}. \tag{8.21}$$

The conclusion $v < c$ follows immediately from (8.21) because the real numbers (x, y, z, t) must transform into the real numbers (x', y', z', t'). No material particle or object at rest in an inertial frame can be accelerated to the speed of light.

It is standard practice in special relativity to define the symbols

$$\beta \equiv \frac{v}{c} < 1 \quad \text{and} \quad \gamma \equiv \frac{1}{\sqrt{1 - \beta^2}} > 1. \tag{8.22}$$

Using these notations, the Lorentz transformation (8.21) takes form

$$x' = x, \quad y' = y, \quad z' = \gamma\left(z - \beta ct\right), \quad ct' = \gamma\left(ct - \beta z\right). \tag{8.23}$$

By symmetry, the inverse transformation from K' to K can be obtained by exchanging the primed and unprimed variables in Eq. (8.23) and letting $v \to -v$:

$$x = x', \quad y = y', \quad z = \gamma\left(z' + \beta ct'\right), \quad ct = \gamma\left(ct' + \beta z'\right). \tag{8.24}$$

Time Dilation and Length Contraction

The mixing of space and time implied by the Lorentz transformation produces a variety of nonintuitive predictions. Consider, for example, two arbitrary events, (x_1, y_1, z_1, t_1) and (x_2, y_2, z_2, t_2), and the difference variables, $\Delta x = x_1 - x_2$, $\Delta y = y_1 - y_2$, $\Delta z = z_1 - z_2$, and $\Delta t = t_1 - t_2$. For the geometry of Figure 8.1, the linearity of Eqs. (8.23) and (8.24) implies that

$$\Delta z' = \gamma\left(\Delta z - \beta c \Delta t\right), \quad c\Delta t' = \gamma\left(c\Delta t - \beta \Delta z\right), \tag{8.25}$$

and

$$\Delta z = \gamma\left(\Delta z' + \beta c \Delta t'\right), \quad c\Delta t = \gamma\left(c\Delta t' + \beta \Delta z'\right). \tag{8.26}$$

The phenomenon of *time dilation* reveals itself when we identify the two events as two readings of a clock at rest in frame K'. Assume that the result of this measurement is the elapsed time $T' = \Delta t'$. The clock does not move in this frame, so $\Delta z' = 0$. There is no need to measure Δz. Therefore, using Eq.(8.26), we obtain for the elapsed time T in frame K:

$$T = \Delta t = \gamma \Delta t' = \frac{T'}{\sqrt{1 - v^2 / c^2}} > T' . \tag{8.27}$$

The observer in the laboratory reports a longer elapsed time than does the observer in the moving frame.

The phenomenon of *length contraction* reveals itself when we identify the two events as detecting the two end points of a rod at rest in frame K'. Assume that the result of this measurement is the length $L' = \Delta z'$. The time lapse $\Delta t'$ needed to measure the length is irrelevant in the rest frame of the rod. By contrast, the only sensible way to measure the "length" of a moving rod is to perform the sightings of its end points simultaneously in the lab frame ($\Delta t = 0$) when we establish that $L = \Delta z$. Using (8.25), we conclude that $L' = \Delta z' = \gamma \Delta z = \gamma L$. Therefore

$$L = \Delta z = \frac{1}{\gamma} \Delta z' = L' \sqrt{1 - v^2 / c^2} < L' . \tag{8.28}$$

The observer in the laboratory reports a shorter length than does the observer in the moving frame. Due to $x = x'$ and $y = y'$, there is no length contraction in the direction transverse to the direction of motion.

Invariant Interval

A relativistic *invariant* is a quantity that takes the same numerical value in every inertial frame. For example, the speed of light is a relativistic invariant. Another relativistic invariant is electric charge. There is no evidence that the charge of an electron (or proton) depends on its speed. In this section, we introduce a third invariant called the *interval* and use it to distinguish past events from future events.

The square of the interval between the two events is defined as

$$(\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - (c\Delta t)^2 . \tag{8.29}$$

The interval combines a distance in space, $d = \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}$ with a distance (or lapse) in time, Δt, into a single quantity. Like d, the interval is invariant to rotations and translations in space. Like Δt, the interval is invariant to translations in time. Most importantly, Δs is invariant to a Lorentz transformation. To prove this, we use the standard configuration (Fig. 8.1) and Eq. (8.25) to write the interval evaluated in K' in terms of the coordinates defined in K. This gives

$$\begin{aligned}(\Delta s')^2 &= (\Delta x)^2 + (\Delta y)^2 + \gamma^2 \left(\Delta z - \beta c \Delta t\right)^2 - \gamma^2 \left(c\Delta t - \beta \Delta z\right)^2 = \\ &= (\Delta x)^2 + (\Delta y)^2 + \gamma^2 \left(1 - \beta^2\right)\left(\Delta z^2 - c^2 \Delta t^2\right) = (\Delta x)^2 + (\Delta y)^2 + (\Delta x)^2 - c^2 \Delta t^2 = (\Delta s)^2 ,\end{aligned} \tag{8.30}$$

where we used $\gamma^2 \left(1 - \beta^2\right) = 1$. This proves that $(\Delta s)^2$ takes the same value in all inertial frames.

It is seen from Eq. (8.29) that $(\Delta s)^2$ can be positive, negative, or zero. The three cases differ in the nature of the "separation" between the events:

$$\begin{aligned} &(\Delta s)^2 > 0, \qquad \textit{space-like separation}, \\ &(\Delta s)^2 = 0, \qquad \textit{null separation}, \qquad . \\ &(\Delta s)^2 < 0, \qquad \textit{time-like separation}. \end{aligned} \tag{8.31}$$

A pair of events with *null separation* can be connected by a signal traveling at the speed of light. An example is the null separation between the origin and every point on the expanding spherical wave front described by Eq. (8.14). For a pair of events with a *space-like separation* $(\Delta s)^2 > 0$, the distance in space is greater than the distance $c\Delta t$ that can be covered by a light beam in the time Δt. For such events, it is always possible to perform a Lorentz transformation to an inertial frame where the event pair are simultaneous. If we call the latter frame K', and locate both events on the z' axis,

$$(\Delta s')^2 = (\Delta z')^2 - c^2 \Delta t'^2 = (\Delta z')^2 , \tag{8.32}$$

The last equality follows from the $\Delta t' = 0$ condition for simultaneity in K' and shows why the label "space-like" is used for this case. We deduce from Eq. (8.25) that

$$\beta = c \frac{\Delta t}{\Delta z} . \tag{8.33}$$

However, $|c\Delta t/\Delta z| < 1$ because $(\Delta s)^2 = (\Delta z)^2 - (c\Delta t)^2 > 0$ for the space-like separation and $\Delta x = \Delta y = 0$. This shows that the boost required to make $\Delta t' = 0$ has $\beta < 1$, which is indeed physically realizable. A similar demonstration shows that a pair of events with a *time-like separation* $(\Delta s)^2 < 0$ can always be made to occur at a single point in space ($\Delta z' = 0$). For such events, the distance in space is less than the distance $c\Delta t$ that can be covered by a light beam in the time Δt.

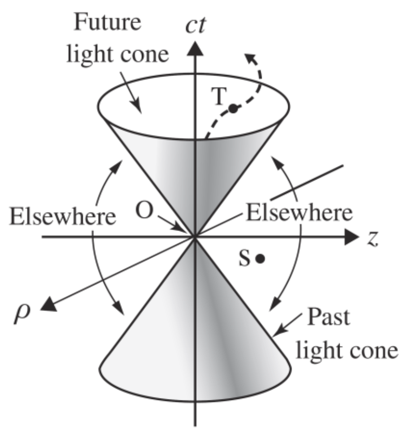

Fig. 8.3

We are now in a position to reconcile the concept of causality with the relativity of simultaneity. Figure 8.3 is a space-time or "Minkowski" diagram where the x- and y-axes are represented by a single axis ρ where $\rho^2 = x^2 + y^2$. An event labeled O occupies the origin of space-time. The two "light cones" drawn in Figure 8.3 are defined by the equation $\rho^2 + z^2 = c^2t^2$. Therefore, the interval (8.29) between O and any event on the surface of either cone is zero. From Eq. (8.31), the corresponding interval is space-like for events which lie *outside* both cones and time-like for events which lie *inside* either cone.

The event labeled S in Figure 8.3 is space-like with respect to O. This event, and all other events outside the light cones, are "absolutely distant" from O because their Euclidean distance from the origin can never be reduced to zero without violating the condition $(\Delta s)^2 > 0$ for a space-like interval. Moreover, these events cannot be said to be earlier or later than the event at O because the time interval between them can have different signs for different observers. For example, if $\Delta t > 0$, we can make $\Delta t' < 0$ in Eq. (8.25) by choosing the boost speed so $|c\Delta t/\Delta z| < \beta < 1$. The possibility of this sign inversion implies that a *cause-and-effect* relationship cannot exist between two space-like events. This is consistent with the impossibility of transforming these events to the same point in space, which would be needed to compare their clocks.

The event labeled T in Figure 8.3 is time-like with respect to O. We say that T lies in the "future light cone" of O because it occurs *later* in time than O in all inertial frames. An event which lies inside the complementary "past light cone" of O occurs *earlier* in time than O in all inertial frames. These statements are true because $\Delta t'$ in Eq. (8.25) has the same sign as Δt for all time-like intervals. This is true, in turn, because the criteria for it *not* to be true is $|\beta| > |c\Delta t/\Delta z|$, which is impossible because $(\Delta s)^2 < 0$ implies that $|c\Delta t/\Delta z| > 1$ when $\Delta x = \Delta y = 0$ in Eq. (8.29). We conclude that causality is a meaningful concept for events with a time-like separation.

Proper Time

The *proper time* is an invariant measure of the motion of a particle along its trajectory in space-time. A definition for proper time follows naturally if we define the "world line" of a particle as the locus of points in space-time which describes the trajectory in question. The dashed curve in Figure 8.3 is a typical world line for a particle with non-uniform velocity $\mathbf{u}(t) = d\mathbf{r}/dt$. All world lines lie inside the light cone because the particle speed is always less than the speed of light.

Focus now on the interval between two points on the world line which lie infinitesimally close to each other. Using Eq. (8.29), this is the time-like quantity

$$(ds)^2 = (dz)^2 - c^2dt^2 = -c^2dt^2\left[1 - \frac{u^2(t)}{c^2}\right]. \tag{8.34}$$

Dividing Eg. (8.34) by the speed of light produces another invariant and allows us to define a differential element of the invariant proper time in an inertial frame K as

$$d\tau = \sqrt{-\frac{(ds)^2}{c^2}} = dt\sqrt{1 - \frac{u^2(t)}{c^2}} = \frac{dt}{\gamma(u)}. \tag{8.35}$$

The invariance of $d\tau$ means that Eq. (8.35) has the same numerical value in any other inertial frame K' where $u' \neq u$ and $t' \neq t$. The definition of $d\tau$ tells us that the "proper time" is the time measured by a clock in its own rest frame. Note that Eq. (8.35) generalizes the meaning of γ so the argument can be a particle speed rather than the speed of one inertial frame to another.

General Lorentz Transformation

The most general Lorentz transformation between two inertial frames differs from the standard configuration transformation (8.21) in two ways. First, the velocity $\mathbf{v}$ need not lie along one of the coordinate axes as it does in the standard configuration (Fig. 8.1). Second, the Cartesian axes of frame K' may not be aligned with the Cartesian axes of frame K. In this case, we can decompose $\mathbf{r}$ into its components $\mathbf{r}_{\parallel}$ and $\mathbf{r}_{\perp}$ which lie parallel and perpendicular to $\boldsymbol{\beta} = \mathbf{v}/c$. Using these variables, the Lorentz transformation and its inverse take the forms, respectively,

$$\mathbf{r}'_{\perp} = \mathbf{r}_{\perp}, \;\; \mathbf{r}'_{\parallel} = \gamma\left(\mathbf{r}_{\parallel} - \boldsymbol{\beta}ct\right), \;\; ct' = \gamma\left(ct - \boldsymbol{\beta}\cdot\mathbf{r}_{\parallel}\right); \tag{8.36}$$

$$\mathbf{r}_{\perp} = \mathbf{r}'_{\perp}, \;\; \mathbf{r}_{\parallel} = \gamma\left(\mathbf{r}'_{\parallel} + \boldsymbol{\beta}ct'\right), \;\; ct = \gamma\left(ct' + \boldsymbol{\beta}\cdot\mathbf{r}'_{\parallel}\right). \tag{8.37}$$

8.4 Four-Vectors

Einstein's first postulate of relativity states that the laws of physics have the same form in every inertial frame. In this section, we introduce the *four-vector* as the first step toward a formalism that will make this form-invariance (or covariance) self-evident.

Let $\mathbf{a} = (a_1, a_2, a_3)$ and $\mathbf{b} = (b_1, b_2, b_3)$ be *three-vectors* in Euclidean space. The scalar product of two three-vectors is invariant to translations and rotations of the coordinate system. In other words, if K and K' are two such systems

$$\mathbf{a}\cdot\mathbf{b} = a_k b_k = a'_k b'_k = \mathbf{a}'\cdot\mathbf{b}', \tag{8.38}$$

where a repeated Latin index indicates summation over the three components of the three-vector. The invariance of the norm $|\mathbf{a}| \equiv \sqrt{\mathbf{a}\cdot\mathbf{a}}$ is a special case of Eq. (8.38).

We denote a *four-vector* in Minkowski space by

$$\vec{a} \equiv (a_1, a_2, a_3, a_4). \tag{8.39}$$

To justify calling (8.39) a "vector", we similarly require the scalar product of two four-vectors be invariant to translations, rotations, and Lorentz transformations from one inertial frame to another. In other words,

$$\vec{a}\cdot\vec{b} = a_\mu b_\mu = a'_\mu b'_\mu = \vec{a}'\cdot\vec{b}', \tag{8.40}$$

where we use a repeated Greek index to indicate summation over the four components of the four-vector. A special case is the invariance of the square length of the four-vector

$$a_\mu a_\mu = a'_\mu a'_\mu. \tag{8.41}$$

By analogy with Eq. (8.31), it is common to say that a four-vector is null, space-like, or time-like depending on whether (8.41) is zero, positive, or negative.

The prototype of a four-vector in special relativity is the space-time coordinate,

$$\vec{r} \equiv (x, y, z, ict) = (\mathbf{r}, ict). \tag{8.42}$$

The fourth component of Eq. (8.42) is a pure imaginary number. This choice ensures that Eq. (8.41) produces the appropriate minus sign when we write Eq. (8.14) as $\vec{r}\cdot\vec{r} = 0$ and the invariant interval (8.29) as $(\Delta s)^2 = \Delta\vec{r}\cdot\Delta\vec{r}$. It was precisely the assumed invariance of these quantities which led us to the standard-configuration Lorentz transformation (8.23) and its inverse (8.24). Using Eq. (8.42), matrix representations for these transformations are

$$r'_\mu = \left[\frac{\partial r'_\mu}{\partial r_\nu}\right] r_\nu \equiv L_{\mu\nu} r_\nu, \quad r_\mu = \left[\frac{\partial r_\mu}{\partial r'_\nu}\right] r'_\nu \equiv L^{-1}_{\mu\nu} r'_\nu, \tag{8.43}$$

where $\overleftrightarrow{\mathbf{L}}$ is the Lorentz transformation matrix

$$\overleftrightarrow{\mathbf{L}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & i\beta\gamma \\ 0 & 0 & -i\beta\gamma & \gamma \end{bmatrix}, \quad \overleftrightarrow{\mathbf{L}}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & -i\beta\gamma \\ 0 & 0 & i\beta\gamma & \gamma \end{bmatrix}. \tag{8.44}$$

Inspection of Eq. (8.44) shows that $\overleftrightarrow{\mathbf{L}}$ is an orthogonal matrix where $\overleftrightarrow{\mathbf{L}}^T = \overleftrightarrow{\mathbf{L}}^{-1}$. Hence,

$$L_{\mu\lambda} L^{-1}_{\lambda\nu} = L_{\mu\lambda} L_{\nu\lambda} = \delta_{\mu\nu}. \tag{8.45}$$

The determinant of these transformation matrices is one:

$$\left|\overleftrightarrow{\mathbf{L}}\right| = \left|\overleftrightarrow{\mathbf{L}}^{-1}\right| = 1. \tag{8.46}$$

By definition, four-vectors are *covariant*. A covariant vector has components that are transformed by the same matrix as the change of basis matrix. In our case, the components of an arbitrary four-vector $\vec{a}$ transform exactly like $\vec{r}$. Hence, a_1, a_2, and a_3 are often called the "space components" of $\vec{a}$, and a_4 is called the "time component" of $\vec{a}$. More precisely, $\vec{a}$ is a four-vector if

$$a'_\mu = L_{\mu\nu} a_\nu \,. \tag{8.47}$$

Referring back to (8.36), Eq. (8.47) is equivalent to

$$\mathbf{a}'_\perp = \mathbf{a}_\perp, \;\; \mathbf{a}'_\parallel = \gamma\left(\mathbf{a}_\parallel + i\boldsymbol{\beta} a_4\right), \;\; a'_4 = \gamma\left(a_4 - i\boldsymbol{\beta}\cdot\mathbf{a}_\parallel\right); \tag{8.48}$$

$$\mathbf{a}_\perp = \mathbf{a}'_\perp, \;\; \mathbf{a}_\parallel = \gamma\left(\mathbf{a}'_\parallel - i\boldsymbol{\beta} a'_4\right), \;\; a_4 = \gamma\left(a'_4 + i\boldsymbol{\beta}\cdot\mathbf{a}'_\parallel\right). \tag{8.49}$$

We leave it as an exercise for the reader to check that the transformation rules (8.48) and (8.49) guarantee that the scalar product of two four-vectors is invariant as indicated in Eq. (8.41).

Four-Velocity and Four-Acceleration

The *four-velocity* $\vec{u}$ is another prototype four-vector. It is defined in a way that it is closely related to the three-velocity $\mathbf{u} = d\mathbf{r}/dt$ of a particle, yet has the Lorentz transformation properties of the four-vector $\vec{r}$ defined in Eq. (8.42). Specifically, $\vec{u}$ is obtained by dividing the four-vector $d\vec{r}$ by a differential element of proper time, which is a Lorentz invariant scalar (see Section 8.6):

$$\vec{u} = \frac{d\vec{r}}{d\tau} = \gamma(u)\frac{d}{dt}\left(\mathbf{r}, ict\right) = \gamma(u)\left(\frac{d\mathbf{r}}{dt}, ic\right) = \gamma(u)\left(\mathbf{u}, ic\right) \equiv \left(\mathbf{U}, U_4\right), \tag{8.50}$$

Regardless of its three-velocity, Eq. (8.50) shows that $\vec{u}$ is a time-like four-vector because it defines the Lorentz invariant scalar

$$\vec{u}\cdot\vec{u} = \mathbf{U}\cdot\mathbf{U} + U_4^2 = \frac{\mathbf{u}\cdot\mathbf{u} - c^2}{1 - u^2/c^2} = -c^2 \,. \tag{8.51}$$

This is sensible because the discussion in Section 8.5 implies that we can always find an inertial frame where **u** is (instantaneously) zero.

Similarly, we define the *four-acceleration* of a particle as

$$\vec{a} = \frac{d\vec{u}}{d\tau} = \gamma(u)\frac{d}{dt}\frac{\left(\mathbf{u}, ic\right)}{\sqrt{1 - u^2/c^2}} \equiv \left(\mathbf{A}, A_4\right). \tag{8.52}$$

With $\mathbf{a} = d\mathbf{u}/dt$, the time derivative in Eq. (8.52) gives

$$\mathbf{A} == \frac{\mathbf{a}}{1 - u^2/c^2} - \frac{\mathbf{u}\left(\mathbf{u}\cdot\mathbf{a}\right)/c^2}{\left(1 - u^2/c^2\right)^2}, \quad A_4 = \frac{i\left(\mathbf{u}\cdot\mathbf{a}\right)/c}{\left(1 - u^2/c^2\right)^2} \,. \tag{8.53}$$

An interesting property of $\vec{a}$ is its orthogonality with the four-velocity $\vec{u}$. This follows immediately from Eq. (8.51) because

$$\vec{u}\cdot\vec{a} = \vec{u}\cdot\frac{d\vec{u}}{d\tau} = \frac{1}{2}\frac{d}{d\tau}\left(\vec{u}\cdot\vec{u}\right) = 0 \,. \tag{8.54}$$

Four-Momentum and Energy

The *four-momentum* $\vec{p}$ plays a central role in relativistic particle dynamics. Given the four-velocity in Eq. (8.50), we define $\vec{p}$ using a scalar $\mathcal{E}$ (not yet defined) and a three-vector **p**:

$$\vec{p} = m\vec{u} = m(\mathbf{U}, U_4) = (\mathbf{p}, i\mathcal{E}/c). \tag{8.55}$$

The mass m in Eq. (8.55) must be a Lorentz invariant scalar if we require $\vec{p}$ to be a four-vector like $\vec{u}$. $\mathcal{E}$ and $\mathbf{p}$ are defined by the following equations:

$$\mathcal{E}/c = -imU_4 = \gamma(u)mc, \quad \mathbf{p} = m\mathbf{U} = \gamma(u)m\mathbf{u}. \tag{8.56}$$

The meaning of $\mathcal{E}$ becomes clear when we Taylor expand $\gamma(u)$ in Eq. (8.56) for $u << c$ to obtain

$$\mathcal{E} = \frac{mc^2}{\sqrt{1-u^2/c^2}} = mc^2 + \frac{mu^2}{2} + \dots. \tag{8.57}$$

The second term on the far right side of Eq. (8.57) is the familiar *low-velocity kinetic energy*. The first term is a constant which may sensibly be called the *rest energy*. The *total energy* is $\mathcal{E}$ and the impossibility of accelerating a massive particle to speed c appears here as the impossibility of accelerating a particle to infinite energy. The exact kinetic energy is

$$T = \mathcal{E} - mc^2 = mc^2\left[\frac{1}{\sqrt{1-u^2/c^2}} - 1\right]. \tag{8.58}$$

The meaning of $\mathbf{p}$ in Eq. (8.56) emerges similarly from a Taylor expansion of $\gamma(u)$ for $u << c$

$$\mathbf{p} = \frac{m\mathbf{u}}{\sqrt{1-u^2/c^2}} = m\mathbf{u}\left[1 + \frac{1}{2}\frac{u^2}{c^2} + \dots\right], \tag{8.59}$$

which shows that $\mathbf{p}$ reduces to the ordinary Newtonian linear momentum, $m\mathbf{u}$, when the particle velocity is very small compared to the speed of light.

Using Eq. (8.51), the Lorentz invariant square length of the energy-momentum four-vector (8.55) is

$$\vec{p}\cdot\vec{p} = m^2\vec{u}\cdot\vec{u} = -m^2c^2. \tag{8.60}$$

8.5 Relativistic Electrodynamics

Now, we are in a position to develop a consistent formulation of relativistic electrodynamics. First, using space-time coordinate $\vec{r} = (\mathbf{r}, ict)$, we define *four-gradient*, as follows:

$$\vec{\nabla} \equiv \left(\nabla, \frac{\partial}{\partial(ict)}\right). \tag{8.61}$$

Since four-vector $\vec{r}$ transforms according to the Lorentz transformation, whose matrix is orthogonal (see Eq. (8.45)), $\vec{\nabla}$ transforms as $\vec{r}$, i.e. as required for a covariant four-vector.

The Lorentz-invariant square length of four-gradient is the wave operator:

$$\vec{\nabla}\cdot\vec{\nabla} = \nabla^2 - \frac{1}{c^2}\frac{\partial}{\partial t^2}. \tag{8.62}$$

The invariance of the wave operator (8.62) implies that waves which propagate at the speed of light in one inertial frame propagate at the speed of light in all inertial frames. This is an important consistency check because we derived the Lorentz transformation assuming precisely this behavior for a spherical wave of light.

Continuity Equation

Einstein's first postulate states that the laws of electromagnetism are valid in every inertial frame. An example is the conservation of charge, which we represent using the continuity equation,

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0 . \tag{8.63}$$

Because $\nabla \cdot \mathbf{J}$ is the scalar product of two three-vectors, and the gradient and time derivative are components of the four-vector (8.61), it is natural to inquire whether (8.63) is the scalar product (8.40) of two four-vectors. Indeed, if we guess

$$\vec{J} \equiv \left(\mathbf{J}, ic\rho\right), \tag{8.64}$$

then Eq. (8.63) assumes the frame-independent form

$$\vec{\nabla} \cdot \vec{J} = 0 . \tag{8.65}$$

The left-hand side of this equation—the scalar product of the four-gradient (8.61) with a four-vector—is called a *four-divergence*.

To confirm that $\vec{J}$ in Eq. (8.64) is a four-vector, consider a bit of electric charge $dq' = \rho' dx'dy'dz'$ at rest in the inertial frame *K′* of Figure 8.1. When viewed from the lab frame *K*, the charge is $dq = \rho dxdydz$ and $dz = dz'/\gamma$ because the volume element is contracted like (8.28) along the direction of motion. On the other hand, the invariance of electric charge requires that $dq' = dq$. To make this so, the relation between the charge densities observed in *K* and *K′* must be $\rho = \gamma\rho'$:

$$dq = \rho dxdydz = \left(\gamma\rho'\right) dx'dy'\left(dz'/\gamma\right) = \rho' dx'dy'dz' = dq' . \tag{8.66}$$

The "charge density dilation" implied by (8.66) follows immediately if (8.64) is indeed a four-vector with its law of transformation:

$$\mathbf{J}'_{\perp} = \mathbf{J}_{\perp}, \;\; \mathbf{J}'_{\parallel} = \gamma\left(\mathbf{J}_{\parallel} - \rho\mathbf{v}\right), \;\; \rho' = \gamma\left(\rho - \mathbf{v}\cdot\mathbf{J}/c^2\right). \tag{8.67}$$

Specifically, since $\mathbf{J} = \rho\mathbf{v}$ in the laboratory frame *K*, the last member of Eq. (8.67) confirms that

$$\rho' = \gamma\left(\rho - \mathbf{v}\cdot\mathbf{v}\rho/c^2\right) = \gamma\left(1-\beta^2\right)\rho = \rho/\gamma . \tag{8.68}$$

Lorenz Gauge Potentials

The gauge freedom enjoyed by the scalar potential Φ(**r**, *t*) and the vector potential **A**(**r**, *t*) implies that these quantities possess no specific transformation properties when we change inertial frames. However, they acquire quite specific transformation properties if we choose a gauge constraint that is preserved by a Lorentz transformation. The Lorenz gauge condition,

$$\nabla \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial \Phi}{\partial t} = 0 , \tag{8.69}$$

has this property if we can define the four-vector

$$\vec{A} \equiv \left(\mathbf{A}, i\Phi/c\right), \tag{8.70}$$

and use Eq. (8.61) to write Eq. (8.69) as an invariant four-divergence:

$$\vec{\nabla} \cdot \vec{A} = 0 . \tag{8.71}$$

To confirm that Eq. (8.70) is indeed a four-vector, we recall that the electromagnetic potentials in the Lorenz gauge satisfy the inhomogeneous wave equations

$$\left[\nabla^2 - \frac{1}{c^2}\frac{\partial}{\partial t^2}\right]\Phi = -\frac{\rho}{\varepsilon_0}, \qquad \left[\nabla^2 - \frac{1}{c^2}\frac{\partial}{\partial t^2}\right]\mathbf{A} = -\mu_0\mathbf{J}\,. \tag{8.72}$$

A four-vector representation of both equations in (8.72) can be written as

$$\left[\nabla^2 - \frac{1}{c^2}\frac{\partial}{\partial t^2}\right]\left(\mathbf{A}, i\Phi / c\right) = -\mu_0\left(\mathbf{J}, ic\rho\right). \tag{8.73}$$

Combining Eq. (8.73) with the Lorentz invariance of the wave operator (8.62) and the four-current character of (8.64) shows that the transformation properties on the left and right sides of Eq. (8.73) will not be the same unless (8.70) is indeed a four-vector.

Field Transformation Laws

The concepts of a "purely electric field" and a "purely magnetic field" do not exist in relativistic electromagnetism and, as we will show in this section, are intrinsically observer dependent. This conclusion is not surprising once we accept the observer-dependent meaning of the charge density and the current density implied by Eq. (8.67).

To demonstrate this explicitly, we start from

$$\mathbf{B} = \nabla\times\mathbf{A}, \qquad \mathbf{E} = -\nabla\Phi - \frac{\partial\mathbf{A}}{\partial t}, \tag{8.74}$$

and the transformation rule (8.48) for the four-vectors $\vec{\nabla}$ in Eq. (8.61) and $\vec{A}$ in Eq. (8.70). We begin with the magnetic field

$$\mathbf{B}' = \nabla'\times\mathbf{A}' = \left(\nabla'_{\parallel} + \nabla'_{\perp}\right)\times\left(\mathbf{A}'_{\parallel} + \mathbf{A}'_{\perp}\right). \tag{8.75}$$

Because $\nabla'_{\parallel}\times\mathbf{A}'_{\parallel} = 0$, the parallel and perpendicular components of $\mathbf{B}'$ are

$$\mathbf{B}'_{\parallel} = \left(\nabla'\times\mathbf{A}'\right)_{\parallel} = \nabla'_{\perp}\times\mathbf{A}'_{\perp}, \tag{8.76}$$

and

$$\mathbf{B}'_{\perp} = \left(\nabla'\times\mathbf{A}'\right)_{\perp} = \nabla'_{\parallel}\times\mathbf{A}'_{\perp} + \nabla'_{\perp}\times\mathbf{A}'_{\parallel}. \tag{8.77}$$

The three-vectors $\nabla'_{\perp}$ and $\mathbf{A}'_{\perp}$ are both transverse components of the space part of a four-vector. Thus, both are invariant under a Lorentz transformation [see Eq. (8.48)] and we conclude from Eq. (8.76) that

$$\mathbf{B}'_{\parallel} = \mathbf{B}_{\parallel}\,. \tag{8.78}$$

A bit more work is needed to evaluate Eq. (8.77) because $\nabla'_{\parallel}$ and $\mathbf{A}'_{\parallel}$ are not Lorentz invariant. Substituting from Eq. (8.48) gives

$$\mathbf{B}'_{\perp} = \gamma\left(\nabla_{\parallel} + \frac{\mathbf{v}}{c^2}\frac{\partial}{\partial t}\right)\times\mathbf{A}_{\perp} + \nabla_{\perp}\times\gamma\left(\mathbf{A}_{\parallel} - \frac{\mathbf{v}}{c^2}\Phi\right), \tag{8.79}$$

which can be rearranged to

$$\mathbf{B}'_{\perp} = \gamma\left(\nabla_{\parallel}\times\mathbf{A}_{\perp} + \nabla_{\perp}\times\mathbf{A}_{\parallel}\right) + \frac{\gamma}{c^2}\left[\mathbf{v}\times\frac{\partial\mathbf{A}_{\perp}}{\partial t} - \nabla_{\perp}\times\left(\mathbf{v}\Phi\right)\right]. \tag{8.80}$$

The structure of Eq. (8.77) and the fact that **v** is a constant vector permit us to rewrite Eq. (8.80) as

$$\mathbf{B}'_{\perp} = \gamma\mathbf{B}_{\perp} - \frac{\gamma}{c^2}\mathbf{v}\times\left(-\frac{\partial\mathbf{A}_{\perp}}{\partial t} - \nabla_{\perp}\Phi\right). \tag{8.81}$$

Finally, the definition of **E** in Eq. (8.74) simplifies Eq. (8.81) to

$$\mathbf{B}'_\perp = \gamma \mathbf{B}_\perp - \gamma \frac{\mathbf{v}}{c^2} \times \mathbf{E}_\perp = \gamma \left(\mathbf{B} - \frac{\mathbf{v}}{c^2} \times \mathbf{E} \right)_\perp . \tag{8.82}$$

For the electric field, using Eq. (8.48) and the four-vector definitions in Eq. (8.61) and (8.70), we transform $\mathbf{E}'_\parallel = -\nabla'_\parallel \Phi' - \frac{\partial \mathbf{A}'_\parallel}{\partial t'}$ to

$$\mathbf{E}'_\parallel = -\gamma \left(\nabla_\parallel + \frac{\mathbf{v}}{c^2} \frac{\partial}{\partial t} \right) \gamma \left(\Phi - \mathbf{v} \cdot \mathbf{A}_\parallel \right) - \gamma \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla_\parallel \right) \gamma \left(\mathbf{A}_\parallel - \frac{\mathbf{v}}{c^2} \Phi \right). \tag{8.83}$$

The right side of (8.83) generates eight terms. Two of these are $\pm \mathbf{v} / c^2 \partial \Phi / \partial t$, which cancel. Because **v** is a constant vector, the terms $\left(\mathbf{v} \cdot \nabla \right) \mathbf{A}_\parallel$ and $\nabla_\parallel \left(\mathbf{v} \cdot \mathbf{A}_\parallel \right)$ cancel also. The four terms which remain can be manipulated to read

$$\mathbf{E}'_\parallel = -\gamma^2 \left(\nabla_\parallel \Phi + \frac{\partial \mathbf{A}_\parallel}{\partial t} \right) \left(1 - \frac{v^2}{c^2} \right). \tag{8.84}$$

Since $\gamma^2 \left(1 - \beta^2 \right) = 1$, we have

$$\mathbf{E}'_\parallel = -\nabla_\parallel \Phi - \frac{\partial \mathbf{A}_\parallel}{\partial t} = \mathbf{E}_\parallel . \tag{8.85}$$

Similarly, we can find the transformation law for $\mathbf{E}'_\perp$ which we leave as an exercise for the reader. Finally, the field transformation laws are as follows:

$$\mathbf{E}'_\parallel = \mathbf{E}_\parallel , \quad \mathbf{E}'_\perp = \gamma \left(\mathbf{E} + \boldsymbol{\beta} \times c\mathbf{B} \right)_\perp , \quad \mathbf{E}_\perp = \gamma \left(\mathbf{E}' - \boldsymbol{\beta} \times c\mathbf{B}' \right)_\perp ; \tag{8.86}$$

$$\mathbf{B}'_\parallel = \mathbf{B}_\parallel , \quad c\mathbf{B}'_\perp = \gamma \left(c\mathbf{B} - \boldsymbol{\beta} \times \mathbf{E} \right)_\perp , \quad c\mathbf{B}_\perp = \gamma \left(c\mathbf{B}' + \boldsymbol{\beta} \times \mathbf{E}' \right)_\perp . \tag{8.87}$$

Plane Waves

Special relativity provides interesting insight into various common optical phenomena. For example, let a monochromatic plane wave propagate in vacuum with speed $c = \omega/k$ in a reference frame *K*. The electromagnetic fields in this frame are

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 \exp(i\mathbf{k} \cdot \mathbf{r} - i\omega t), \quad c\mathbf{B} = \hat{\mathbf{k}} \times \mathbf{E} . \tag{8.88}$$

The Lorentz invariance of the wave operator implies that the plane wave has exactly the same form when observed in a frame *K′* which moves with uniform speed **v** with respect to *K*:

$$\mathbf{E}'(\mathbf{r}', t') = \mathbf{E}'_0 \exp(i\mathbf{k}' \cdot \mathbf{r}' - i\omega' t'), \quad c\mathbf{B}' = \hat{\mathbf{k}}' \times \mathbf{E}' . \tag{8.89}$$

On the other hand, Eq. (8.89) can be obtained directly transforming Eq. (8.88) from *K* to the moving frame *K′*. The field amplitudes are transformed according to Eq. (8.86), so that the components of $\mathbf{E}_0$ are changed to $\mathbf{E}'_0$ and $\mathbf{B}_0$ to $\mathbf{B}'_0$. At the same time, variables **r** and *t* are transformed to *r′* and t' according to Eq. (8.37). This transformations lead to

$$\mathbf{E}'(\mathbf{r}', t') = \mathbf{E}'_0 \exp\left[i\mathbf{k} \cdot \left(\mathbf{r}'_\perp + \gamma \mathbf{r}'_\parallel + \gamma \boldsymbol{\beta} ct' \right) - i\omega\gamma \left(t' + \boldsymbol{\beta} \cdot \mathbf{r}'_\parallel / c \right) \right]. \tag{8.90}$$

Comparing Eqs. (8.90) and (8.89), we find that

$$\mathbf{k}'_\perp = \mathbf{k}_\perp , \quad \mathbf{k}'_\parallel = \gamma \left(\mathbf{k}_\parallel - \boldsymbol{\beta} \omega / c \right), \quad \omega' = \gamma \left(\omega - \mathbf{v} \cdot \mathbf{k}_\parallel \right). \tag{8.91}$$

Result (8.91) is significant because comparison with Eq. (8.48) shows that the frequency and the wave vector of a plane wave are the components of a four-vector,

$$\vec{k} = (\mathbf{k}, i\omega / c). \tag{8.92}$$

An immediate consequence of Eq. (8.92) is that the phase of a plane wave is a Lorentz invariant scalar. This is so because the phase can be written as the scalar product of two four-vectors:

$$\phi(\mathbf{r},t) = \mathbf{k}\cdot\mathbf{r} - \omega t = \vec{k}\cdot\vec{r}. \tag{8.93}$$

The invariant length of (8.92) is zero ($\vec{k}\cdot\vec{k} = 0$) because $\omega = c|\mathbf{k}|$ for vacuum waves. The transformation properties of $\vec{k}$ generate two well-known optical effects when there is relative motion between a source of waves and a detector of waves: the *Doppler effect* and *stellar aberration*. The Doppler effect refers to the fact that the frequency of the observed waves differs from the frequency of the emitted waves. Stellar aberration refers to the shift in the apparent position of a star because the direction of the observed light's wave vector differs from the direction of the emitted light's wave vector. Below we discuss some aspects of the Doppler effect.

Reflection from a Moving Mirror

We consider normal-incidence reflection from a large mirror in the *x-y* plane which moves with velocity v $\mathbf{v} = v\hat{\mathbf{z}}$ (Figure 8.4). In the laboratory, the incident wave fields with wave vector $\mathbf{k}_i = k_i\hat{\mathbf{z}}$ are

$$\mathbf{E}_i = \hat{\mathbf{x}}E_i \exp(i\mathbf{k}_i\cdot\mathbf{r} - \omega_i t), \quad c\mathbf{B}_i = \hat{\mathbf{z}}\times\mathbf{E}_i. \tag{8.94}$$

The incident wave fields in the rest frame of the mirror are

$$\mathbf{E}'_i = \hat{\mathbf{x}}E'_i \exp(i\mathbf{k}'_i\cdot\mathbf{r}' - \omega'_i t'), \quad c\mathbf{B}'_i = \hat{\mathbf{z}}\times\mathbf{E}'_i. \tag{8.95}$$

Using (8.86) and (8.91), the wave parameters in (8.95) are related to those in (8.94) as follows:

$$\begin{aligned} \omega'_i &= \gamma(\omega_i - \mathbf{v}\cdot\mathbf{k}_i) = \gamma(1-\beta)\omega_i = \sqrt{\frac{1-\beta}{1+\beta}}\,\omega_i = ck'_i, \\ \mathbf{k}'_i &= \gamma(\mathbf{k}_i - \mathbf{v}\omega_i / c^2) = \gamma(1-\beta)k_i\hat{\mathbf{z}} = \sqrt{\frac{1-\beta}{1+\beta}}\,k_i\hat{\mathbf{z}} = k'_i\hat{\mathbf{z}}, \\ E'_i &= \gamma(E_i - vB_i) = \gamma(1-\beta)E_i = \sqrt{\frac{1-\beta}{1+\beta}}\,E_i. \end{aligned} \tag{8.96}$$

The factor γ in the frequency formula is a relativistic correction to the Doppler effect formula within Newtonian physics. The correction is small when $v \ll c$, but it produces the entire *transverse Doppler effect* ($\omega'_i = \gamma\omega_i$) when $\mathbf{v}\cdot\mathbf{k}_i = 0$. This transverse effect has no counterpart in Newtonian physics.

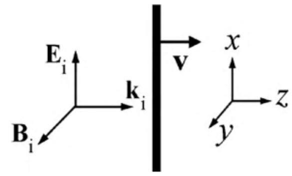

Fig. 8.4

In the rest frame of the mirror, the reflected fields have wave vector $\mathbf{k}'_r = -\mathbf{k}'_i$ and oscillate at frequency $\omega'_r = \omega'_i$. The electric field amplitude changes sign upon reflection. Therefore, with $E'_r = -E'_i$, the wave fields in the rest frame of the mirror are

$$\mathbf{E}'_r = \hat{\mathbf{x}} E'_r \exp(i\mathbf{k}'_r \cdot \mathbf{r}' - \omega'_r t'), \quad c\mathbf{B}'_r = -\hat{\mathbf{z}} \times \mathbf{E}'_r , \tag{8.97}$$

where

$$\begin{aligned}
\omega_r &= \gamma\left(\omega'_r + \mathbf{v} \cdot \mathbf{k}'_r \right) = \gamma c k'_i (1-\beta) = \left(\frac{1-\beta}{1+\beta} \right) \omega_i < \omega_i , \\
\mathbf{k}_r &= \gamma\left(\mathbf{k}'_r + \mathbf{v}\omega'_r / c^2 \right) = -\gamma k'_i (1-\beta)\,\hat{\mathbf{z}} = -\left(\frac{1-\beta}{1+\beta} \right) \mathbf{k}_i , \\
E_r &= \gamma\left(E'_r - vB'_r \right) = \gamma(1-\beta) E'_r = -\left(\frac{1-\beta}{1+\beta} \right) E_i .
\end{aligned} \tag{8.98}$$

The first line of Eq. (8.98) shows that the reflected wave frequency is “red shifted” compared to the incident wave frequency for the receding mirror situation shown in Figure 8.4. A “blue shift” ($\omega_r > \omega_i$) occurs when the mirror approaches the incident wave and $\beta \to -\beta$ in Eq. (8.98). The last line of Eq. (8.98) shows that the energy and momentum densities of the reflected wave decrease similarly (compared to the incident wave) because they are proportional to the square of the field amplitude:

$$u_{EM} = \varepsilon_0 \left| E^2 \right|, \quad c\mathbf{g} = u_{EM} \hat{\mathbf{k}} . \tag{8.99}$$

The force of radiation pressure mediates the exchange of energy and momentum between the plane wave and the moving mirror. No violation of conservation of energy or linear momentum occurs because an external agent maintains the constant speed of the mirror.

8.5 Covariant Formulation of Electrodynamics

The *covariant* formulation of classical electromagnetism refers to the way of writing the laws of classical electromagnetism (in particular, Maxwell's equations) in a form that is invariant under the Lorentz transformations. These equations prove that the laws of classical electromagnetism take the same form in any inertial coordinate system.

Lorentz Tensors

Lorentz tensors are defined in complete analogy with the rotational Cartesian tensors. Thus, a Lorentz tensor of rank 0: a one-component quantity which is invariant to a change of inertial frame

$$c' = c . \tag{8.100}$$

A Lorentz tensor of rank 1 is what we have previously called a four-vector: an object whose four components transform according to the Lorentz transformation matrix (8.44):

$$a'_\mu = L_{\mu\nu} a_\nu . \tag{8.101}$$

A Lorentz tensor of rank 2 is an object whose sixteen components transform according to the rule

$$s'_{\mu\nu} = L_{\mu\alpha} L_{\nu\beta} s_{\alpha\beta} . \tag{8.102}$$

Lorentz tensors of higher rank are defined similarly.

Three expressions derived earlier illustrate the use of Lorentz tensors manifesting covariance. These are the continuity equation (8.65), the Lorenz gauge condition (8.71), and the inhomogeneous wave equation for the Lorentz gauge potentials (8.73):

$$\frac{\partial J_\mu}{\partial x_\mu} = 0, \quad \frac{\partial A_\mu}{\partial x_\mu} = 0, \quad \frac{\partial^2 A_\nu}{\partial x_\mu \partial x_\mu} = \mu_0 J_\nu . \tag{8.103}$$

Here we defined the four-gradient $\vec{\nabla} \equiv \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}, \frac{\partial}{\partial x_4} \right) = \left(\nabla, \frac{\partial}{\partial (ict)} \right)$, and used for the four-current $\vec{J} \equiv (J_1, J_2, J_3, J_4) = (\mathbf{J}, ic\rho)$ and the four-potential $\vec{A} \equiv (A_1, A_2, A_3, A_4) = (\mathbf{A}, i\Phi / c)$.

Maxwell's Equations

Maxwell's equations can be written using Lorentz tensors. Using $\mathbf{B} = \nabla \times \mathbf{A}$, the magnetic field components are

$$
\begin{aligned}
B_x &= \frac{\partial}{\partial y} A_z - \frac{\partial}{\partial z} A_y = \frac{\partial}{\partial x_2} A_3 - \frac{\partial}{\partial x_3} A_2 , \\
B_y &= \frac{\partial}{\partial z} A_x - \frac{\partial}{\partial x} A_y = \frac{\partial}{\partial x_3} A_1 - \frac{\partial}{\partial x_1} A_3 , \\
B_z &= \frac{\partial}{\partial x} A_y - \frac{\partial}{\partial y} A_x = \frac{\partial}{\partial x_1} A_2 - \frac{\partial}{\partial x_2} A_1 .
\end{aligned}
\tag{8.104}
$$

Using $\mathbf{E} = -\nabla\Phi - \frac{\partial \mathbf{A}}{\partial t}$, the electric field components are

$$
\begin{aligned}
\frac{iE_x}{c} &= \frac{\partial A_x}{\partial (ict)} - \frac{\partial}{\partial x}\left(\frac{i\Phi}{c} \right) = \frac{\partial}{\partial x_4} A_1 - \frac{\partial}{\partial x_1} A_4 , \\
\frac{iE_y}{c} &= \frac{\partial A_y}{\partial (ict)} - \frac{\partial}{\partial y}\left(\frac{i\Phi}{c} \right) = \frac{\partial}{\partial x_4} A_2 - \frac{\partial}{\partial x_2} A_4 , \\
\frac{iE_z}{c} &= \frac{\partial A_z}{\partial (ict)} - \frac{\partial}{\partial x}\left(\frac{i\Phi}{c} \right) = \frac{\partial}{\partial x_4} A_3 - \frac{\partial}{\partial x_3} A_4 .
\end{aligned}
\tag{8.105}
$$

Equations (8.104) and (8.105) show that the Cartesian components of **E** and **B** are components of a second-rank Lorentz tensor with the form of a "generalized curl":

$$
F_{\mu\nu} = \frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu} . \tag{8.106}
$$

The electromagnetic field-strength tensor $F_{\mu\nu}$ has only 6 (rather than 16) independent components because it is asymmetric ($F_{\mu\nu} = -F_{\nu\mu}$) and the diagonal elements ($\mu = \nu$) are zero. In matrix form,

$$
\overleftrightarrow{\mathbf{F}} = \begin{bmatrix} 0 & B_z & -B_y & -iE_x / c \\ -B_z & 0 & B_x & -iE_y / c \\ B_y & -B_x & 0 & -iE_z / c \\ iE_x / c & iE_y / c & iE_z / c & 0 \end{bmatrix} . \tag{8.107}
$$

Using the field-strength tensor, the field components can be written as follows:

$$
E_k = icF_{k4} , \quad B_k = \tfrac{1}{2}\epsilon_{klm} F_{lm} , \tag{8.108}
$$

where ϵ_{klm} is the Levi-Civita symbol. It can be shown that that the tensor transformation rule applied to $F_{\mu\nu}$ reproduces the field transformation formulae (8.86) and (8.87) derived earlier.

Using $F_{\mu\nu}$ the two inhomogeneous Maxwell's equations can be written in the covariant form as follow:

$$\frac{\partial F_{\mu\nu}}{\partial x_\nu} = \mu_0 J_\mu . \tag{8.109}$$

Specifically, the μ = 4 component of Eq. (8.109) is Gauss' law, $\nabla \cdot \mathbf{E} = \rho / \varepsilon_0$, because $J_4 = ic\rho$ and $\mu_0 \varepsilon_0 c^2 = 1$:

$$i\mu_0 c\rho = \frac{\partial F_{4\mu}}{\partial x_\mu} = \frac{\partial}{\partial x}\left(\frac{iE_x}{c}\right) + \frac{\partial}{\partial y}\left(\frac{iE_y}{c}\right) + \frac{\partial}{\partial z}\left(\frac{iE_z}{c}\right) + 0 . \tag{8.110}$$

Similarly, the μ = 1, 2, 3 components of Eq. (8.109) are the *x*-, *y*-, *z* components of Ampere-Maxwell's equation, $\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \varepsilon_0 \partial \mathbf{E} / \partial t$. For example, μ = 1 gives

$$\mu_0 J_x = \frac{\partial F_{1\mu}}{\partial x_\mu} = 0 + \frac{\partial}{\partial y} B_z + \frac{\partial}{\partial z}\left(-B_y\right) + \frac{\partial}{\partial (ict)}\left(-\frac{iE_x}{c}\right) = \left(\nabla \times \mathbf{B}\right)_x - \frac{1}{c^2}\frac{\partial E_x}{\partial t} . \tag{8.111}$$

The homogeneous Maxwell's equations can also be written in terms of $F_{\mu\nu}$. However, it appears to be more convenient to use a different second-rank Lorentz tensor for that. The Lorentz transformation formulae (8.86) and (8.87) are invariant to the duality transformation $\mathbf{B} \rightarrow -\mathbf{E}/c$ and $\mathbf{E}/c \rightarrow \mathbf{B}$. Applying this symmetry operation to the elements of $F_{\mu\nu}$ we obtain an independent second-rank Lorentz tensor $G_{\mu\nu}$. The matrix form of this *dual tensor* is

$$\overleftrightarrow{\mathbf{G}} = \begin{bmatrix} 0 & -E_z / c & E_y / c & -iB_x \\ E_z / c & 0 & -E_x / c & -iB_y \\ -E_y / c & E_x / c & 0 & -iB_z \\ iB_x & iB_y & iB_z & 0 \end{bmatrix} . \tag{8.112}$$

It is straightforward to confirm that homogeneous Maxwell's equations $\nabla \cdot \mathbf{B} = 0$ and $\nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t$ are contained in the single manifestly covariant equation

$$\frac{\partial G_{\mu\nu}}{\partial x_\nu} = 0 . \tag{8.113}$$

Indeed, the μ = 4 component of Eq. (8.113) reads

$$0 = \frac{\partial G_{4\mu}}{\partial x_\mu} = i\frac{\partial B_x}{\partial x} + i\frac{\partial B_y}{\partial y} + i\frac{\partial B_z}{\partial z} = i\nabla \cdot \mathbf{B} . \tag{8.114}$$

Similarly, the μ = 1, 2, 3 components of Eq. (8.113) are the *x*-, *y*-, *z* components of Faraday's law. For example, μ = 1 gives

$$0 = \frac{\partial G_{1\mu}}{\partial x_\mu} = 0 - \frac{1}{c}\frac{\partial E_z}{\partial y} + \frac{1}{c}\frac{\partial E_y}{\partial z} - i\frac{\partial B_x}{\partial (ict)} = \frac{1}{c}\left[\left(\nabla \times \mathbf{E}\right)_x - \frac{\partial B_x}{\partial t}\right] . \tag{8.115}$$

Thus, all Maxwell's equations can be written in the covariant form as

$$\frac{\partial F_{\mu\nu}}{\partial x_\nu} = \mu_0 J_\mu , \quad \frac{\partial G_{\mu\nu}}{\partial x_\nu} = 0 . \tag{8.116}$$