Physics 221B
Spring 2020
Notes 46
Lorentz Transformations in Special Relativity†

1. Introduction

Before we examine how the Dirac equation and Dirac wave function transform under Lorentz transformations we present some material on the Lorentz transformations themselves. In these notes we will work at the level of classical special relativity, without reference to quantum mechanics, but the presentation is tailored to our needs in the next set of notes when we examine the transformation properties of the Dirac equation. The discussion in these notes is similar to that in Notes 11, which presented the theory of spatial rotations. Indeed, since spatial rotations are examples of Lorentz transformations, these notes constitute a generalization of Notes 11 in many respects. You may wish to review Notes 11 at this point to orient yourself. You may also wish to look over Appendix E for some general information on tensor analysis in special relativity, especially Secs. E.18–E.20. In the following we call on some basic knowledge of special relativity and tensor analysis.

2. Lorentz Transformations and the Lorentz Group

The usual approach to Lorentz transformations in introductory treatments of special relativity is to work out the transformations that represent boosts, usually following Einstein's postulate that the speed of light is the same in all reference frames, and then to show that these preserve the Minkowski scalar product shown in Eq. (3) below. It is assumed that you are familiar with this approach. Along the way it is sometimes mentioned that ordinary spatial rotations are also Lorentz transformations. In the following we will reverse this sequence of concepts, and define a Lorentz transformation as any linear transformation that satisfies Eq. (3). This has the advantage of unifying parity and time reversal with ordinary (proper) Lorentz transformations, which consist of boosts and rotations. In our approach, a general Lorentz transformation is a product of operations that can include rotations, boosts, parity and time reversal.

We begin with a space-time diagram, Fig. 1, which shows the coordinate axes $txyz$ of a Lorentz frame as well as a light cone. The light cone has the equation,

$$c^2t^2 = x^2 + y^2 + z^2. \tag{1}$$

It consists of two parts, the forward and backward light cones, the parts $t > 0$ and $t < 0$, respectively. The forward light cone is the locus of events generated by a light flash at $x = y = z = t = 0$, that is,

† Links to the other sets of notes can be found at:
`http://bohr.physics.berkeley.edu/classes/221/1920/221.html`.

by the spherical shell of the light pulse created by the flash, expanding at velocity c. The backwards light cone consists of another light pulse in the form of a spherical shell, this one collapsing inward at velocity c in such a way that the light reaches the spatial origin $x = y = z = 0$ at time $t = 0$. As usual in relativity theory, the interior and surface of the forward light cone are the set of events that can be reached from $x = y = z = t = 0$ by a signal traveling with velocity $v \le c$, that is, events that lie in the causal future of the event $x = y = z = t = 0$. The interior and surface of the backwards light cone consists of events that can communicate with the event $x = y = z = t = 0$ by means of a signal traveling at velocity $v \le c$, that is, events that lie in the causal past of the event $x = y = z = t = 0$.

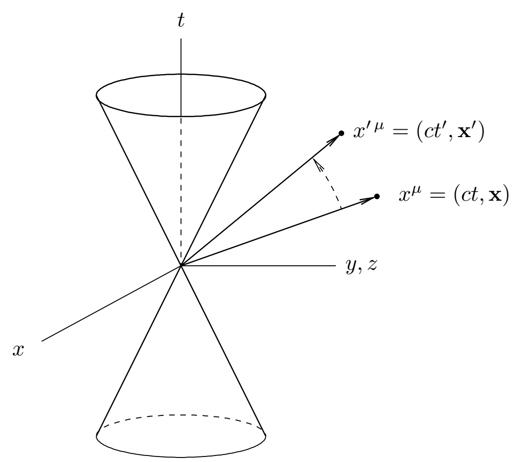

Fig. 1. A Lorentz frame with light cone. One of the spatial axes is suppressed to make a drawing possible. Two space time events, x^μ and $x'^{\,\mu}$, are shown.

Also shown in the figure is an event with coordinates $x^\mu = (ct, \mathbf{x})$. We consider a mapping of space-time onto itself, taking the (old) event x^μ into a (new) event $x'^{\,\mu}$, as shown. We suppose the mapping is linear and specified by a matrix $\Lambda^\mu{}_\nu$,

$$x'^{\,\mu} = \Lambda^\mu{}_\nu \, x^\nu. \tag{2}$$

Only linear transformations map free particle orbits (straight lines at constant velocity) into other free particle orbits. We define the mapping to be a *Lorentz transformation* if the Minkowski scalar product is preserved, that is, if

$$x'^{\,\alpha} \, g_{\alpha\beta} \, x'^{\,\beta} = x^\mu \, g_{\mu\nu} \, x^\nu \tag{3}$$

for all x^μ, where $g_{\mu\nu}$ is the Minkowski metric with components

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{4}$$

By this definition, the nature of the vector x^μ (space-like, time-light or light-like) is preserved by the transformation (3). In particular, the surface of the light cone (the set of light-like events x^μ) is mapped into itself by the transformation, as is the interior (the set of time-like events) and the exterior (the set of space-like events). As we shall see, the interior and surface of the forward light cone may be mapped into itself by the Lorentz transformation, in which case the interior and surface of the backwards light cone is also mapped into itself, or the forward and backward light cones may be mapped into each other, depending on the Lorentz transformation.

Substituting Eq. (2) into Eq. (3) and juggling indices slightly, we obtain

$$x^\mu \, \Lambda^\alpha{}_\mu \, g_{\alpha\beta} \, \Lambda^\beta{}_\nu \, x^\nu = x^\mu \, g_{\mu\nu} \, x^\nu, \tag{5}$$

or, since x^μ is arbitrary,

$$\Lambda^\alpha{}_\mu \, g_{\alpha\beta} \, \Lambda^\beta{}_\nu = g_{\mu\nu}. \tag{6}$$

A Lorentz transformation, that is, a matrix $\Lambda^\mu{}_\nu$ specifying a Lorentz transformation, is one that satisfies this equation.

Let us put this into matrix form. Let

$$\begin{aligned} \mathsf{\Lambda} &= \text{the matrix with components } \Lambda^\mu{}_\nu, \\ \mathsf{g} &= \text{the matrix with components } g_{\mu\nu}, \\ \mathsf{g}^{-1} &= \text{the matrix with components } g^{\mu\nu}, \end{aligned} \tag{7}$$

in which the first (second) index is understood as the row (column) index. As matrices, g and g^{-1} are actually the same [see Eqs. (E.74) and (E.81)], but we shall keep them notationally distinct. Then the condition (6) defining a Lorentz transformation can be written,

$$\mathsf{\Lambda}^t \mathsf{g} \mathsf{\Lambda} = \mathsf{g}, \tag{8}$$

where t indicates the transpose.

Compare this to the definition of a 3×3 orthogonal matrix, which can be written

$$\mathsf{R}^t \mathsf{I} \mathsf{R} = \mathsf{I}, \tag{9}$$

where all matrices are understood to be 3×3 and where I is the identity matrix, representing the Euclidean metric δ_{ij} on three-dimensional space. A matrix R satisfying this equation is said to belong to $O(3)$, the group of 3×3 orthogonal matrices. The properties of 3×3 orthogonal matrices are discussed in Notes 11. Similarly, Eq. (8) shows that a Lorentz transformation can be regarded as a matrix that is orthogonal with respect to the Minkowski metric $g_{\mu\nu}$. A matrix $\mathsf{\Lambda}$ satisfying Eq. (8) is said to belong to the group $O(3,1)$, which indicates the group of matrices orthogonal with respect to a metric with three space-like and one time-like directions. That is, by our definition, $O(3,1)$ is the group of Lorentz transformations.

The proof of the group property works like this. First, note that $\mathsf{\Lambda} = \mathsf{I}$ (now the 4×4 identity) satisfies Eq. (8). Next, by taking determinants of Eq. (8), we obtain

$$(\det \mathsf{\Lambda})^2 = 1, \tag{10}$$

since the determinant of the transpose of a matrix is the same as the determinant of the matrix, and since $\det \mathsf{g}$ cancels from both sides. This shows that

$$\det \Lambda = \pm 1, \tag{11}$$

and, in particular, $\det \Lambda \neq 0$, so Λ^{-1} exists. To obtain an expression for Λ^{-1} we multiply Eq. (8) by g^{-1},

$$(\mathsf{g}^{-1}\Lambda^t \mathsf{g})\Lambda = \mathsf{I}, \tag{12}$$

which shows that

$$\Lambda^{-1} = \mathsf{g}^{-1}\Lambda^t \mathsf{g}. \tag{13}$$

This is not quite as simple as the rule $\mathsf{R}^{-1} = \mathsf{R}^t$ that applies to ordinary orthogonal matrices, but it does make it easy to invert a Lorentz transformation. If Eq. (13) is translated into index language, it becomes

$$(\Lambda^{-1})^{\mu}{}_{\nu} = g^{\mu\alpha}\,\Lambda^{\beta}{}_{\alpha}\,g_{\beta\nu}. \tag{14}$$

Having shown that Λ^{-1} exists, we now show that it is a Lorentz transformation. We do this by multiplying Eq. (8) from the left by $(\Lambda^{-1})^t$, and from the right by Λ^{-1}. This gives

$$\mathsf{g} = (\Lambda^{-1})^t \mathsf{g} \Lambda^{-1}, \tag{15}$$

which shows that Λ^{-1} is a Lorentz transformation. Finally, if Λ_1 and Λ_2 are two Lorentz transformations, then

$$(\Lambda_1\Lambda_2)^t \mathsf{g}(\Lambda_1\Lambda_2) = \Lambda_2^t\Lambda_1^t \mathsf{g}\Lambda_1\Lambda_2 = \Lambda_2^t \mathsf{g}\Lambda_2 = \mathsf{g}, \tag{16}$$

which shows that the product $\Lambda_1\Lambda_2$ is also a Lorentz transformation.

In this discussion and in the rest of these notes on the Dirac equation we view Lorentz transformations as mappings of space-time onto itself, that is, we will take the active point of view. See Sec. 11.4 for a discussion of the active and passive points of view. Most introductions to relativity theory take the passive point of view, in which coordinates or tensors are transformed from one coordinate system to another. But we used the active point of view in our development of rotations in quantum mechanics in earlier sets of notes, and for consistency we must continue it with Lorentz transformations. We will give some examples presently of how the active point of view works out in practical examples, as well as some comments about how to think about it.

3. The Group Manifold of $O(3,1)$; Proper and Improper Lorentz Transformations

We now discuss the distinction between proper and improper Lorentz transformations, which is explained geometrically in terms of the group manifold of $O(3,1)$. You may wish to review Sec. 11.11, which explains the analogous situation for the group $O(3)$ of orthogonal transformations on 3-dimensional space.

The space of Lorentz transformations is 6-dimensional, that is, it takes six parameters to specify a Lorentz transformation. As we shall see, those parameters can be identified with the Euler angles of a rotation and the velocity of a boost. Recall that the space of rotations alone is 3-dimensional, and that it can be parameterized by the three Euler angles. The fact that we need six parameters to specify a Lorentz transformation follows from the definition (8). This equation can be regarded as a constraint that the matrix Λ must satisfy in order to be a Lorentz transformation. In terms of components, there are 10 independent constraints, because both sides of Eq. (8) are symmetric, 4×4 matrices, which have 10 independent components. But a general 4×4 matrix has 16 components, so the number of independent parameters in a Lorentz transformation is $16-10=6$.

The *group manifold* of $O(3,1)$ can be thought of as the 6-dimensional surface in 16-dimensional matrix space (the space of 4×4, real matrices) on which Eq. (8) is satisfied. Recall that the group manifold $O(3)$ is the 3-dimensional surface in the 9-dimensional space of 3×3 matrices such that $\mathsf{R}^t\mathsf{R}=\mathsf{I}$, and that it consists of two disconnected components containing the proper and improper rotations. Recall also that the improper rotations contain the spatial inversion operation represented by the matrix P [see Eq. (11.16)].

It turns out that the group manifold $O(3,1)$ consists of four disconnected components, which are distinguished by $\det\Lambda$, which can only take on the values ±1 [see Eq. (11)], and by the sign of the time-time component of Λ, that is, $\Lambda^0{}_0$. The meaning of this component is the following. First of all, it is easy to show from the definition (7) that

$$(\Lambda^0{}_0)^2\ge1. \tag{17}$$

Next, in a Lorentz transformation $x'^{\,\mu}=\Lambda^\mu{}_\nu\,x^\nu$, the 0-th component can be written,

$$t'=\Lambda^0{}_0\,t+\text{other terms}, \tag{18}$$

that is, $\Lambda^0{}_0$ is the coefficient connecting the old time t with the new time t'. This coefficient is the partial derivative,

$$\Lambda^0{}_0=\frac{\partial t'}{\partial t}, \tag{19}$$

where it is understood that $\mathbf{x}$ (the old spatial point) is held fixed in the differentiation. To within a sign, $\Lambda^0{}_0$ is the relativistic factor of time dilation, given in terms of the velocity $\mathbf{v}$ of the boost contained in Λ by

$$|\Lambda^0{}_0|=\gamma=\frac{1}{\sqrt{1-\dfrac{\mathbf{v}^2}{c^2}}}. \tag{20}$$

If $\Lambda^0_0>0$, then as the old time increases, so does the new time, and $\Lambda^0{}_0=\gamma\ge1$. In this case the Lorentz transformation maps the forward light cone into the forward light cone, and the backward into the backward. But if $\Lambda^0{}_0<0$, then as the old time increases, the new time decreases. This means that the Lorentz transformation contains a time-reversal operation, and that $\Lambda^0{}_0=-\gamma\le-1$. In this case the Lorentz transformation maps the forward light cone into the backward light cone and vice versa.

We can exhibit some matrices taken from the four components of the Lorentz group $O(3,1)$. In the following P is the matrix that brings about a spatial inversion but which does nothing to the time:

$$\mathsf{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{21}$$

It is a generalization of the matrix P discussed in Notes 11 and 20 [see Eq. (11.16)]. Also, T is the matrix that inverts time but does nothing to the spatial components:

$$\mathsf{T} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{22}$$

Finally, the matrix PT inverts all the space-time coordinates:

$$\mathsf{PT} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{23}$$

These matrices appear in Table 1, which gives an example of a matrix taken from each of the four components of $O(3,1)$. It is easily checked that all four matrices in Table 1 are Lorentz transformations, according to the definition (8).

Matrix	$\det \mathsf{\Lambda}$	$\operatorname{sign} \Lambda^0{}_0$
I	$+1$	$+1$
P	-1	$+1$
T	-1	-1
PT	$+1$	-1

Table 1. The four components of the Lorentz group $O(3,1)$ are distinguished by $\det \mathsf{\Lambda}$ and by $\operatorname{sign} \Lambda^0{}_0$, which amounts to specifying whether the Lorentz transformation contains a spatial inversion and/or a time-reversal. An example of a matrix taken from each component is shown.

We can picture the group manifold of $O(3,1)$ geometrically as illustrated in Fig. 2. The four components are each connected sets that are disconnected from one another. See Fig. 11.4 for the analogous diagram in the case of the group $O(3)$.

The component containing the identity I is called the *identity component*. It consists of Lorentz transformations that can be continuously connected with the identity, that is, transformations that can be built up out of products of infinitesimal Lorentz transformations. We will refer to the Lorentz transformations belonging to the identity component as *proper*, while all others will be considered

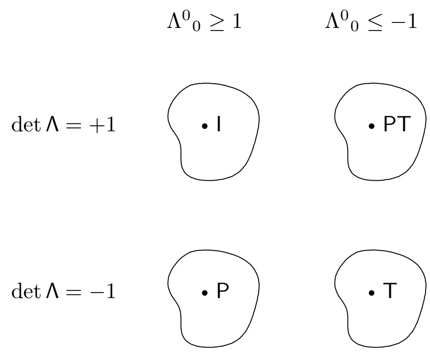

Fig. 2. The group manifold of $O(3,1)$ consists of four disconnected components (each component is a connected set, but disconnected from the other components). Each component can be viewed as a 6-dimensional surface inside 16-dimensional matrix space (the space of real, 4×4 matrices). The four components are distinguished by $\det \Lambda$ and by $\operatorname{sign} \Lambda^0{}_0$. A sample matrix taken from each component, as in Table 1, is illustrated for each component.

improper. The idea is that proper Lorentz transformations consist of proper rotations or boosts, or products of these, but do not contain either a spatial inversion or a time reversal operation. The proper Lorentz transformations constitute a group by themselves, a subgroup of the full Lorentz group $O(3,1)$.

For the time being we will concentrate on proper Lorentz transformations, and defer a discussion of parity and time-reversal to a subsequent set of notes.

4. Pure Rotations

Pure rotations are examples of proper Lorentz transformations. They are specified by rotation matrices, which are exactly the same as those presented in Notes 11, except that now we append a row and column corresponding to the time coordinate, which is not affected by the rotation. As in Notes 11, we will denote a rotation by angle θ about axis $\hat{\mathbf{n}}$ by $\mathsf{R}(\hat{\mathbf{n}},\theta)$. For example, the rotations about the three coordinate axes are

$$\mathsf{R}(\hat{\mathbf{x}},\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta & -\sin\theta \\ 0 & 0 & \sin\theta & \cos\theta \end{pmatrix}, \qquad \mathsf{R}(\hat{\mathbf{y}},\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & \sin\theta \\ 0 & 0 & 1 & 0 \\ 0 & -\sin\theta & 0 & \cos\theta \end{pmatrix}$$

$$\mathsf{R}(\hat{\mathbf{z}},\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{24}$$

These may be compared to Eqs. (11.18). These matrices are interpreted in the active sense, as in Notes 11, that is, they map an (old) unrotated point to a (new) rotated one.

5. Pure Boosts

Pure boosts are another example of proper Lorentz transformations. Boosts can be specified by the velocity $\mathbf{v}$ of the boost, but in the following we shall use a different parameterization, which is convenient for bringing out the analogy with rotations more clearly. First, we define the usual relativistic parameters β and γ in terms of the magnitude $v = |\mathbf{v}|$ of the velocity (a la Jackson, *Classical Electrodynamics*),

$$\beta = \frac{v}{c}, \qquad \gamma = \frac{1}{\sqrt{1-\beta^2}}. \tag{25}$$

By these definitions $0 \le v < c$ and $0 \le \beta < 1$, but sometimes it is convenient to let β and v take on negative values, so that $\mathbf{v}/c = (v/c)\hat{\mathbf{b}} = \beta\hat{\mathbf{b}}$, where $\hat{\mathbf{b}}$ is a unit vector such that $\mathbf{v}$ is in the same direction as $\hat{\mathbf{b}}$ (when $\beta > 0$) or opposite the direction of $\hat{\mathbf{b}}$ (when $\beta < 0$). We call $\hat{\mathbf{b}}$ the *axis* of the boost. With these understandings, $-c < v < +c$, $-1 < \beta < +1$ and $1 \le \gamma < \infty$. Next, it is convenient to introduce the *rapidity* parameter λ, defined by

$$\beta = \tanh\lambda, \qquad \gamma = \cosh\lambda, \qquad \beta\gamma = \sinh\lambda. \tag{26}$$

Then the range $-\infty < \lambda < +\infty$ corresponds to $-1 < \beta < +1$. The rapidity λ is a kind of hyperbolic angle, which plays the same role for boosts as does θ for rotations. We will write $\mathsf{B}(\hat{\mathbf{b}}, \lambda)$ for the boost along axis $\hat{\mathbf{b}}$ with rapidity λ. The velocity of the boost is given by

$$\frac{\mathbf{v}}{c} = \hat{\mathbf{b}}\tanh\lambda. \tag{27}$$

We will now work out the explicit form of the matrices $\mathsf{B}(\hat{\mathbf{b}}, \lambda)$. Here we call on some experience with Lorentz transformations in special relativity.

In any boost along a direction $\hat{\mathbf{b}}$, the two spatial coordinates orthogonal to $\hat{\mathbf{b}}$ are not affected, while the coordinate $x^0 = ct$ and the spatial coordinate parallel to $\hat{\mathbf{b}}$ transform according to the 2×2 matrix,

$$\begin{pmatrix} \gamma & \gamma\beta \\ \gamma\beta & \gamma \end{pmatrix} = \begin{pmatrix} \cosh\lambda & \sinh\lambda \\ \sinh\lambda & \cosh\lambda \end{pmatrix}. \tag{28}$$

We interpret this matrix in an active sense. An example will illustrate the meaning of active boosts, as well as showing how the signs in the matrix (28) were determined. Suppose we have a particle of mass m at rest. Then its energy-momentum 4-vector [see Eq. (44.19)] is given by

$$p_0^\mu = \begin{pmatrix} mc \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{29}$$

where the 0-subscript means "at rest." If we multiply the $x^0 = ct$ and $x^1 = x$ coordinates of this 4-vector by the matrix (28), keeping the $x^2 = y$ and $x^3 = z$ components fixed, we obtain

$$p^\mu = \begin{pmatrix} mc\gamma \\ mc\gamma\beta \\ 0 \\ 0 \end{pmatrix}, \tag{30}$$

which corresponds to the correct energy $E = mc^2\gamma$ and momentum $p_x = mv\gamma$ for a particle moving down the x-axis with velocity v. The particle has been boosted with velocity v, which is just what we want in the active sense. This allows us to write down the matrix $\mathsf{B}(\hat{\mathbf{x}}, \lambda)$.

Carrying out the same procedure for boosts down the three coordinate axes, we obtain explicit matrices $\mathsf{B}(\hat{\mathbf{b}}, \lambda)$ when $\hat{\mathbf{b}}$ is one of the coordinate directions. These are

$$\mathsf{B}(\hat{\mathbf{x}}, \lambda) = \begin{pmatrix} \cosh\lambda & \sinh\lambda & 0 & 0 \\ \sinh\lambda & \cosh\lambda & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \mathsf{B}(\hat{\mathbf{y}}, \lambda) = \begin{pmatrix} \cosh\lambda & 0 & \sinh\lambda & 0 \\ 0 & 1 & 0 & 0 \\ \sinh\lambda & 0 & \cosh\lambda & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$
$$\mathsf{B}(\hat{\mathbf{z}}, \lambda) = \begin{pmatrix} \cosh\lambda & 0 & 0 & \sinh\lambda \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh\lambda & 0 & 0 & \cosh\lambda \end{pmatrix}. \tag{31}$$

These boosts may be compared to the rotations (24).

We can also obtain an explicit expression for the boost $\mathsf{B}(\hat{\mathbf{b}}, \lambda)$ along an arbitrary direction $\hat{\mathbf{b}}$. Let $X^\mu = (X^0, \mathbf{X})$ be an arbitrary 4-vector, and let

$$X'^{\,\mu} = B(\hat{\mathbf{b}}, \lambda)^\mu{}_\nu \, X^\nu. \tag{32}$$

The boost transforms the time component X^0 and the component of the spatial vector $\mathbf{X}$ along $\hat{\mathbf{b}}$, that is, $\hat{\mathbf{b}} \cdot \mathbf{X}$, according to the matrix (28), while leaving the component of $\mathbf{X}$ orthogonal to $\hat{\mathbf{b}}$ invariant. That is, Eq. (32) is equivalent to

$$\begin{aligned} X'^{\,0} &= \cosh\lambda \, X^0 + \sinh\lambda (\hat{\mathbf{b}} \cdot \mathbf{X}), \\ \hat{\mathbf{b}} \cdot \mathbf{X}' &= \sinh\lambda \, X^0 + \cosh\lambda (\hat{\mathbf{b}} \cdot \mathbf{X}), \\ \mathbf{X}' - \hat{\mathbf{b}}(\hat{\mathbf{b}} \cdot \mathbf{X}') &= \mathbf{X} - \hat{\mathbf{b}}(\hat{\mathbf{b}} \cdot \mathbf{X}). \end{aligned} \tag{33}$$

Combining the last two equations, we write the overall transformation as

$$\begin{aligned} X'^{\,0} &= \cosh\lambda \, X^0 + \sinh\lambda (\hat{\mathbf{b}} \cdot \mathbf{X}), \\ \mathbf{X}' &= \mathbf{X} + \hat{\mathbf{b}}(\hat{\mathbf{b}} \cdot \mathbf{X})(\cosh\lambda - 1) + \sinh\lambda \, \hat{\mathbf{b}} X^0. \end{aligned} \tag{34}$$

This may be compared to the analogous formula for pure rotations, Eq. (11.44). It may also be compared to Jackson, *Classical Electrodynamics*, 3rd ed, Eq. (11.19). The formula in Jackson has minus signs compared to our Eq. (34), because he uses the passive interpretation.

A final remark about rapidities is that boosts along the same axis commute with one another, and the rapidities just add when boosts are composed. That is, we have

$$\mathsf{B}(\hat{\mathbf{b}}, \lambda_1)\mathsf{B}(\hat{\mathbf{b}}, \lambda_2) = \mathsf{B}(\hat{\mathbf{b}}, \lambda_2)\mathsf{B}(\hat{\mathbf{b}}, \lambda_1) = \mathsf{B}(\hat{\mathbf{b}}, \lambda_1 + \lambda_2). \tag{35}$$

This may compared with Eq. (11.17), which shows the analogous property for rotations. We note that boosts along different axes do not commute, in general.

Property (35) may be proved by composing two transformations of the form (34), but since the coordinates orthogonal to the direction of a boost are not affected by the boost, the essence of the proof is just the multiplication of 2×2 matrices of the form (28). In fact, calling that matrix $M(\lambda)$, it is easily verified that

$$M(\lambda_1)M(\lambda_2) = M(\lambda_1+\lambda_2). \tag{36}$$

Rapidities and velocities are related by

$$\frac{v}{c} = \tanh\lambda, \tag{37}$$

so the velocity that results on composing two boosts with rapidities λ_1 and λ_2 along the same axis is

$$v = c\tanh(\lambda_1+\lambda_2) = c\frac{\tanh\lambda_1+\tanh\lambda_2}{1+\tanh\lambda_1\tanh\lambda_2}, \tag{38}$$

where we use the addition formula for the tanh function. But this is equivalent to the usual rule in special relativity for composing velocities,

$$v = \frac{v_1+v_2}{1+\frac{v_1v_2}{c^2}}. \tag{39}$$

This rule has the property that if $|v_1|<c$ and $|v_2|<c$, then $|v|<c$. It is equivalent to the statement $|\tanh\lambda|<1$ for all λ.

6. A Useful Theorem, and Parameters of Proper Lorentz Transformations

We have presented boosts and rotations as examples of proper rotations, but there is a theorem that states that every proper Lorentz transformation can be factored in a unique way into the product of a rotation times a boost,

$$\mathsf{\Lambda} = \mathsf{R}\mathsf{B}, \tag{40}$$

where R has some axis $\hat{\mathbf{n}}$ and angle θ, and where B has some axis $\hat{\mathbf{b}}$ and rapidity λ. We will not prove this theorem (the proof is accomplished most easily in terms of the spinor representations of the proper Lorentz group), but it is a useful fact. It shows in particular that the six parameters specifying an arbitrary proper Lorentz transformation can be taken to be the three inherent in $(\hat{\mathbf{n}},\theta)$ (or, equivalently, the three Euler angles), plus the three inherent in $(\hat{\mathbf{b}},\lambda)$ (or, equivalently, the velocity $\mathbf{v}$).

7. Exponential Notation for Rotations and Boosts

It was shown in Notes 11 that proper rotations can be written in exponential form,

$$\mathsf{R}(\hat{\mathbf{n}},\theta) = \exp(\theta\hat{\mathbf{n}}\cdot\mathbf{J}) \tag{41}$$

[see Eq. (11.40)]. Here $\mathbf{J}$ is a "vector" of matrices, given by

$$\mathsf{J}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad \mathsf{J}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \qquad \mathsf{J}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{42}$$

These are the same as the matrices (11.22), except that we have appended a row and column of zeroes to take care of the time component. That is, these matrices have a $1+3$ block diagonal structure, where the 3×3 lower block is the same as in Eq. (11.22).

Boosts can also be represented in exponential form. First we consider an infinitesimal boost, in which $\lambda \ll 1$. Notice that when λ is small, then $\lambda = v/c$, which gives a physical interpretation to the rapidity in this case. Specializing Eq. (34) to the case of small λ and neglecting terms of order λ^2, we have

$$\begin{pmatrix} X'^0 \\ X'^1 \\ X'^2 \\ X'^3 \end{pmatrix} = \begin{pmatrix} X^0 \\ X^1 \\ X^2 \\ X^3 \end{pmatrix} + \lambda \begin{pmatrix} 0 & b_1 & b_2 & b_3 \\ b_1 & 0 & 0 & 0 \\ b_2 & 0 & 0 & 0 \\ b_3 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} X^0 \\ X^1 \\ X^2 \\ X^3 \end{pmatrix}, \tag{43}$$

where (b_1, b_2, b_3) are the spatial components of $\hat{\mathbf{b}}$. This can also be written,

$$\mathsf{B}(\hat{\mathbf{b}}, \lambda) = \mathsf{I} + \lambda \hat{\mathbf{b}} \cdot \mathbf{K}, \qquad (\lambda \ll 1) \tag{44}$$

where $\mathbf{K}$ is another "vector" of 4×4 matrices, given explicitly by

$$\mathsf{K}_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \mathsf{K}_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \mathsf{K}_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{45}$$

Equation (44) is the form of an infinitesimal boost, which may be compared to Eq. (11.32), which applies to infinitesimal rotations.

Now following the logic of Sec. 11.8, starting with Eq. (11.42), we can obtain an exponential form for boosts $\mathsf{B}(\hat{\mathbf{b}}, \lambda)$ when λ is not small. If λ is not small then we can use Eq. (35) to write

$$\mathsf{B}(\hat{\mathbf{b}}, \lambda) = \mathsf{B}\Big(\hat{\mathbf{b}}, \frac{\lambda}{N}\Big)^N, \tag{46}$$

where we can choose N large enough to make λ/N as small as we like. This makes it plausible that

$$\mathsf{B}(\hat{\mathbf{b}}, \lambda) = \lim_{N\to\infty} \Big(\mathsf{I} + \frac{\lambda}{N}\hat{\mathbf{b}} \cdot \mathbf{K}\Big)^N, \tag{47}$$

which by the limit discussed in Sec. 11.8 gives

$$\mathsf{B}(\hat{\mathbf{b}}, \lambda) = \exp(\lambda \hat{\mathbf{b}} \cdot \mathbf{K}). \tag{48}$$

This is the exponential form of boosts, which may be compared to Eq. (41) for rotations.

8. Infinitesimal Lorentz Transformations

Infinitesimal rotations were studied in Notes 11; when the angle θ is small, then

$$\mathsf{R}(\hat{\mathbf{n}},\theta) = \mathsf{I} + \theta(\hat{\mathbf{n}}\cdot\mathbf{J}) \qquad (\theta \ll 1) \tag{49}$$

[see Eq. (11.32)]. This can be regarded as the first two terms of the exponential series in Eq. (41). Similarly, Eq. (44) gives the form of infinitesimal boosts, which can be regarded as the first two terms of the exponential series in Eq. (48). Now multiplying the infinitesimal rotation (49) by the infinitesimal boost (44) and using the theorem (40), we obtain a general, proper, infinitesimal Lorentz transformation,

$$\Lambda = \mathsf{I} + \theta(\hat{\mathbf{n}}\cdot\mathbf{J}) + \lambda(\hat{\mathbf{b}}\cdot\mathbf{K}) \qquad (\theta,\lambda \ll 1), \tag{50}$$

where we have neglected terms that are second order in the small corrections. We see that an infinitesimal Lorentz transformation is the identity plus a correction matrix that is a linear combination of the six matrices J_i, K_i, $i=1,2,3$.

These six matrices constitute the *Lie algebra* of the Lorentz group. They satisfy commutation relations that can be worked out directly from the definitions (42) and (45),

$$[\mathsf{J}_i,\mathsf{J}_j] = \epsilon_{ijk}\,\mathsf{J}_k, \qquad [\mathsf{J}_i,\mathsf{K}_j] = \epsilon_{ijk}\,\mathsf{K}_k, \qquad [\mathsf{K}_i,\mathsf{K}_j] = -\epsilon_{ijk}\,\mathsf{J}_k. \tag{51}$$

The first of these commutators was already given in Eq. (11.34), and we have seen that it is a classical version of the standard angular momentum commutation relations in quantum mechanics, as explained in Notes 13. The second of these equations is a classical statement of the fact that $\mathbf{K}$ transforms as a vector under rotations [see Eq. (19.20) for the definition of a vector operator in quantum mechanics]. The third of these commutation relations shows that rotations can be generated as products of pure boosts. For example, when λ is small, the Lorentz transformation

$$\mathsf{B}(\hat{\mathbf{x}},\lambda)\mathsf{B}(\hat{\mathbf{y}},\lambda)\mathsf{B}(\hat{\mathbf{x}},-\lambda)\mathsf{B}(\hat{\mathbf{y}},-\lambda) \tag{52}$$

is a pure rotation about the z-axis. See the discussion in Sec. 11.13. The fact that pure rotations can be generated by products of noncommuting pure boosts lies behind Thomas precession, an effective rotation that is experienced by accelerated frames. Thomas precession is responsible for canceling one half of the spin precession expected due to spin-orbit coupling. See the discussion in Sec. 24.2.

9. A Covariant Approach to Infinitesimal Lorentz Transformations

Infinitesimal Lorentz transformations are important because an arbitrary, proper Lorentz transformation can be built up as products of them. But the decomposition of them into rotations and boosts, as in Eq. (50), is not entirely covariant. That is, what is purely a boost in one frame is partly a boost, partly a rotation in another. We now develop a completely covariant approach to infinitesimal Lorentz transformations (that is, to the Lie algebra of the group), which will be useful to us in Notes 47 for establishing the covariance of the Dirac equation. (Note that the words "covariance" and "covariant" have two distinct meanings, as explained in Sec. E.1.)

Let a Lorentz transformation have the form,

$$\Lambda = \mathsf{I} + \epsilon \mathsf{C}, \tag{53}$$

where C is the correction matrix and ϵ is a small scale factor used to remind us that the correction is small. This Lorentz transformation is infinitesimal. Plugging this into the definition (8) of a Lorentz transformation, we have

$$(\mathsf{I} + \epsilon \mathsf{C}^t)\mathsf{g}(\mathsf{I} + \epsilon \mathsf{C}) = \mathsf{g}, \tag{54}$$

or, through first order in ϵ,

$$\mathsf{C}^t\mathsf{g} + \mathsf{g}\mathsf{C} = (\mathsf{g}\mathsf{C})^t + \mathsf{g}\mathsf{C} = 0. \tag{55}$$

This shows that gC is an antisymmetric, 4×4 matrix. Compare this to what we did with ordinary rotations in Sec. 11.7.

As in Sec. 11.7, it is convenient to introduce a basis of antisymmetric matrices. The difference is that now we are dealing with 4×4 matrices, so the basis consists of six matrices altogether. Let $\mathsf{A}^{\mu\nu}$ be the 4×4 matrix with a 1 in component $\mu\nu$ and a -1 in component $\nu\mu$ (the A just stands for "antisymmetric"). The indices $\mu\nu$ attached to $\mathsf{A}^{\mu\nu}$ are labels of the matrix, not component indices. If we wish to specify the components of the matrix $\mathsf{A}^{\mu\nu}$, we will write $(A^{\mu\nu})_{\alpha\beta}$. According to the definition, these components are

$$(A^{\mu\nu})_{\alpha\beta} = \delta^\mu_\alpha\, \delta^\nu_\beta - \delta^\nu_\alpha\, \delta^\mu_\beta, \tag{56}$$

where the two terms put a 1 in slot $(\alpha\beta) = (\mu\nu)$, and a -1 in slot $(\alpha\beta) = (\nu\mu)$. The matrices $\mathsf{A}^{\mu\nu}$ satisfy

$$(\mathsf{A}^{\mu\nu})^t = -\mathsf{A}^{\mu\nu} = \mathsf{A}^{\nu\mu}, \tag{57}$$

in particular, they are antisymmetric. If we let μ and ν run over $0, 1, 2, 3$, we get 16 matrices altogether. We can think of these matrices as constituting an antisymmetric, 4×4 "tensor" of antisymmetric matrices, much as J_i constitutes a "vector" of matrices. Since this "tensor" is antisymmetric, that is, $\mathsf{A}^{\mu\nu} = -\mathsf{A}^{\nu\mu}$, four of these matrices vanish and the remaining twelve consist of six pairs that are equal and opposite. We see that there are only six linearly independent matrices $\mathsf{A}^{\mu\nu}$, which we get by restricting μ and ν to the range $0 \le \mu < \nu \le 3$.

Thus a given antisymmetric matrix, such as gC in Eq. (55), can be written as a linear combination of the six independent matrices $\mathsf{A}^{\mu\nu}$ with coefficients that we will call $\theta_{\mu\nu}$:

$$\mathsf{gC} = \sum_{\mu<\nu} \theta_{\mu\nu} \mathsf{A}^{\mu\nu}. \tag{58}$$

In this equation, you must remember that $\theta_{\mu\nu}$ are numbers and $\mathsf{A}^{\mu\nu}$ are matrices. Given gC, the numbers $\theta_{\mu\nu}$ are uniquely determined for $\mu < \nu$. We now define $\theta_{\mu\nu}$ for other values of μ, ν by requiring,

$$\theta_{\mu\nu} = -\theta_{\nu\mu}, \tag{59}$$

so that $\theta_{\mu\nu}$ can be thought of as an antisymmetric, 4×4 matrix or space-time tensor. Then using Eq. (57), the expansion (58) can be written,

$$\mathsf{gC} = \frac{1}{2}\theta_{\mu\nu}\mathsf{A}^{\mu\nu}, \tag{60}$$

where we drop the sum and use the summation convention, since now we have an unrestricted sum over μ and ν.

Finally, we multiply Eq. (60) by g^{-1} and write the result as

$$\mathsf{C} = \frac{1}{2}\theta_{\mu\nu}\mathsf{J}^{\mu\nu}, \tag{61}$$

where

$$\mathsf{J}^{\mu\nu} = \mathsf{g}^{-1}\mathsf{A}^{\mu\nu}. \tag{62}$$

Again, the indices μ, ν on $\mathsf{J}^{\mu\nu}$ are labels of the matrices, not component indices. You may compare Eq. (61), the expansion of the correction matrix for an infinitesimal Lorentz transformation in terms of a set of basis matrices, with Eq. (11.21), the analogous expression for pure rotations.

Combining Eqs. (53) and (61), we can write the infinitesimal Lorentz transformation itself as

$$\Lambda = \mathsf{I} + \frac{1}{2}\theta_{\mu\nu}\,\mathsf{J}^{\mu\nu} \qquad (\theta_{\mu\nu} \ll 1), \tag{63}$$

where we have absorbed the ϵ into the parameters $\theta_{\mu\nu}$, which, as noted, are small. Since both this expression and Eq. (50) represent general, infinitesimal Lorentz transformations, they must be equal to each other.

To establish the connection between these two expressions, we first work out the components of $\mathsf{J}^{\mu\nu}$, which, according to Eqs. (7) and (62), are the same as the components of $\mathsf{A}^{\mu\nu}$ with the first index raised. The components of $\mathsf{A}^{\mu\nu}$ are given by Eq. (56). Raising the index α in that equation, we obtain

$$(\mathsf{J}^{\mu\nu})^{\alpha}{}_{\beta} = g^{\mu\alpha}\,\delta^{\nu}_{\beta} - g^{\nu\alpha}\,\delta^{\mu}_{\beta}. \tag{64}$$

If we set $(\mu\nu) = (0i)$ or $(\mu\nu) = (ij)$ in this expression and work out the components of the matrices, we find

$$\mathsf{J}^{0i} = \mathsf{K}_i, \qquad \mathsf{J}^{ij} = \epsilon_{ijk}\,\mathsf{J}_k, \tag{65}$$

where J_i and K_i are given by Eqs. (42) and (45). Thus, the antisymmetric "tensor" of matrices $\mathsf{J}^{\mu\nu}$ is equivalent to the two "vectors" of matrices $\mathbf{J}$ and $\mathbf{K}$, and it contains the Lie algebra of the Lorentz group. The commutation relations (51), expressed in terms of the matrices $\mathbf{J}$, $\mathbf{K}$, can be expressed in terms of the matrices $\mathsf{J}^{\mu\nu}$. Alternatively, the commutation relations among the $\mathsf{J}^{\mu\nu}$ can be worked out directly from Eq. (64). Although we will not use the result we note it for reference:

$$[\mathsf{J}^{\mu\nu}, \mathsf{J}^{\sigma\tau}] = g^{\nu\sigma}\,\mathsf{J}^{\mu\tau} - g^{\nu\tau}\,\mathsf{J}^{\mu\sigma} - g^{\mu\sigma}\,\mathsf{J}^{\nu\tau} + g^{\mu\tau}\,\mathsf{J}^{\nu\sigma}. \tag{66}$$

Breaking the sum on μ, ν in Eq. (63) into its space and time components and using Eq. (65), we can write the infinitesimal Lorentz transformation as

$$\Lambda = \mathsf{I} + \frac{1}{2}(\theta_{0i}\,\mathsf{J}^{0i} + \theta_{i0}\,\mathsf{J}^{i0}) + \frac{1}{2}\theta_{ij}\,\mathsf{J}^{ij} = \mathsf{I} + \theta_{0i}\,\mathsf{K}_i + \frac{1}{2}\epsilon_{ijk}\,\theta_{ij}\,\mathsf{J}_k. \tag{67}$$

Comparing this to Eq. (50), we can read off the relations between the six parameters θ, $\hat{\mathbf{n}}$, λ and $\hat{\mathbf{b}}$ of an infinitesimal Lorentz transformation, and the six independent parameters contained in $\theta_{\mu\nu}$. To do this it is convenient to define two vectors,

$$\boldsymbol{\theta} = \theta\hat{\mathbf{n}} \tag{68}$$

and

$$\boldsymbol{\lambda} = \lambda\hat{\mathbf{b}}, \tag{69}$$

which are equivalent to

$$\hat{\mathbf{n}} = \frac{\boldsymbol{\theta}}{\theta}, \qquad \theta = |\boldsymbol{\theta}|, \tag{70}$$

and

$$\hat{\mathbf{b}} = \frac{\boldsymbol{\lambda}}{\lambda}, \qquad \lambda = |\boldsymbol{\lambda}|. \tag{71}$$

Then these vectors are related to the coefficients $\theta_{\mu\nu}$ by

$$\theta_i = \frac{1}{2}\epsilon_{ijk}\,\theta_{jk} \quad \text{or} \quad \theta_{ij} = \epsilon_{ijk}\,\theta_k, \tag{72}$$

and

$$\lambda_i = \theta_{0i} = -\theta_{i0}. \tag{73}$$

If the infinitesimal Lorentz transformation is a pure rotation, then

$$\Lambda = \mathsf{I} + \theta\hat{\mathbf{n}}\cdot\mathbf{J} = \mathsf{I} + \boldsymbol{\theta}\cdot\mathbf{J}, \tag{74}$$

and

$$\theta_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \theta_3 & -\theta_2 \\ 0 & -\theta_3 & 0 & \theta_1 \\ 0 & \theta_2 & -\theta_1 & 0 \end{pmatrix}. \tag{75}$$

If the infinitesimal Lorentz transformation is a pure boost, then

$$\Lambda = \mathsf{I} + \lambda\hat{\mathbf{b}}\cdot\mathbf{K} = \mathsf{I} + \boldsymbol{\lambda}\cdot\mathbf{K}, \tag{76}$$

and

$$\theta_{\mu\nu} = \begin{pmatrix} 0 & \lambda_1 & \lambda_2 & \lambda_3 \\ -\lambda_1 & 0 & 0 & 0 \\ -\lambda_2 & 0 & 0 & 0 \\ -\lambda_3 & 0 & 0 & 0 \end{pmatrix}. \tag{77}$$

10. Active and Passive Lorentz Transformations

In Sec. 5 we made a few comments about how to interpret Lorentz transformations in an active sense. Here we will elaborate somewhat on this question. We mainly think in terms of boosts, since the active interpretation of rotations was covered in Notes 11.

Let the world line of a particle be given by $x^\mu(\tau)$, where τ is the proper time. Then under an active, proper Lorentz transformation specified by matrix $\Lambda^\mu{}_\nu$, the state of motion of the particle specified by $x^\mu(\tau)$ mapped into a new state of motion $x'^\mu(\tau)$, given by

$$x'^\mu(\tau) = \Lambda^\mu{}_\nu \, x^\nu(\tau). \tag{78}$$

In Sec. 5 we considered the special case in which the original motion $x^\mu(\tau)$ was that of a particle at rest, but here $x^\mu(\tau)$ can represent any state of motion, including accelerated motion. The momentum p^μ transforms by the same matrix $\Lambda^\mu{}_\nu$ as the space-time vector x^μ, since

$$p^\mu = m\frac{dx^\mu}{d\tau}. \tag{79}$$

See Eq. (B.49). Notice that τ is invariant under proper Lorentz transformations.

In the passive point of view which is used in most introductions to special relativity one speaks of two frames, the "lab" and "moving" frames, and one studies the relation between the space-time coordinates of the particle as seen in the two frames. To connect the present, active interpretation with the passive interpretation, first consider boosting a particle with space time coordinates x^μ in the active sense. Let the Lorentz transformation be parameterized by the velocity $\mathbf{v}_a$ of the boost, so that the space-time coordinates of the particle after the boost are

$$x'^\mu = \Lambda(\mathbf{v}_a)^\mu{}_\nu \, x^\nu, \tag{80}$$

where we put an a subscript on $\mathbf{v}$ to indicate that it is the velocity of the boost in the active interpretation. The meaning of $\mathbf{v}_a$ is that if we boost a particle at rest, then $\mathbf{v}_a$ is the velocity of the particle after the boost.

To obtain the passive interpretation we need a second frame. To obtain it we boost not only the particle but also the original Lorentz frame, both by the same Lorentz transformation $\Lambda(\mathbf{v}_a)^\mu{}_\nu$. This produces a boosted frame, the velocity of which with respect to the original frame is $\mathbf{v}_a$. Also, the space-time coordinates of the particle with respect to the boosted frame are the same as the space-time coordinates of the original particle with respect to the original frame, that is, they are x^μ in Eq. (80). Now we interpret the boosted frame as the "lab" frame and the original frame as the "moving" frame. Then Eq. (80) can be written,

$$x^\mu_{\text{moving}} = \Lambda(\mathbf{v}_a)^\mu{}_\nu \, x^\nu_{\text{lab}}. \tag{81}$$

However, in the passive interpretation we usually parameterize the boost by the velocity of the moving frame with respect to the lab frame, which is $-\mathbf{v}_a$ in the present notation.

In summary, the active equation (80) is put into passive form by interpreting x'^μ as x^μ_{moving} and x^μ as x^μ_{lab}, and changing the sign on the velocity parameter of the boost. You may compare our Eq. (34) with Jackson's Eq. (11.19) with this rule in mind.

11. Transforming Fields

Equation (78) gives the rule for Lorentz transforming the state of a particle in an active sense, but how do we transform fields? For example, what does it mean to boost a field? The answer can be given in a physical sense, for example, in the case of the electromagnetic field $F^{\mu\nu}$. The electromagnetic field is produced by charges and currents, according to the Maxwell equations,

$$\partial_\mu F^{\mu\nu} = \frac{4\pi}{c} J^\nu, \tag{82}$$

where J^μ is the electric current 4-vector. The current in turn is produced by charges in some state of motion. It is logical to define the Lorentz transformed current as the current produced by the charges in the Lorentz transformed state of motion, according to Eq. (78), and the Lorentz transformed electromagnetic field as the field produced by the Lorentz transformed current. See Problem 2 where these ideas are explored.

The same results can be obtained in another way. Let us begin with the example of a scalar field $S(\mathbf{x})$ on 3-dimensional space. If we apply an active rotation specified by a matrix R, then it is logical to define the rotated field as the one that we get by allowing the points to carry the value of the scalar field with them as they are rotated. That is, if $\mathbf{x}$ is an "old" point and $\mathbf{x}' = \mathsf{R}(\mathbf{x})$ is a "new" or rotated point, then we require that the rotated field S' satisfy

$$S'(\mathbf{x}') = S(\mathbf{x}). \tag{83}$$

In words, the value of the new field at the new point is equal to the value of the old field at the old point. This same rule was discussed in Sec. 15.2 in connection with the transformation properties of the wave function of a spinless particle, see Eq. (15.15). Although that discussion was framed in terms of the transformation of a quantum wave function, the same concepts apply to any scalar field, for example, the temperature or pressure of a fluid. If we substitute $\mathbf{x}' = \mathsf{R}\mathbf{x}$ into Eq. (83) we get

$$S'(\mathsf{R}\mathbf{x}) = S(\mathbf{x}), \tag{84}$$

or, making the replacement $\mathbf{x} \to \mathsf{R}^{-1}\mathbf{x}$,

$$S'(\mathbf{x}) = S\big(\mathsf{R}^{-1}\mathbf{x}\big). \tag{85}$$

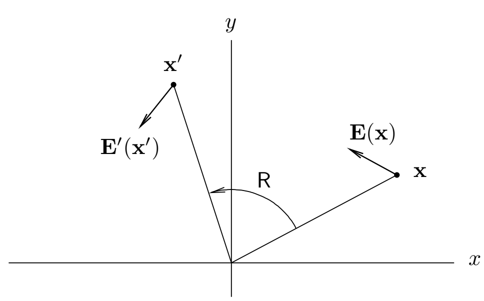

Fig. 3. When we rotate a vector field such as an electric field, we rotate the value of the field along with the point at which the field is evaluated.

What about a vector field, such as an electric field $\mathbf{E}(\mathbf{x})$? The most obvious definition of the rotated field from a geometrical standpoint is illustrated in Fig. 3, in which the value of the field is rotated along with the point at which the field is evaluated. That is, the rule is

$$\mathbf{E}'(\mathbf{x}') = \mathsf{R}\mathbf{E}(\mathbf{x}), \tag{86}$$

where again $\mathbf{x}' = \mathsf{R}\mathbf{x}$. Again making the replacement $\mathbf{x} \to \mathsf{R}^{-1}\mathbf{x}$, this becomes

$$\mathbf{E}'(\mathbf{x}) = \mathsf{R}\mathbf{E}(\mathsf{R}^{-1}\mathbf{x}). \tag{87}$$

Now let $S(x)$ be a scalar field on space-time, where x stands for all four space-time coordinates x^μ or $(\mathbf{x}, t)$. The transformation law (85) for scalar fields under ordinary rotations suggests that scalar fields should transform under Lorentz transformations according to

$$S'(x) = S\big(\Lambda^{-1}x\big). \tag{88}$$

Examples of scalar fields in electromagnetism include

$$F_{\mu\nu}F^{\mu\nu} = 2(B^2 - E^2), \tag{89}$$

and

$$\epsilon_{\alpha\beta\mu\nu}\, F^{\alpha\beta}\, F^{\mu\nu} = 4\mathbf{E}\cdot\mathbf{B}, \tag{90}$$

the two classic invariants of the electromagnetic field. [Actually, the quantity (90) is a pseudo-scalar, not a scalar.] Another scalar field is the Klein-Gordon wave function $\psi(x)$ studied in Notes 44.

Similarly, let $X^\mu(x)$ be a vector field on space-time. Then the transformation law (87) for vector fields under ordinary rotations suggests that vector fields should transform under Lorentz transformations according to

$$X'^{\,\mu}(x) = \Lambda^\mu{}_\nu\, X^\nu\big(\Lambda^{-1}x\big). \tag{91}$$

Indeed, it is shown in Prob. 2 that the electric current J^μ transforms in exactly this way, when the particles creating the current are boosted by Λ. Similarly, the electromagnetic field tensor $F^{\mu\nu}$ transforms according to

$$F'^{\,\mu\nu}(x) = \Lambda^\mu{}_\alpha\, \Lambda^\nu{}_\beta\, F^{\alpha\beta}\big(\Lambda^{-1}x\big). \tag{92}$$

In the next set of notes we will consider how Dirac spinors transform under Lorentz transformations, a necessary step in discussing the covariance of the Dirac equation.

Problems

1. Some general properties of Lorentz transformations, following from the definition (6) or (8).

(a) Show that if a matrix Λ satisfies Eq. (8), then the component $\Lambda^0{}_0$ satisfies Eq. (17). Hint: Examine the 00 component of Eq. (8).

(b) Below Eq. (19) we referred to "the velocity $\mathbf{v}$ of the boost contained in Λ." In this part of the problem we will explain what this phrase means. Let a particle have a world line specified by $y^\mu(\tau)$, where τ is the proper time, that is, the time as indicated by a clock carried by the particle. We allow the particle to be in any state of motion, in particular, the particle may be accelerated. When we apply a Lorentz transformation specified by $\Lambda^\mu{}_\nu$ to this particle, interpreted in the active sense, we obtain the same particle in a different state of motion specified by $x^\mu(\tau)$, where

$$x^\mu(\tau) = \Lambda^\mu{}_\nu \, y^\nu(\tau). \tag{93}$$

The proper time is the same before and after the Lorentz transformation, because

$$c^2 \, d\tau^2 = g_{\mu\nu} \, dy^\mu \, dy^\nu = g_{\mu\nu} \, dx^\mu \, dx^\nu. \tag{94}$$

This implies that $d\tau^2$ is invariant.

Now suppose the particle is at rest before the Lorentz transformation is applied. Express the four components y^μ as explicit functions of τ. Let $\mathbf{v}$ be the velocity of the particle after the Lorentz transformation, and express its components v_i in terms of $\Lambda^i{}_0$ and $\Lambda^0{}_0$. We regard this $\mathbf{v}$ as the velocity of the boost contained in Λ. Use this expression to prove Eq. (20).

2. Boosting particles to get the boosted current. In this problem we choose units so that $c = 1$.

(a) Consider a particle of charge q whose trajectory through space is given by $\mathbf{y}(t)$. Let $\mathbf{x}$ be a field point. Then the charge and current densities are

$$\begin{aligned} \rho(\mathbf{x}, t) &= q\delta\big(\mathbf{x} - \mathbf{y}(t)\big), \\ \mathbf{J}(\mathbf{x}, t) &= q\mathbf{v}\delta\big(\mathbf{x} - \mathbf{y}(t)\big), \end{aligned} \tag{95}$$

where $\mathbf{v} = d\mathbf{y}/dt$ is the velocity of the particle.

Now let τ be the proper time of the particle as measured by a clock carried along with the particle. Also let the proper time τ and the coordinate time t be related by a function y^0,

$$t = y^0(\tau), \tag{96}$$

and let

$$\mathbf{y}(t) = \mathbf{y}\big(y^0(\tau)\big) \equiv \mathbf{y}(\tau). \tag{97}$$

This is a way of expressing the 4-vector giving the space-time coordinates of the particle as a function of the proper time,

$$y^\mu(\tau) = \begin{pmatrix} y^0(\tau) \\ \mathbf{y}(\tau) \end{pmatrix}. \tag{98}$$

Show that the current 4-vector is given by an integral along the particle's orbit (and remember that $c = 1$),

$$J^\mu(\mathbf{x}, t) = \begin{pmatrix} \rho(\mathbf{x}, t) \\ \mathbf{J}(\mathbf{x}, t) \end{pmatrix} = q \int d\tau \, \frac{dy^\mu}{d\tau} \, \delta^4\big(x - y(\tau)\big), \tag{99}$$

where if X^μ is a 4-vector then

$$\delta^4(X) = \delta(X^0)\delta(X^1)\delta(X^2)\delta(X^3). \tag{100}$$

Hint: Use the τ-integration to eliminate the delta function $\delta\bigl(x^0 - y^0(\tau)\bigr)$.

(b) If the particle motion $y^\mu(\tau)$ is mapped into a new motion $y'^{\,\mu}(\tau)$ by a Lorentz transformation $\Lambda^\mu{}_\nu$,

$$y'^{\,\mu}(\tau) = \Lambda^\mu{}_\nu \, y^\nu(\tau), \tag{101}$$

then show that the 4-current produced by the Lorentz transformed particles is given by

$$J'^{\,\mu}(x) = \Lambda^\mu{}_\nu \, J^\nu\bigl(\Lambda^{-1}x\bigr). \tag{102}$$

Hint: Use the fact that if M is a 4×4 matrix and X is a 4-vector, then

$$\delta^4(MX) = \frac{\delta^4(X)}{|\det M|}. \tag{103}$$

3. Let a^μ and b^μ be two 4-vectors. Let us write

$$a \cdot b = a^\mu \, b_\mu = a^\mu \, g_{\mu\nu} \, b^\nu \tag{104}$$

as a shorthand for the Minkowski scalar product. Let Λa be a shorthand for the 4-vector whose μ-th component is $\Lambda^\mu{}_\nu \, a^\nu$. Show that

$$a \cdot (\Lambda^{-1} b) = (\Lambda a) \cdot b. \tag{105}$$