12.1.4 The Structure of Spacetime

(i) Four-vectors. The Lorentz transformations take on a simpler appearance when expressed in terms of the quantities

$$x^0 \equiv ct, \quad \beta \equiv \frac{v}{c}. \tag{12.21}$$

Using x^0 (instead of t) and β (instead of v) amounts to changing the unit of time from the *second* to the *meter*—1 meter of x^0 corresponds to the time it takes light to travel 1 meter (in vacuum). If, at the same time, we number the x, y, z coordinates, so that

$$x^1 = x, \quad x^2 = y, \quad x^3 = z, \tag{12.22}$$

then the Lorentz transformations read

$$\left.\begin{aligned} \bar{x}^0 &= \gamma(x^0 - \beta x^1), \\ \bar{x}^1 &= \gamma(x^1 - \beta x^0), \\ \bar{x}^2 &= x^2, \\ \bar{x}^3 &= x^3. \end{aligned}\right\} \tag{12.23}$$

Or, in matrix form:

$$\begin{pmatrix} \bar{x}^0 \\ \bar{x}^1 \\ \bar{x}^2 \\ \bar{x}^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}. \tag{12.24}$$

Letting Greek indices run from 0 to 3, this can be distilled into a single equation:

$$\bar{x}^\mu = \sum_{\nu=0}^{3} (\Lambda^\mu_\nu) x^\nu, \tag{12.25}$$

where Λ is the **Lorentz transformation matrix** in Eq. 12.24 (the superscript μ labels the row, the subscript ν labels the column). One virtue of writing things in this abstract manner is that we can handle in the same format a more general transformation, in which the relative motion is *not* along a common $x\,\bar{x}$ axis; the matrix Λ would be more complicated, but the structure of Eq. 12.25 is unchanged.

If this reminds you of the *rotations* we studied in Sect. 1.1.5, it's no accident. There we were concerned with the change in components when you switch to a *rotated* coordinate system; here we are interested in the change of components when you go to a *moving* system. In Chapter 1 we defined a (3-) vector as any set of three components that transform under rotations the same way (x, y, z) do; by extension, we now define a **4-vector** as any set of *four* components that transform in the same manner as (x^0, x^1, x^2, x^3) under Lorentz transformations:

$$\bar{a}^\mu = \sum_{\nu=0}^{3} \Lambda^\mu_\nu a^\nu. \tag{12.26}$$

For the particular case of a transformation along the x axis:

$$\left.\begin{aligned} \bar{a}^0 &= \gamma(a^0 - \beta a^1), \\ \bar{a}^1 &= \gamma(a^1 - \beta a^0), \\ \bar{a}^2 &= a^2, \\ \bar{a}^3 &= a^3. \end{aligned}\right\} \tag{12.27}$$

There is a 4-vector analog to the dot product ($\mathbf{A} \cdot \mathbf{B} \equiv A_x B_x + A_y B_y + A_z B_z$), but it's not just the sum of the products of like components; rather, the zeroth components have a minus sign:

$$-a^0 b^0 + a^1 b^1 + a^2 b^2 + a^3 b^3. \tag{12.28}$$

This is the **four-dimensional scalar product**; you should check for yourself (Prob. 12.17) that it has the same value in all inertial systems:

$$-\bar{a}^0 \bar{b}^0 + \bar{a}^1 \bar{b}^1 + \bar{a}^2 \bar{b}^2 + \bar{a}^3 \bar{b}^3 = -a^0 b^0 + a^1 b^1 + a^2 b^2 + a^3 b^3. \tag{12.29}$$

Just as the ordinary dot product is **invariant** (unchanged) under rotations, this combination is invariant under Lorentz transformations.

To keep track of the minus sign it is convenient to introduce the **covariant** vector a_μ, which differs from the **contravariant** a^μ only in the sign of the zeroth component:

$$a_\mu = (a_0, a_1, a_2, a_3) \equiv (-a^0, a^1, a^2, a^3). \tag{12.30}$$

You must be scrupulously careful about the placement of indices in this business: *upper* indices designate *contravariant* vectors; *lower* indices are for *covariant* vectors. Raising or lowering the temporal index costs a minus sign ($a_0 = -a^0$); raising or lowering a spatial index changes nothing ($a_1 = a^1$, $a_2 = a^2$, $a_3 = a^3$). The scalar product can now be written with the summation symbol,

$$\sum_{\mu=0}^{3} a_\mu b^\mu, \tag{12.31}$$

or, more compactly still,

$$a_\mu b^\mu. \tag{12.32}$$

Summation is *implied* whenever a Greek index is repeated in a product—once as a covariant index and once as contravariant. This is called the **Einstein summation convention**, after its inventor, who regarded it as one of his most important contributions. Of course, we could as well take care of the minus sign by switching to covariant b:

$$a_\mu b^\mu = a^\mu b_\mu = -a^0b^0 + a^1b^1 + a^2b^2 + a^3b^3. \tag{12.33}$$

• **Problem 12.17** Check Eq. 12.29, using Eq. 12.27. [This only proves the invariance of the scalar product for transformations along the x direction. But the scalar product is also invariant under *rotations*, since the first term is not affected at all, and the last three constitute the three-dimensional dot product $\mathbf{a}\cdot\mathbf{b}$. By a suitable rotation, the x direction can be aimed any way you please, so the four-dimensional scalar product is actually invariant under *arbitrary* Lorentz transformations.]

Problem 12.18

(a) Write out the matrix that describes a *Galilean* transformation (Eq. 12.12).

(b) Write out the matrix describing a Lorentz transformation along the y axis.

(c) Find the matrix describing a Lorentz transformation with velocity v along the x axis followed by a Lorentz transformation with velocity $\bar{v}$ along the y axis. Does it matter in what order the transformations are carried out?

Problem 12.19 The parallel between rotations and Lorentz transformations is even more striking if we introduce the **rapidity**:

$$\theta \equiv \tanh^{-1}(v/c). \tag{12.34}$$

(a) Express the Lorentz transformation matrix Λ (Eq. 12.24) in terms of θ, and compare it to the rotation matrix (Eq. 1.29).

In some respects rapidity is a more natural way to describe motion than velocity. [See E. F. Taylor and J. A. Wheeler, *Spacetime Physics* (San Francisco: W. H. Freeman, 1966).] For one thing, it ranges from $-\infty$ to $+\infty$, instead of $-c$ to $+c$. More significantly, rapidities add, whereas velocities do not.

(b) Express the Einstein velocity addition law in terms of rapidity.

(ii) The invariant interval. Suppose event A occurs at $(x_A^0, x_A^1, x_A^2, x_A^3)$, and event B at $(x_B^0, x_B^1, x_B^2, x_B^3)$. The difference,

$$\Delta x^\mu \equiv x_A^\mu - x_B^\mu, \tag{12.35}$$

is the **displacement 4-vector**. The scalar product of Δx^μ with itself is a quantity of special importance; we call it the **interval** between two events:

$$I \equiv (\Delta x)_\mu (\Delta x)^\mu = -(\Delta x^0)^2 + (\Delta x^1)^2 + (\Delta x^2)^2 + (\Delta x^3)^2 = -c^2t^2 + d^2, \tag{12.36}$$

where t is the time difference between the two events and d is their spatial separation. When you transform to a moving system, the *time* between A and B is altered ($\bar{t} \neq t$), and so is the *spatial separation* ($\bar{d} \neq d$), but the interval I remains the same.