Chapter 3

The Lorentz Group

Shortly after Einstein presented his special theory of relativity in 1905, Minkowski interpreted Einstein's theory in terms of the notion of *space-time*, a four-dimensional manifold with an unusual scalar product.[1] From this geometrical way of thinking the general theory of relativity evolved.

Consider a neighborhood of some particular "point" in spacetime. Such a point represents an *event* (e.g., the snapping of my fingers). A neighborhood of such an event would comprise all "points" sufficiently close to the given one in regard to both spatial location and time. Special relativity is just the aspect of general relativity that is concerned with how different observers passing one another describe the neighborhood of that spacetime event (i.e., their passing).

To fix in your mind precisely what is meant by two *local frames of reference*, think of two artificial earth satellites that are in orbits which periodically carry them past one another. We shall assume that the occupants of each space vehicle use its thrusters to stop any tumbling. From the perspective of the inhabitants of each vehicle objects jettisoned from either will appear to travel in straight lines at uniform velocity. It is important that one perform one's observations only in the immediate spatial neighborhood of the location where the earth satellites are when they pass one another, and that all the observations be completed within a short time. Otherwise, the fact that the space ships are in orbit about the earth would become evident. The stronger the gravitational field of the central planet, the more stringent will be the conditions imposed upon the choice of spacetime neighborhood that can be considered.

The relationship between how one observer describes phenomena and

[1]See translation of a talk by Minkowski in *The Principle of Relativity*, Dover, 1923.

how another observer moving with respect to the first describes the same phenomena can best be described in terms of groups. The Lorentz group is that group which leaves the whole Maxwell theory of electromagnetism invariant in form. Fortunately, the group can be identified without looking at the full Maxwell theory at once. One may concentrate upon the fact that the Maxwell theory implies that light waves emitted at a certain point in space at a certain time propagate outward in every direction at speed c as spherical waves. If the Maxwell theory is not going to single out one particular inertial frame, this conclusion must hold in all inertial frames, strange as that may seem to anyone who is unable to shed the prejudices of pre-Einsteinian thinking.

If the equation of an expanding spherical wave has the form

$$(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - (c\Delta t)^2 = 0$$

with respect to one inertial observer, then it must have the form

$$(\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2 - (c\Delta t')^2 = 0$$

with respect to another inertial observer. One is thus led [2] to consider the group $O(3,1)$, which leaves scalar products of the type

$$A_1B_1 + A_2B_2 + A_3B_3 - A_4B_4$$

invariant. $O(3,1)$ differs from $O(4)$ because of the one minus sign.

Vectors at a given "point" in the spacetime manifold can be referred to a basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$. The one pointing in the timelike direction is normalized so that $\mathbf{e}_4 \cdot \mathbf{e}_4 = -1$, and it is made orthogonal to the other three, i.e., $\mathbf{e}_i \cdot \mathbf{e}_4 = 0$ $(i = 1, 2, 3)$. An arbitrary vector $\mathbf{A}$ at the given point of spacetime can be expressed in the form

$$\mathbf{A} = A_1\mathbf{e}_1 + A_2\mathbf{e}_2 + A_3\mathbf{e}_3 - A_4\mathbf{e}_4.$$

This is in agreement with the requirement that

$$\mathbf{A} \cdot \mathbf{B} = A_1B_1 + A_2B_2 + A_3B_3 - A_4B_4.$$

We shall be dealing with such scalar products a lot. It is convenient to introduce the following convention: $A^1 = A_1, A^2 = A_2, A^3 = A_3$, but $A^4 = -A_4$. Then

$$\mathbf{A} \cdot \mathbf{B} = A_1B_1 + A_2B_2 + A_3B_3 - A_4B_4 = A^aB_a,$$

[2] Consideration of light rays alone does not suffice to pin down the group. For that other aspects of the Maxwell theory must be used.

where summation over the repeated index a is understood. Similarly, we may write $\mathbf{A} = A^a \mathbf{e}_a$ and $\mathbf{B} = B^a \mathbf{e}_a$.

The scalar product

$$g_{ab} = \mathbf{e}_a \cdot \mathbf{e}_b = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

is called the *metric tensor*. Note that

$$\mathbf{A} \cdot \mathbf{B} = (A^a \mathbf{e}_a) \cdot (B^b \mathbf{e}_b) = (\mathbf{e}_a \cdot \mathbf{e}_b) A^a B^b = g_{ab} A^a B^b = A^a B_a.$$

The metric tensor is used to lower indices, while the inverse metric tensor, denoted by g^{ab}, is used to raise indices. The Lorentz group $O(3,1)$ corresponds to transformations

$$\mathbf{e}_a \to \mathbf{e}'_a = \Lambda_a{}^b \mathbf{e}_b \text{ (summation convention!)}$$

which preserve the scalar product. The 4×4 matrices $\Lambda_a{}^b$ satisfy the condition

$$g_{ab} = \mathbf{e}'_a \cdot \mathbf{e}'_b = (\Lambda_a{}^c \mathbf{e}_c) \cdot (\Lambda_b{}^d \mathbf{e}_d) = g_{cd} \Lambda_a{}^c \Lambda_b{}^d.$$

In matrix language this equation becomes simply $g = \Lambda g \Lambda^T$. It is the analog of the condition $RR^T = I$ for the $O(n)$ transformation matrices.

Following the procedure which we employed in studying the simpler groups, we write Λ in the form $\Lambda = e^{-iX}$, noting that this time the matrix $g^{-1}X$ (not simply X) is skew-symmetric and pure imaginary. It follows that the matrix X can be expressed as a linear combination with real coefficients of the six pure imaginary matrices enumerated below:

$$J_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$J_2 = \begin{pmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$J_3 = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

and

$$K_1 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix},$$
$$K_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix},$$
$$K_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \end{pmatrix}.$$

The commutation relations satisfied by these matrices are the following:

$$\begin{aligned}
J_2J_3 - J_3J_2 &= iJ_1 \\
J_3J_1 - J_1J_3 &= iJ_2 \\
J_1J_2 - J_2J_1 &= iJ_3 \\
J_2K_3 - K_3J_2 &= iK_1 \\
J_3K_1 - K_1J_3 &= iK_2 \\
J_1K_2 - K_2J_1 &= iK_3 \\
J_1K_1 - K_1J_1 &= 0 \\
J_2K_2 - K_2J_2 &= 0 \\
J_3K_3 - K_3J_3 &= 0 \\
K_2K_3 - K_3K_2 &= -iJ_1 \\
K_3K_1 - K_1K_3 &= -iJ_2 \\
K_1K_2 - K_2K_1 &= -iJ_3
\end{aligned}$$

Notice in particular that the commutation relations satisfied by J_1, J_2, J_3 are identical to the angular momentum commutation relations of quantum theory. This is related to the fact that the rotation group $O(3)$ is a *subgroup* of the Lorentz group.

Of course, any six linearly independent linear combinations of the matrices J_1, J_2, J_3, K_1, K_2 and K_3 will serve as well as a basis for the Lie algebra. In particular, we may employ the Hermitian matrices

$$\begin{aligned}
\Sigma_i &= J_i + iK_i \\
\bar{\Sigma}_i &= J_i - iK_i,
\end{aligned}$$

where $i = 1, 2, 3$. (It should be noted that $\bar{\Sigma}_i$ is *not* the Hermitian conjugate (or adjoint) of Σ_i.)

One advantage of the matrices Σ_i and $\bar{\Sigma}_i$ is that each Σ_i commutes with each $\bar{\Sigma}_j$. Furthermore,

$$\Sigma_1{}^2 = \Sigma_2{}^2 = \Sigma_3{}^2 = \bar{\Sigma}_1{}^2 = \bar{\Sigma}_2{}^2 = \bar{\Sigma}_3{}^2 = I,$$

and

$$\begin{array}{lcl} \Sigma_2\Sigma_3 = i\Sigma_1 & , & \bar{\Sigma}_2\bar{\Sigma}_3 = i\bar{\Sigma}_1 \ , \\ \Sigma_3\Sigma_1 = i\Sigma_2 & , & \bar{\Sigma}_3\bar{\Sigma}_1 = i\bar{\Sigma}_2 \ , \\ \Sigma_1\Sigma_2 = i\Sigma_3 & , & \bar{\Sigma}_1\bar{\Sigma}_2 = i\bar{\Sigma}_3 \ . \end{array}$$

Thus, the matrices $\Sigma_1, \Sigma_2, \Sigma_3$ and $\bar{\Sigma}_1, \bar{\Sigma}_2, \bar{\Sigma}_3$ both have the same algebraic properties as the Pauli spin matrices $\sigma_1, \sigma_2, \sigma_3$, although the former are 4×4 matrices while the latter are 2×2 matrices. Because the barred and unbarred matrices commute with one another, we may write

$$\Lambda = \exp(\frac{i}{2}\bar{\Sigma}_k\xi_k^*)\exp(\frac{i}{2}\Sigma_\ell\xi_\ell),$$

thus factoring the Λ matrices into two factors. Here ξ_1, ξ_2, ξ_3 denote three arbitrary complex parameters, whose complex conjugates are $\xi_1^*, \xi_2^*, \xi_3^*$.

The matrices $\exp(\frac{i}{2}\sigma_k\xi_k)$ themselves constitute a group, which is called the *special linear group*, and which is designated $SL(2, C)$, the C standing for "complex" and the 2 for two-dimensional. The Lorentz group is isomorphic to the direct product of two $SL(2, C)$ groups whose parameters are related by complex conjugation.

Notice that the matrices $\exp(\frac{i}{2}\Sigma_k\xi_k)$ provide a 4×4 representation of the abstract group $SL(2, C)$. On the other hand, when the parameters ξ_k are restricted to real values, then one has a 4×4 representation of the abstract group $SU(2)$. Such real values of the parameters correspond to transformations that are just spatial rotations.

We shall now consider some very special Lorentz transformations, just to get a feeling for their nature. Suppose, for example, that

$$\Lambda = \exp(i\psi K_1),$$

where

$$K_1 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}.$$

It is a simple matter to sum the series which defines the exponential function. The result is that

$$\Lambda = \begin{pmatrix} \cosh\psi & 0 & 0 & \sinh\psi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh\psi & 0 & 0 & \cosh\psi \end{pmatrix}.$$

Ex. 8 *Verify the above result.*

This special Lorentz transformation corresponds to the transformation

$$\begin{aligned} \mathbf{e}_1' &= \cosh\psi\mathbf{e}_1 + \sinh\psi\mathbf{e}_4, \\ \mathbf{e}_2' &= \mathbf{e}_2, \\ \mathbf{e}_3' &= \mathbf{e}_3, \\ \mathbf{e}_4' &= \sinh\psi\mathbf{e}_1 + \cosh\psi\mathbf{e}_4. \end{aligned}$$

When the basic tetrad is transformed in this way, the components of a given vector transform as follows:

$$\begin{aligned} \Delta x' &= \cosh\psi\ \Delta x - \sinh\psi\ c\Delta t, \\ \Delta y' &= \Delta y, \\ \Delta z' &= \Delta z, \\ -c\Delta t' &= \sinh\psi\ \Delta x - \cosh\psi\ c\Delta t. \end{aligned}$$

For definiteness I have used a position vector as an example. The fourth component is equal to the velocity of light times the time interval. (Most relativists adopt units such that $c = 1$.)

Suppose we consider two observers in relative motion. Let one of the observers designate a spacetime event by x, y, z, t, while the other observer designates the same spacetime event by x', y', z', t'. A point at rest in the primed system will have $\Delta x' = \Delta y' = \Delta z' = 0$. The velocity of the primed system relative to the unprimed system will therefore be given by

$$\begin{aligned} v_x &= c\tanh\psi, \\ v_y &= 0, \\ v_z &= 0. \end{aligned}$$

One may eliminate the parameter ψ in favor of v_x and thus obtain

$$\Delta x' = \frac{\Delta x - v_x\Delta t}{\sqrt{1-(v_x/c)^2}},$$

$$\begin{aligned} \Delta y' &= \Delta y, \\ \Delta z' &= \Delta z, \\ \Delta t' &= \frac{\Delta t - v_x \Delta x / c^2}{\sqrt{1 - (v_x/c)^2}}. \end{aligned}$$

These are the transformation formulas for a simple Lorentz transformation in the x-direction. When $|v_x| \ll c$ and $|\Delta x| \ll c|\Delta t|$, the Lorentz transformation reduces to a Galilean transformation. Thus, as long as the two observers are traveling at relative speeds which are small compared to the speed of light, and as long as one considers distances that are small compared to the distance light travels in the time intervals considered, the Galilean transformations suffice to relate how these two observers describe the same events.

Let us now look at the situation when the relative speed of the two observers is no longer insignificant compared to the speed of light. It then becomes important to appreciate the fact that different observers moving with respect to one another will attribute different coordinate labels (x, y, z, t) and (x', y', z', t') to one and the same event.

Fitzgerald contraction Suppose you have caught an especially fast fish which you would like to measure. As the fish swims by, you make marks on the glass wall of the fish tank indicating where the tail is and where the head is. Then you use your standard meter stick to determine the separation between the marks. It is especially important that the marks be made *simultaneously*, for if you should make the tail mark before making the head mark the fish will have had an opportunity to swim forward some distance in the interim, and you will obtain an exaggerated estimate of the length of the fish.

Suppose the fish is the primed observer and you are the unprimed observer. Let us consider as the two events in question the making of the two marks on the glass. You endeavor to make the marks simultaneously; i.e., you make $\Delta t = 0$. But then it is clear from the Lorentz transformation formula that the separation $\Delta x'$ between the marks from the perspective of the fish, i.e., the *proper length* of the fish, will not be the separation Δx which you measured, but rather it will be given by

$$\Delta x' = \frac{\Delta x}{\sqrt{1 - (v_x/c)^2}}.$$

Letting the proper length be denoted by L, we have the Fitzgerald contraction formula

$$\Delta x = L\sqrt{1-(v_x/c)^2}.$$

As the speed of the fish approaches the speed of light, its length as determined by you shrinks toward zero!

Time dilation Now let us consider the relationship between the clocks employed by two observers who are moving relative to one another. If a clock is at rest in the primed system, and we consider as two events two successive ticks of the clock, the primed observer will attribute to these two events no spatial separation; i.e., $\Delta x' = \Delta y' = \Delta z' = 0$. The *proper time* interval which elapses between two successive ticks of the clock will by $\Delta t' = \Delta\tau$ (one second?). What, however, will be the time interval Δt attributed to the very same events by the unprimed observer? The easiest way to evaluate this time interval is to observe that

$$\Delta t = \frac{\Delta t' + v_x \Delta x'/c^2}{\sqrt{1-(v_x/c)^2}},$$

and this implies

$$\Delta t = \frac{\Delta\tau}{\sqrt{1-(v_x/c)^2}}.$$

Since the time interval Δt is longer than $\Delta\tau$, this effect is called *time dilation*. Clocks moving with respect to the observer run more slowly than clocks at rest!

Velocity addition formula Suppose now you consider *three observers*, with relative motions along the x-direction. Call the observers A, B and C, and let the velocity of B relative to A be denoted by v_{AB} while the velocity of C relative to B is denoted by v_{CB}. Now, the velocity of C relative to A is *not* given by $v_{AC} = v_{AB} + v_{BC}$. In fact, it is the parameter ψ which adds when you compound simple Lorentz transformations, and $v = c\tanh\psi$. Thus, the correct velocity addition rule is

$$\begin{aligned} v_{AC} &= c\tanh\psi_{AC} \\ &= c\tanh(\psi_{AB}+\psi_{BC}) \\ &= c\frac{\tanh\psi_{AB}+\tanh\psi_{BC}}{1+\tanh\psi_{AB}\tanh\psi_{BC}} \\ &= \frac{v_{AB}+v_{BC}}{1+v_{AB}v_{BC}/c^2}. \end{aligned}$$

If the magnitudes of v_{AB} and v_{BC} are less than c, then the magnitude of v_{AC} will be less than c too. Moreover, if $v_{BC} = c$, then $v_{AC} = c$ too, regardless of the value of v_{AB}. The speed of light is the same in all reference frames!

There is some very old experimental support for the velocity addition formula. When the speed of light is measured in liquid of index of refraction n and speed v, the result depends upon whether the light is traveling with or against the flow of the liquid. Fizeau's experiment yielded data consistent with the empirical formula

$$u = \frac{c}{n} \pm v(1 - \frac{1}{n^2}).$$

Ex. 9 *Using the relativistic velocity addition formula, establish that when the speed v of the liquid is much less than c the theory does in fact predict the empirical result of Fizeau. (Simply identify $v_{AC} = u, v_{AB} = c/n$ and $v_{BC} = \pm v$ in the relativistic velocity addition formula and then consistently throw away terms that are negligible.)*