Lorentz Transformation of Weyl Spinors

January 11, 2012

William O. Straub, PhD

When Dirac first derived his relativistic electron equation in 1928, he was puzzled by the fact that the solution involved a four-component object (subsequently identified as the Dirac spinor Ψ) rather than a two-component one. The reason for this was Dirac's expectation that the electron, already known at the time to have two spin orientations (up and down), should be completely expressible in terms of 2×2 matrices, not 4×4 matrices. Consequently, the theory required four solutions (ψ_1,ψ_2, ψ_3,ψ_4) for a complete description of the electron. What did the two additional solutions refer to?

Dirac eventually realized that while ψ_1 and ψ_2 do indeed describe spin-up and spin-down electrons, ψ_3 and ψ_4 describe spin-up and spin-down *positrons*, the antimatter equivalent of electrons. The positron (or *antielectron*) was discovered experimentally by Anderson in 1932. Dirac's theory, arguably the greatest achievement of the human mind, has since become *the* cornerstone of relativistic quantum mechanics.

All physics students learn that physical quantities can be described as scalars, vectors and tensors. But Dirac's spinor is a completely different animal whose existence lies somewhere between scalars and vectors. Given the fact that all observable matter in the universe is composed of fermions (quarks and leptons) and their composites, it is disconcerting to learn that they are all spin-1/2 objects that must be described mathematically by the spinor formalism, as least as far as their fundamental properties are concerned. Here we will concentrate on one such property, which is their behavior under Lorentz transformations.

Review of the Lorentz Transformation

Basically, a Lorentz transformation redefines quantities (such as vectors) under rotations in 3-space and "boosts" in four-dimensional spacetime (a boost refers to how measurements of space and time are compared by observers in two different inertial reference frames traveling at a constant velocity with respect to one another). The familiar Lorentz transformation of high school physics texts refers invariably to boosts alone, but the full Lorentz transformation includes rotations as well. For example, the Lorentz boost equations for inertial observers moving in the x^1 direction with relative velocity v are given by

$$\begin{aligned} x^{0\prime} &= \gamma\left(x^0 - \beta x^1\right) \\ x^{1\prime} &= \gamma(x^1 - \beta x^0) \end{aligned}$$

where $x^0 = ct$, $\beta = v/c$ and

$$\gamma = \frac{1}{\sqrt{1-\beta^2}}$$

Given the last definition, it is convenient to relabel the boost parameter by substituting $\tanh\phi = \beta$, giving us the resulting expressions $\cosh\phi = \gamma$ and $\sinh\phi = \beta\gamma$. We then have the familiar matrix expression

$$\begin{bmatrix} x^{0\prime} \\ x^{1\prime} \\ x^{2\prime} \\ x^{3\prime} \end{bmatrix} = \begin{bmatrix} \cosh\phi & -\sinh\phi & 0 & 0 \\ -\sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix}$$

However, since the Lorentz transformation also includes rotations, it is best to understand the transformation as derivable from the *generators* of rotations and well as boosts, which themselves are described as rotations in spacetime. For example, consider the matrix

$$\Omega = \begin{bmatrix} \cosh\phi_1 & -\sinh\phi_1 & 0 & 0 \\ -\sinh\phi_1 & \cosh\phi_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The generator associated with this matrix is defined by

$$K_1 = -i\frac{d\Omega}{d\phi_1}|_{\phi_1=0} = -i\begin{bmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We now define the Lorentz boost transformation with respect to x_1 as the exponential matrix quantity $e^{iK_1\phi_1}$:

$$\begin{aligned} e^{iK_1\phi_1} &= I + iK_1\phi_1 + \frac{1}{2!}(iK_1\phi_1)^2 + \frac{1}{3!}(iK_1\phi_1)^3 + \ldots \\ &= I + iK_1\left[\phi_1 + \frac{1}{3!}\phi_1^3 + \ldots\right] + (iK_1)^2\left[\frac{\phi_1^2}{2!} + \frac{\phi_1^4}{4!}\ldots\right] \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \sinh\phi_1 \begin{bmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\left(I\cosh\phi_i - \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\right) \\ &= \begin{bmatrix} \cosh\phi_1 & -\sinh\phi_1 & 0 & 0 \\ -\sinh\phi_1 & \cosh\phi_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

For boosts in the x^2 and x^3 directions, we have

$$\begin{aligned} K_2 &= -i\frac{d\Omega}{d\phi_2}|_{\phi_2=0} = -i\begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ K_3 &= -i\frac{d\Omega}{d\phi_3}|_{\phi_3=0} = -i\begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

with corresponding results for the transformation matrices, which we will provide shortly.

Similarly, there are generators for ordinary 3-space rotations. For example, a rotation about the x^1 axis is generated by the matrix

$$J_1 = -i\frac{d\Omega}{d\theta_1}|_{\theta_1=0} = -i\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

Expansion of $e^{iJ_1\theta_1}$ then gives the familiar result (this time we skip the details)

$$\begin{aligned} e^{iJ_1\theta_1} &= I + iJ_1\theta_1 + \frac{1}{2!}(iJ_1\theta_1)^2 + \frac{1}{3!}(iJ_1\theta_1)^3 + \ldots \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta_1 & \sin\theta_1 \\ 0 & 0 & -\sin\theta_1 & \cos\theta_1 \end{bmatrix} \end{aligned}$$

while rotations about the other axes are generated by

$$J_2 = -i\frac{d\Omega}{d\theta_2}|_{\theta_2=0} = -i\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$J_3 = -i\frac{d\Omega}{d\theta_3}|_{\theta_3=0} = -i\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Let us now put all this together. The full matrix of Lorentz transformations of boosts and rotations is given by

$$e^{i\boldsymbol{J}\cdot\boldsymbol{\theta}+i\boldsymbol{K}\cdot\boldsymbol{\phi}} \equiv e^{\boldsymbol{\Omega}} = \exp\begin{bmatrix} 0 & -\phi_1 & -\phi_2 & -\phi_3 \\ -\phi_1 & 0 & \theta_3 & -\theta_2 \\ -\phi_2 & -\theta_3 & 0 & \theta_1 \\ -\phi_3 & \theta_2 & -\theta_1 & 0 \end{bmatrix}$$

(Please remember that only one rotation or one boost angle can be applied at any one time – you can't do two or more simultaneously!)

Note that the boost elements of the matrix $\boldsymbol{\Omega} = \boldsymbol{\Omega}^\mu{}_\nu$ (where μ denotes row and ν denotes column) are symmetric while the rotation elements are antisymmetric. It is conventional to deal with this matrix with both indices lowered by contraction with the metric tensor $\eta_{\mu\nu}$. For example,

$$\begin{aligned} \Omega_{10} &= \eta_{1\mu}\Omega^\mu{}_0 = -\Omega^1{}_0 = \phi_1 \quad \text{and} \\ \Omega_{12} &= \eta_{1\mu}\Omega^\mu{}_2 = -\Omega^1{}_2 = -\theta_3 \end{aligned}$$

Thus, contraction of all the elements in $\boldsymbol{\Omega}$ produces the *completely antisymmetric* matrix

$$\Omega_{\mu\nu} = \begin{bmatrix} 0 & -\phi_1 & -\phi_2 & -\phi_3 \\ \phi_1 & 0 & -\theta_3 & \theta_2 \\ \phi_2 & \theta_3 & 0 & -\theta_1 \\ \phi_3 & -\theta_2 & \theta_1 & 0 \end{bmatrix}$$

This result is highly significant, because we will pair this form of the Lorentz matrix with the Dirac matrices to show how Dirac and Weyl spinors behave under Lorentz transformations.

Lorentz Transformation of Spinors

We start with Dirac's electron equation in free space,

$$i\hbar\gamma^\mu\partial_\mu\Psi = mc\Psi$$

where Ψ is the Dirac spinor and the four γ^μ are the Dirac (gamma) matrices. There are several representations of these matrices available, but we will use what is commonly referred to as the Weyl (or chiral) representation:

$$\gamma^0 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \gamma^1 = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$$\gamma^2 = \begin{bmatrix} 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix}, \quad \gamma^3 = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

or, equivalently,

$$\gamma^0 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \gamma^j = \begin{bmatrix} 0 & -\sigma^j \\ \sigma^j & 0 \end{bmatrix}$$

where each entry is a 2×2 matrix – I is the unit matrix and the σ^j are the three Pauli matrices. Consequently, Ψ must be a 4-component column matrix, which we can also express in two-dimensional form via

$$\Psi = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{bmatrix} = \begin{bmatrix} \psi_R \\ \psi_L \end{bmatrix}$$

where the subscripts refer arbitrarily (for the time being) to *right* and *left.*

The Lorentz invariance of the Dirac equation is evident, but it can be viewed in one of two ways. Since the differentiation operator ∂_μ is obviously a covariant vector, we appear to have the choice of taking the set of matrices γ^μ as a contravariant vector, making the combination $\gamma^\mu \partial_\mu$ a scalar, with Ψ itself also a scalar. The other view is to assume that the γ^μ are *fixed* in the forms given above; Lorentz invariance of the Dirac equation would then require that Ψ display some unusual property under Lorentz transformations that counterbalances the transformation of ∂_μ.

In 1955, R.H. Good showed that any choice of Dirac matrices γ^μ can be made equivalent to the fixed matrices by a suitable unitary transformation, meaning that both moving and stationary observers can use the same set of matrices. This necessarily forces us to accept the fact that the quantity Ψ is not a scalar after all, but somehow transforms appropriately under a Lorentz transformation. The job now is to find out how.

Let us first assume that a Dirac spinor changes under a Lorentz transformation by

$$\Psi'(x') = S\Psi(x) \tag{1}$$

where S is a constant 4×4 matrix to be determined. We will also assume that S has an inverse such that

$$\Psi = S^{-1}\Psi'$$

and $S^{-1}S = 1$. The Dirac equation

$$i\hbar\gamma^\mu \partial_\mu \Psi = mc\Psi$$

in the primed system is then

$$i\hbar\gamma^\mu \partial'_\mu \Psi' = mc\Psi'$$

with the Dirac matrices unchanged. Using (1) and $\partial'_\mu = \Omega^\alpha_{\ \mu}\partial_\alpha$ we have

$$\begin{aligned} i\hbar\gamma^\mu \Omega^\alpha_{\ \mu}\partial_\alpha S\Psi &= mcS\Psi \\ &= i\hbar S\gamma^\alpha \partial_\alpha \Psi \end{aligned}$$

Left-multiplication by S^{-1} then gives

$$S^{-1}\gamma^\mu S\Omega^\alpha_{\ \mu} = \gamma^\alpha$$

where we have used the constancy of S to write $\partial_\alpha S = S\partial_\alpha$. Bringing the Lorentz factor over to the right then leaves

$$S^{-1}\gamma^\alpha S = \Omega^\alpha_{\ \beta}\gamma^\beta \tag{2}$$

Up to this point we have no idea what the operator S should look like. But from (2) it should be obvious that it must contain the components of the Lorentz matrix $\Omega^\alpha_{\ \mu}$ along with the Dirac matrices, so we venture to write

$$S \sim \Omega^{\alpha}_{\ \nu}\gamma_\alpha\gamma^\nu = \Omega_{\mu\nu}\gamma^\mu\gamma^\nu \tag{3}$$

To get a definitive form for S, we use the standard trick of taking infinitesimal approximations for S and $\exp\Omega$. Thus,

$$\begin{aligned} S &= 1 + k\Delta\Omega_{\mu\nu}\gamma^\mu\gamma^\nu, \\ S^{-1} &= 1 - k\Delta\Omega_{\mu\nu}\gamma^\mu\gamma^\nu, \\ e^{\boldsymbol{\omega}} &= \delta^{\mu}_{\ \alpha} + \Delta\Omega^{\mu}_{\ \alpha} \end{aligned}$$

where k is some constant to be determined, $\delta^{\mu}_{\ \alpha}$ is the Kronecker symbol and $\Delta\Omega^{\mu}_{\ \alpha}$ is the Lorentz matrix of infinitesimal boosts and angles. Plugging these expressions into (3), we have

$$(1 - k\Delta\Omega_{\alpha\beta}\gamma^\alpha\gamma^\beta)\gamma^\mu(1 + k\Omega_{\lambda\rho}\gamma^\lambda\gamma^\rho) = (\delta^{\mu}_{\ \alpha} + \Delta\Omega^{\mu}_{\ \alpha})\gamma^\alpha$$

which, after some reduction and relabeling, gives

$$k\Delta\Omega_{\alpha\beta}\gamma^\mu\gamma^\alpha\gamma^\beta - k\Delta\Omega_{\alpha\beta}\gamma^\alpha\gamma^\beta\gamma^\mu = \Delta\Omega^{\mu}_{\ \alpha}\gamma^\alpha$$

We must now move the γ^α matrix on the two terms on the left all the way to the right to match the γ^α term on the right-hand side. This is done by using identities like

$$\gamma^\alpha\gamma^\beta\gamma^\mu = (2\eta^{\alpha\beta} - \gamma^\beta\gamma^\alpha)\gamma^\mu$$

along with the antisymmetry of $\Omega_{\mu\nu}$. After several such operations, we easily find that $4k\Delta\Omega^{\mu}_{\ \alpha}\gamma^\alpha = \Delta\Omega^{\mu}_{\ \alpha}\gamma^\alpha$, so that $k = 1/4$. Thus, the infinitesimal form for S is

$$S = 1 + \frac{1}{4}\Delta\Omega_{\mu\nu}\gamma^\mu\gamma^\nu$$

For any *finite* Lorentz transformation we can build up S using n successive applications of the above expression. Using $\Delta\Omega_{\mu\nu} = \Omega_{\mu\nu}/n$, we have

$$\begin{aligned} S &= \lim_{n\to\infty}(1 + \frac{1}{4}\frac{\Omega_{\mu\nu}}{n}\gamma^\mu\gamma^\nu)^n \quad \text{or} \\ S &= \exp(\frac{1}{4}\Omega_{\mu\nu}\gamma^\mu\gamma^\nu) \end{aligned}$$

which is the correct definition for the finite Lorentz transformation operator. Finally, let us note that the summation of indices in $\Omega_{\mu\nu}\gamma^\mu\gamma^\nu$ includes terms involving $\mu = \nu$ and $\mu \neq \nu$. Using the antisymmetry of $\Omega_{\mu\nu}$ and the anticommutation properties of the Dirac matrices, we see that an apparent *double counting* occurs; this can be eliminated by writing

$$S = \exp(\frac{1}{4}\Omega_{\mu\nu}\gamma^\mu\gamma^\nu) = \exp(\frac{1}{2}\Omega_{\mu\nu}\gamma^\mu\gamma^\nu) \qquad (\text{for } \mu < \nu)$$

Thus, the ubiquitous $1/2$ factor found in spinorial calculations results because of the way the Dirac spinor Ψ changes under a Lorentz transformation. An immediate consequence of this is seen in ordinary rotations – a spinor rotated by 360 degrees does not return to itself, but merely changes sign; a full 720-degree rotation is required to bring the spinor back.

Perhaps the most interesting aspect of the above definition for S (at least in the Weyl representation) is the fact that it is *block diagonal*. This can be seen by a straightforward expansion of the exponential in terms of the Pauli matrices:

$$\begin{aligned} \frac{1}{2}\Omega_{\mu\nu}\gamma^\mu\gamma^\nu &= \frac{1}{2}\left[\Omega_{0k}\gamma^0\gamma^k + \Omega_{12}\gamma^1\gamma^2 + \Omega_{13}\gamma^1\gamma^3 + \Omega_{23}\gamma^2\gamma^3\right] \\ &= \frac{1}{2}\phi_k \begin{bmatrix} \sigma^k & 0 \\ 0 & -\sigma^k \end{bmatrix} + \frac{1}{2}i\theta_k \begin{bmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{bmatrix} \end{aligned}$$

so that

$$S = \exp \begin{bmatrix} \frac{1}{2} i\boldsymbol{\sigma} \cdot \boldsymbol{\theta} - \frac{1}{2}\boldsymbol{\sigma} \cdot \boldsymbol{\phi} & 0 \\ 0 & \frac{1}{2} i\boldsymbol{\sigma} \cdot \boldsymbol{\theta} + \frac{1}{2}\boldsymbol{\sigma} \cdot \boldsymbol{\phi} \end{bmatrix} \tag{4}$$

Note the all-important sign change for the boost angles in the lower right 2×2 matrix of (4) – it makes the lower row *distinct* from the upper row, so we have two *inequivalent* representations of 2-component spinors. To see this, recall the earlier form of the Dirac spinor

$$\Psi = \begin{bmatrix} \psi_R \\ \psi_L \end{bmatrix}$$

where each entry in the column is a two-component spinor. By virtue of (4), these spinors transform differently in accordance with

$$\begin{aligned} \psi'_R &= \exp\left(\frac{1}{2} i\boldsymbol{\sigma} \cdot \boldsymbol{\theta} - \frac{1}{2}\boldsymbol{\sigma} \cdot \boldsymbol{\phi}\right) \psi_R \\ \psi'_L &= \exp\left(\frac{1}{2} i\boldsymbol{\sigma} \cdot \boldsymbol{\theta} + \frac{1}{2}\boldsymbol{\sigma} \cdot \boldsymbol{\phi}\right) \psi_L \end{aligned}$$

The quantities ψ_R and ψ_L are called *Weyl spinors* and, unlike the Dirac spinor, they are *irreducible*; in this sense they are more fundamental than the Dirac spinor. However, it can be shown that Weyl spinors do not preserve parity (that is, they are not invariant with respect to the change $\boldsymbol{x} \to -\boldsymbol{x}$), and for this reason Weyl spinors were assumed to represent *neutrinos*, which are all left-handed (described by ψ_L) while antineutrinos are all right-handed (described by ψ_R). The Dirac spinor, being composed of both spinors, is fully parity-preserving.

The Dirac (or Standard) Representation

The above analysis is based on the Weyl (or chiral) representation of the Dirac matrices. For many applications of quantum field theory, the Dirac (or standard) representation of the matrices is preferred. In this representation we have

$$\bar{\gamma}^0 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

with the other Dirac matrices the same (except for a sign change). This representation necessarily mixes the Weyl spinors under Lorentz transformations, so their distinction is not so noticeable. To see this more clearly, consider the transformation matrix S in non-exponential form (x^1 direction only, no rotations):

$$S = \begin{bmatrix} \cosh \frac{1}{2}\phi_1 - \sigma^1 \sinh \frac{1}{2}\phi & 0 \\ 0 & \cosh \frac{1}{2}\phi_1 + \sigma^1 \sinh \frac{1}{2}\phi \end{bmatrix}$$

The decoupling of the Weyl spinors is of course preserved with this form. However, in the Dirac representation this changes in accordance with the transformation $S \to \bar{S}$, where

$$\begin{aligned} A &= A^{-1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ \gamma^0 &\to \bar{\gamma}^0 = A^{-1}\gamma^0 A \\ \gamma^k &\to \bar{\gamma}^k = A^{-1}\gamma^k A \\ S &\to \bar{S} = A^{-1} S A \end{aligned}$$

so that

$$
\begin{aligned}
\bar{\gamma}^0 &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\
\bar{\gamma}^k &= \begin{bmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{bmatrix} \\
\bar{S} &= \begin{bmatrix} \cosh\frac{1}{2}\phi_1 & -\sigma^1\sinh\frac{1}{2}\phi_1 \\ -\sigma^1\sinh\frac{1}{2}\phi_1 & \cosh\frac{1}{2}\phi_1 \end{bmatrix}
\end{aligned}
$$

which mixes the spinors ψ_R and ψ_L.

Comments

Because of the irreducibility of Weyl spinors, they have been used as the fermion basis of *supersymmetry* theories, where either ψ_R or ψ_L (or both) can be used in the construction. It is a simple matter to construct Lorentz-invariant Lagrangians out of either spinor; for example, the quantity $\psi_L^T(i\sigma^2)\psi_L$ (where the T stands for *transpose*) is a Lorentz-invariant spinor scalar. Supersymmetry is an attempt to provide a unified approach to bosonic and fermionic field theories; one of its predictions is the existence of a host of hypothetical new particles (spin-1/2 photinos, squarks, selectrons, etc.). However, as of this writing the results of high-energy scattering experiments at the Large Hadron Collider appear to rule out the existence of such particles, placing doubt on the validity of supersymmetry.

Reference

Lewis H. Ryder, *Quantum Field Theory,* Cambridge University Press, 2nd Edition, 1996.