1. Introduction to Lorentz Spinors

Spinors are analogous to vectors and tensors, but they transform according to the "covering group" SL$(2, \mathbb{C})$ of the homogeneous Lorentz group rather than the Lorentz group itself. SL$(2, \mathbb{C})$ will be discussed in some detail later—for now we just mention that it acts most directly on two-component, complex vectors, the basic spinors that are analogous to four-vectors for the Lorentz group. The aim of these notes is not particularly to develop the so-called "spinor calculus," which is a way of describing the finite dimensional representations of the Lorentz group, but rather to develop some basic facts about the groups themselves.

After some thought about whether it is really appropriate for students approaching Lorentz spinors for the first time to see the complex Lorentz group, we have decided to include it here. The need to actually deal with complex Lorentz transformations occurs only in rather technical situations, especially involving questions of analyticity or continuation from the Minkowski to the Euclidean domain; but the cost of including them in the discussion is so minimal that there seems no reason to deny the potential benefit of the richer context. It is a recurring theme in symmetry considerations where a given group sits relative to its subgroups, and to the groups that naturally contain it as a subgroup. But the essential results for relativistic physics are indeed those for the real Lorentz group and its "covering group" SL$(2, \mathbb{C})$.

1.1 The Homogeneous Lorentz Group

1.1.a Real Lorentz Group

The real, homogeneous Lorentz group L$(\mathbb{R})$ is the set of $4{\times}4$ real matrices that satisfy the equation

$$\Lambda^{Tr} G \Lambda \; = \; G \,, \qquad\qquad G \; = \; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \tag{1.1}$$

It is easy to check that this does define a group. Our notation is $\mu = 0, 1, 2, 3$ for Minkowski indices (latter part of the Greek alphabet); and the Minkowski metric $g_{\mu\nu} = g^{\mu\nu}$ is given by the matrix elements of G. In tensor notation, we have

$$\Lambda^{\mu}{}_{\mu'} \, \Lambda^{\nu}{}_{\nu'} \, g^{\mu'\nu'} \; = \; g^{\mu\nu} \,, \tag{1.2}$$

where $\Lambda^{\mu}{}_{\nu}$ is the "natural" notation for the matrix elements of Λ, corresponding to a contravariant vector transformation law such as

$$x' = \Lambda x \,, \qquad x'^{\mu} = \Lambda^{\mu}{}_{\nu} \, x^{\nu'} \,, \qquad x \in \mathrm{I\!R}^4 \,; \tag{1.3}$$

and $g^{\mu\nu}$ and $g_{\mu\nu}$ are equally natural notations for the matrix elements of G.*

Since we have mentioned contravariant vectors, recall that the metric symbol $g_{\mu\nu}$ can be used to lower the Minkowski index of a contravariant vector and turn it into a covariant vector (like the gradient $\partial/\partial x^{\mu}$):

$$x_{\mu} = g_{\mu\nu} x^{\nu} \,, \qquad x'_{\mu} = \Lambda_{\mu}{}^{\nu'} \, x_{\nu'} \,, \tag{1.4}$$

where we have used the fairly common, but we think bad notation,

$$\Lambda_{\mu}{}^{\nu} = g_{\mu\mu'} \, g^{\nu\nu'} \, \Lambda^{\mu'}{}_{\nu'} \,, \tag{1.5}$$

for the matrix elements of the contragredient matrix†

$$\Lambda^{-1Tr} = G \, \Lambda \, G \,. \tag{1.6}$$

While it may seem economical, the notation $\Lambda_{\mu}{}^{\nu}$ for the matrix elements of the contragredient matrix encourages the misconception that Λ is a tensor. It is better to write the

* Unless otherwise stated, when we write $x \in \mathrm{I\!R}^4$ for a four-vector, we take it to be contravariant.

† The *contragredient* of a matrix is its inverse transpose (or transpose inverse). It has the special form $G\Lambda G$ here because of Eq. (1.1).

covariant transformation law as

$$x'_\mu = x_\nu \, \Lambda^\nu{}_\mu \,, \tag{1.7}$$

or to make it explicit that Λ^{-1Tr} is a different matrix whose matrix elements are naturally written with down and up indices by giving it a name,

$$\tilde{\Lambda} \equiv \Lambda^{-1Tr} \,, \qquad \tilde{\Lambda}_\mu{}^\nu = (G\Lambda G)_{(\mu\nu)} \,. \tag{1.8}$$

The parenthesis notation for the indices on the r.h.s. in the second equation is handy for some calculations which are easier to follow in matrix than in index notation.

It follows from Eq. (1.1) that

$$\det \Lambda = \pm 1 \,, \tag{1.9}$$

and that

$$\Lambda^0{}_0 \, \Lambda^0{}_0 - \sum_{i=1}^{3} \Lambda^0{}_i \Lambda^0{}_i = 1 \,, \qquad\qquad |\Lambda^0{}_0| \geq 1 \,. \tag{1.10}$$

Thus L(IR) splits into four disjoint classes, named according to

Name	$\det \Lambda$	$\Lambda^0{}_0$
$L^\uparrow_+$ (proper, orthochronous)	$+1$	$\geq \;\; 1$
$L^\uparrow_-$ (improper, orthochronous)	-1	$\geq \;\; 1$
$L^\downarrow_+$ (proper, nonorthochronous)	$+1$	≤ -1
$L^\downarrow_-$ (proper, orthochronous)	-1	≤ -1

The special Lorentz transformations I (identity), P (space inversion), T (time inversion), Y (total inversion), defined by

$$P = G \,, \qquad T = -G \,, \qquad Y = -I \,, \tag{1.11}$$

show that none of the four categories is empty; and in fact the sets of transformations are related as follows:

$$L^\uparrow_- = L^\uparrow_+ \, P = P \, L^\uparrow_+ \,, \qquad L^\downarrow_+ = L^\uparrow_+ \, Y = Y \, L^\uparrow_+ \,, \qquad L^\downarrow_+ = L^\uparrow_+ \, T = T \, L^\uparrow_+ \,. \tag{1.12}$$

Only $L^\uparrow_+$ is a subgroup of L(IR). It can be shown to be continuously connected to the identity, and is called the identity component of L(IR). This property implies that the other three pieces are continuously connected to the matrices P, Y, and T, respectively, and from the defining properties of the four pieces listed in the table, it is clear that all are disconnected from each other.

1.1.b Complex Lorentz Group

The set of complex, 4×4 matrices that satisfy Eq. (1.1) is also a group, called the complex, homogeneous Lorentz group, $\mathrm{L}(\mathbb{C})$.

Its elements also satisfy $\det\Lambda = \pm 1$, but the orthochronicity conditions no longer make sense for all Λ. The group $\mathrm{L}(\mathbb{C})$ has two connected pieces, differing by space or time inversion from each other:

$$\begin{aligned} \mathrm{L}_+(\mathbb{C}) &= \{\Lambda \in \mathrm{L}(\mathbb{C}) : \det\Lambda = +1\}\ , \\ \mathrm{L}_-(\mathbb{C}) &= \{\Lambda \in \mathrm{L}(\mathbb{C}) : \det\Lambda = -1\}\ , \\ &= \mathrm{L}_+(\mathbb{C})\, P = P\, \mathrm{L}_+(\mathbb{C}) = \mathrm{L}_+(\mathbb{C})\, T = T\, \mathrm{L}_+(\mathbb{C}). \end{aligned} \tag{1.13}$$

Of the two pieces, only $\mathrm{L}_+(\mathbb{C})$ is a subgroup; it is the identity component of $\mathrm{L}(\mathbb{C})$. The matrices I and $-I$ are connected in $\mathrm{L}_+(\mathbb{C})$, and P and T are connected in $\mathrm{L}_-(\mathbb{C})$.

The equation

$$\Lambda^\dagger\, G\, \Lambda \;=\; G\,, \tag{1.14}$$

where "†" means *Hermitean conjugation*,[⋆] defines another complex Lorentz group, but it is the wrong one for considerations of analyticity. As defined at the beginning of this section, $\mathrm{L}(\mathbb{C})$ is the *analytic complexification* of $\mathrm{L}_+^\uparrow$,[†] a technical notion which means that $\mathrm{L}_+(\mathbb{C})$ is a complex analytic manifold (of complex dimension six), which contains $\mathrm{L}_+^\uparrow$ as a real analytic submanifold (of real dimension six). $\mathrm{L}_+(\mathbb{C})$ is the only group extension of $\mathrm{L}_+^\uparrow$ with that property.

1.1.c Minkowski Scalar Product

It is known from undergraduate physics that $\mathrm{L}(\mathbb{R})$ is the group of real linear transfor-

⋆ The *Hermitean conjugate* of a matrix is its complex conjugate transpose.

† A classic reference on group manifolds is L. Pontrjagin, *Topological Groups*, Princeton University Press, 1939.

mations of $\mathbb{R}^4$ that leave the Minkowski scalar product invariant:

$$x \cdot x \equiv (x^0)^2 - \mathbf{x} \cdot \mathbf{x} = (\Lambda x) \cdot (\Lambda x) . \tag{1.15}$$

It is also true that $\mathrm{L}(\mathbb{C})$ is the group of linear transformations of $\mathbb{C}^4$ that leave the complex Minkowski scalar product invariant. The complex scalar product is defined by

$$z \cdot z = (z^0)^2 - \mathbf{z} \cdot \mathbf{z} , \tag{1.16}$$

without complex conjugation of one of the complex four-vectors.

1.1.d Rotation Subgroups

The proper three-dimensional group of real rotations, which is the proper orthogonal group $\mathrm{O}_+(3, \mathbb{R})$, is isomorphic⋆ to a subgroup of $\mathrm{L}_+^\uparrow$, namely, the set of real, 4×4 matrices of the form

$$\Lambda = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & & & \\ 0 & & R & \\ 0 & & & \end{array} \right) , \qquad \det R = 1 , \qquad R^{Tr} R = I . \tag{1.17}$$

When it doesn't seem confusing, we may also use the notation R for the 4×4 matrix to which it corresponds.

The subgroup $\mathrm{O}_+(3, \mathbb{R}) \subset \mathrm{L}_+^\uparrow$ is connected.

If we let R be complex, but still orthogonal and with determinant unity, we get the group of proper complex rotations of $\mathbb{C}^3$, called $\mathrm{O}_+(3, \mathbb{C})$, as a connected subgroup of $\mathrm{L}_+(\mathbb{C})$.

⋆ A group *isomorphism* is a one-to-one map of one group onto another which preserves the group multiplication law. A *homomorphism* is a possibly several-to-one map of one group into another that preserves the group law. A homomorphism of a group into another group is also called a *representation* of the group, especially when it is into a group of matrices. A representation is called *faithful* when it is one-to-one. A group *automorphism* is a group isomorphism onto the group itself.

1.2 Matrix Representation of Minkowski Space

The set of complex four-vectors $\mathbb{C}^4$, with the Minkowski metric, can be put in one-to-one linear correspondence with the set of complex 2×2 matrices, since both have the same complex dimension. To make the correspondence explicit, we introduce a complete set of 2×2 matrices, the *Pauli* matrices:

$$\sigma_\mu \;=\; (I, \overrightarrow{\sigma}) , \tag{1.18}$$

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} , \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{1.19}$$

The Pauli Matrices are naturally defined as a covariant vector. An alternative complete set is

$$\tilde{\sigma}_\mu = (I, -\overrightarrow{\sigma}) . \tag{1.20}$$

To every complex four-vector $z \in \mathbb{C}^4$ corresponds the 2×2 matrix

$$Z \;=\; z \cdot \sigma \;\equiv\; z^\mu \sigma_\mu \;=\; z^0 I + \mathbf{z} \cdot \overrightarrow{\sigma} . \tag{1.21}$$

Conversely, to each complex 2×2 matrix Z corresponds the four-vector

$$z_\mu \;=\; 1/2\, \mathrm{Tr}\, (\tilde{\sigma}_\mu Z) . \tag{1.22}$$

These linear correspondences are one-to-one, onto inverses of each other, because of the orthogonality of the Pauli matrices:

$$g_{\mu\nu} \;=\; 1/2\, \mathrm{Tr}\, (\tilde{\sigma}_\mu \sigma_\nu) , \tag{1.23}$$

which is readily derived from the anticommutation rules

$$\{\sigma_i, \, \sigma_j\} \;=\; 2\, \delta ij\, I , \qquad i, j = 1, 2, 3 , \tag{1.24}$$

and the fact that

$$\mathrm{Tr}\, \sigma_i \;=\; 0 . \tag{1.25}$$

While we're at it, let's quote

$$[\sigma_i, \, \sigma_j] \;=\; 2\, i\, \epsilon_{ijk}\, \sigma_k , \tag{1.26}$$

$$\sigma_i \sigma_j \;=\; \delta_{ij}\, I + i\, \epsilon_{ijk}\, \sigma_k\,. \tag{1.27}$$

The real four-vectors $x \in \mathrm{I\!R}^4$ are put in one-to-one correspondence with the *Hermitean*[*] 2×2 matrices:

$$X \;=\; x \cdot \sigma \;=\; x \cdot \sigma^\dagger \;=\; X^\dagger \tag{1.28}$$

because of the Hermitean property

$$\sigma_\mu^\dagger \;=\; \sigma_\mu\,. \tag{1.29}$$

The Minkowski scalar product goes over into the determinant in 2×2 space:

$$z \cdot z \;=\; \det(z \cdot \sigma)\,, \tag{1.30}$$

which is a trivial calculation from the formula

$$z \cdot \sigma \;=\; \begin{pmatrix} z^0 + z^3 & z^1 - i z^2 \\ z^1 + i z^2 & z^0 - z^3 \end{pmatrix}\,. \tag{1.31}$$

Finally, the reader should check some useful identities, good for any a, b, $c \in \mathbb{C}^4$:

$$a \cdot \sigma\, b \cdot \tilde\sigma + b \cdot \sigma\, a \cdot \tilde\sigma \;=\; 2\, a \cdot b\,, \tag{1.32}$$

$$a \cdot \sigma\, a \cdot \tilde\sigma \;=\; a \cdot \tilde\sigma\, a \cdot \sigma \;=\; 2\, a \cdot a\,, \tag{1.33}$$

$$a \cdot \sigma\, b \cdot \tilde\sigma\, c \cdot \sigma \;=\; a \cdot \sigma\, b \cdot c - b \cdot \sigma\, c \cdot a + c \cdot \sigma\, a \cdot b + i\, \epsilon^{\mu\nu\lambda\rho}\, \sigma_\mu\, a_\nu\, b_\lambda\, c_\rho\,, \tag{1.34}$$

where $\epsilon_{0123} = -1$. Note that Eq. (1.34) has another version where σ and $\tilde\sigma$ are interchanged and i is replaced by $-i$, which can be derived by evaluating Eq. (1.34) at Pa, Pb, and Pc, and applying the definitions (1.20) of $\tilde\sigma$ and (1.11) of P, and the pseudoscalar nature of the Levi-Civita alternating symbol:

$$\Lambda^\mu{}_{\mu'}\, \Lambda^\nu{}_{\nu'}\, \Lambda^\lambda{}_{\lambda'}\, \Lambda^\rho{}_{\rho'}\, \epsilon^{\mu'\nu'\lambda'\rho'} \;=\; \det\Lambda\; \epsilon^{\mu\nu\lambda\rho}\,. \tag{1.35}$$

This formula holds, by the way, for any complex 4×4 matrix Λ.

---

[*] A *Hermitean* matrix is equal to its Hermitean conjugate.

1.3 Covering Groups of $L_+^\uparrow$ and $L_+(\mathbb{C})$

The notion of a *universal covering group* is a topological construction in Lie group theory whose details do not concern us here; but since the language is very common, we use it. Generally speaking, there are two mathematical issues:

- The first is that the connected subgroups of the real and complex homogeneous groups are multiply connected, and the covering groups are constructed as simply connected. The covering groups are distinct from the groups they "cover," not isomorphic, but homomorphic to them. In our case the homomorphism from the covering group to the group is two-to-one.
- The second is that the Lie algebra of infinitesimal generators of a Lie group is isomorphic to the Lie algebra of infinitesimal generators of its universal covering group. An equivalent technical statement is that a Lie group is *locally* isomorphic to its universal covering group. The Lie algebra is an essential focus for physics, because the infinitesimal generators are observables, and the algebraic relations they obey are an essential part of their physical meaning in quantum mechanics.

For physicists, a rough but useful characterization is that the Lorentz group has to do with the symmetry of vectors (tensors), while its covering group has to do with that of spinors.

1.3.a Complex Lorentz Transformations

Consider the linear transformations of the space of 2×2 matrices that have the form

$$Z' \;=\; A\,Z\,B^{Tr}\,, \tag{1.36}$$

where A and B are 2×2 matrices. We restrict ourselves to the class of pairs (A, B) that preserve the Minkowski metric; i.e.,

$$\det Z' \;=\; \det Z \tag{1.37}$$

for every Z. It follows that

$$\det A\,\det B \;=\; 1\,. \tag{1.38}$$

Clearly the pair $(cA, B/c)$ gives the same linear transformation of 2×2 matrices as (A, B), if c is any nonzero complex number. Equation (1.38) tells us that we can choose c to arrange

$$\det(cA) \;=\; \det(B/c) \;=\; 1\,; \tag{1.39}$$

for example, $c = (\det A)^{-1/2}$.

In other words, we need only consider A and B that are *unimodular* (determinant unity), if we want the pair (A, B) to preserve the Minkowski metric. From now on, we do that. There is still the freedom of a sign; (A, B) and $(-A, -B)$ give the same metric-preserving transformation; and $\det A = \det(-A) = \det B = \det(-B) = 1$.

The set of these transformations forms a group, with multiplicatioin law

$$(A_1, B_1)\,(A_2, B_2) \;=\; (A_1 A_2, B_1 B_2)\,. \tag{1.40}$$

It is just the direct product group SL(2, $\mathbb{C}$)×SL(2, $\mathbb{C}$), where SL(2, $\mathbb{C}$) denotes the group of 2×2 matrices with unit determinant (special linear group of transformations in 2 complex dimensions).

Since the correspondence we defined between four-vectors and 2×2 matrices is linear, the linear transformation (A, B) of 2×2 matrices induces a linear transformation $\Lambda(A, B)$ of four-vectors. That correspondence is easy to compute; let

$$\begin{aligned} z' \cdot \sigma \;&=\; A\, z \cdot \sigma\, B^{Tr} \\ &= [\Lambda(A, B)\, z] \cdot \sigma\,. \end{aligned} \tag{1.41}$$

This holds for every z, so

$$\sigma_\mu\, \Lambda(A, B)^\mu{}_\nu \;=\; A\, \sigma_\nu\, B^{Tr}\,. \tag{1.42}$$

From the orthogonality relation for Pauli matrices (1.23), it follows that

$$\Lambda(A, B)^\mu{}_\nu \;=\; 1/2\, \mathrm{Tr}\, (\tilde{\sigma}^\mu\, A\, \sigma_\nu\, B^{Tr})\,. \tag{1.43}$$

Because $\Lambda(A, B)$ preserves the Minkowski metric, it must belong to L($\mathbb{C}$). Moreover, it is easy to show from Eq. (1.41) that

$$\Lambda(A_1, B_1)\, \Lambda(A_2, B_2) \;=\; \Lambda(A_1 A_2, B_1 B_2)\,. \tag{1.44}$$

What we have done, then, is to construct a homomorphism of the group SL(2, $\mathbb{C}$)×SL(2, $\mathbb{C}$) into L($\mathbb{C}$). One of our jobs is to find out which Lorentz transformations we have arrived at in this way. We are going to see that we don't get them all. But before we do that, let's find out how many transformations (A, B) correspond to a given Λ. The homomorphism is at least two-to-one, because

$$\Lambda(A, B) \;=\; \Lambda(-A, -B)\,. \tag{1.45}$$

We claim that the homomorphism is not more than two-to-one, by the following argument. Suppose two pairs, (A, B) and (A', B') induce the same Lorentz transformation:

$$A\,\sigma_\mu\,B^{Tr} \;=\; A'\,\sigma_\mu\,B'^{Tr}\,. \tag{1.46}$$

Let us show that $(A', B') = (A, B)$ or $(-A, -B)$. First we write

$$\begin{aligned} \sigma_\mu \;&=\; A^{-1}\,A'\,\sigma_\mu\,B'^{Tr}\,B^{Tr-1} \\ &=\; C\,\sigma_\mu\,D^{Tr}\,, \qquad\qquad C \equiv A^{-1}A'\,, \quad D \equiv B^{-1}B'\,. \end{aligned} \tag{1.47}$$

If we can show that C and D are multiples of I, we are finished, for if

$$C \;=\; \lambda_1\,I\,, \qquad D \;=\; \lambda_2\,I\,, \tag{1.48}$$

then we have $\det C = \det D = 1 = \lambda_1^2 = \lambda_2^2$, so that $\lambda_1 = \pm 1$ and $\lambda_2 = \pm 1$. Also, we cannot have $\lambda_1\lambda_2 = -1$, for that contradicts (1.47). We then have $C = D = \pm I$, which proves that $(A', B') = (A, B)$ or $(-A, -B)$, as claimed.

It remains only to show that C and D are multiples of I. To do that, we prove a lemma.

Lemma. For any 2×2 matrix M,

$$\tilde{\sigma}_\mu\,M\,\sigma^\mu \;=\; 2\,(\mathrm{Tr}\,M)\,I\,. \tag{1.49}$$

Proof. Because the trace is linear, it is sufficient to prove the lemma for each of the four

Pauli matrices. Consider $M = \sigma_0 = I$. Applying the identity (1.32), we get

$$\begin{aligned} \tilde{\sigma}_\mu \, \sigma_0 \, \sigma^\mu &= \tilde{\sigma}_\mu \, \sigma^\mu = \sigma_\mu \, \tilde{\sigma}^\mu \\ &= {}^1\!/_2 \, g^{\mu\nu} \, (\sigma_\mu \, \tilde{\sigma}_\nu + \sigma_\nu \, \tilde{\sigma}_\mu) \\ &= {}^1\!/_2 \, g^{\mu\nu} \, 2 \, g_{\mu\nu} \, I = 4I = 2 \, (\mathrm{Tr} \, \sigma_0) \, I \, . \end{aligned} \tag{1.50}$$

Next, consider $M = \sigma_i$, $i = 1, 2, 3$:

$$\begin{aligned} \tilde{\sigma}_\mu \, \sigma_i \, \sigma^\mu &= \sigma_0 \, \sigma_i \, \sigma^0 + \sigma_j \, \sigma_i \, \sigma_j \\ &= \sigma_i + \delta_{ji} \, \sigma_j + i \, \epsilon_{jik} \, \sigma_k \, \sigma_j \\ &= 2 \, \sigma_i - 2 \, \sigma_i = 0 = 2 \, (\mathrm{Tr} \, \sigma_i) \, I \, , \end{aligned} \tag{1.51}$$

where we used (1.25) and (1.26). ▮

Now apply the lemma to (1.47), after multiplying from the left or right by $\tilde{\sigma}^\mu$ and contracting. For example,

$$\begin{aligned} \tilde{\sigma}^\mu \, \sigma_\mu &= \tilde{\sigma}^\mu \, C \, \sigma_\mu \, D^{Tr} \, , \\ 4 \, I &= 2 \, (\mathrm{Tr} \, C) \, D^{Tr} \, . \end{aligned} \tag{1.52}$$

This shows that D is a multiple of I, and a similar argument works for C.

The reader should check that the lemma above can be used to prove the following refinement of the statements in Eqs. (1.47) and (1.48):

Theorem. Let M and N be any 2×2 matrices that obey

$$\sigma_\mu \, M = N \, \sigma_\mu$$

for all four Pauli matrices. Then $M = N$ and both are the same multiple of the identity.

This theorem contains the statement that the three Pauli matrices σ_i are *irreducible*; namely, any 2×2 matrix that commutes with all three is a multiple of the identity.

So far we have found a two-to-one homomorphism

$$\mathrm{SL}(2, \mathbb{C}) \times \mathrm{SL}(2, \mathbb{C}) \xrightarrow{\Lambda(A,B)} \mathrm{L}(\mathbb{C}) \, .$$

The next claim is that $\Lambda(A, B)$ is in fact a proper Lorentz transformation. We give a heuristic argument; namely, you have to believe that $\mathrm{SL}(2, \mathbb{C})$ is connected. We shall see that explicitly later; but for now we point out that $\mathrm{SL}(2,\mathbb{C})$ is the set of matrices $A = a\cdot\sigma$ with $\det A = a\cdot a = 1$; and that the complex hyperboloid $a\cdot a = 1$ is a connected set in $\mathbb{C}^4$. It follows that $\mathrm{SL}(2,\mathbb{C})\times\mathrm{SL}(2,\mathbb{C})$ is also connected. Note that $\Lambda(I,I) = I$. Since $\Lambda(A,B)$ as defined in (1.43) is a continuous function of A and B,* and any pair (A,B) is continuously connected to (I,I), we see that $\Lambda(A,B)$ is continuously connected to the identity in $\mathrm{L}(\mathbb{C})$. Therefore $\det\Lambda = 1$, for $\det\Lambda$ is a continuous function of Λ, and there is no continuous way to reach group elements with $\det\Lambda = -1$ from the identity.

That shows that $\Lambda(A,B)\in \mathrm{L}_+(\mathbb{C})$. A stronger statement is true: the homomorphism $\Lambda(A,B)$ maps $\mathrm{SL}(2,\mathbb{C})\times\mathrm{SL}(2,\mathbb{C})$ two-to-one *onto* $\mathrm{L}_+(\mathbb{C})$. The proof that *every* $\Lambda\in\mathrm{L}_+(\mathbb{C})$ has the representation $\Lambda = \Lambda(A,B)$ is algebraically nontrivial, and we omit it. One could do it by inverting the equation to find A and B, given Λ, as Joos does in his famous article;* or one could do it by parametrizing $\mathrm{L}_+(\mathbb{C})$ in terms of a product of six elementary transformations in the six $\mu\nu$ planes, and putting that in correspondence with a parametrization of $\mathrm{SL}(2,\mathbb{C})\times\mathrm{SL}(2,\mathbb{C})$.

1.3.b Complex Rotations

Now we want to find out which pairs (A,B) correspond to the most important subgroups of $\mathrm{L}_+(\mathbb{C})$. First, let's consider the proper complex rotations, $\mathrm{O}_+(3,\mathbb{C})$. The answer is easy; $\mathrm{O}_+(3,\mathbb{C})$ is the set of proper Lorentz transformations that leave the $\mu = 0$ component of any four-vector unchanged, so

$$A\,\sigma_0\,B^{Tr} = \sigma_0 \qquad \Longrightarrow \qquad B^{Tr} = A^{-1}\,. \tag{1.53}$$

Thus

$$\mathrm{O}_+(3,\mathbb{C}) = \left\{\Lambda(A,A^{-1Tr}) : A\in\mathrm{SL}(2,\mathbb{C})\right\}\,. \tag{1.54}$$

It follows that $\mathrm{SL}(2,\mathbb{C})$ is two-to-one homomorphic to $\mathrm{O}_+(3,\mathbb{C})$.

1.3.c Real Lorentz Transformations

To find those pairs (A,B) which give real Lorentz transformations, we use the fact

* That is, of their matrix elements.
* H. Joos, *Fortschritte der Physik* **11** (1962) 65.

that real transformations preserve real four-vectors; and hence for any Hermitean X,

$$\begin{aligned}(A\,X\,B^{Tr})^\dagger\; &= B^{Tr\dagger}\,X\,A^\dagger\; = A\,X\,B^{Tr}\,,\\ X\;&=\;A^{-1}B^{Tr\dagger}\,X\,A^\dagger\,B^{Tr-1}\,.\end{aligned} \tag{1.55}$$

In particular, this must hold for $X=\sigma_\mu$; and we have already shown in Section 1.3.c that any $2{\times}2$ matrix that commutes with all four σ_μ's must be a multiple of the identity. More precisely in this case:

$$\begin{aligned}A^{-1}B^{Tr\dagger}\;&=\;B^{-1}A^{\dagger Tr}\;=\;\pm I\,,\\ A\;&=\;\pm B^{Tr\dagger}\;=\;\pm\bar{B}\,.\end{aligned} \tag{1.56}$$

Thus, the pairs that give real Lorentz transformations fall into two classes, those of the form $(A,\bar{A})$ and those of the form $(A,-\bar{A})$. Each class is connected, because SL$(2,\mathbb{C})$ is. The first is connected to the identity (I,I), which is outside the second class; so the two classes are disconnected from each other.

It is correct to guess that these two sets correspond to $\mathrm{L}_+^\uparrow$ and $\mathrm{L}_+^\downarrow$, and we can guess at once that $(A,\bar{A})$ goes with $\mathrm{L}_+^\uparrow$ while $(A,-\bar{A})$ goes with $\mathrm{L}_+^\downarrow$, because the $(A,\bar{A})$ are a subgroup of SL$(2,\mathbb{C})\times$SL$(2,\mathbb{C})$ (isomorphic to SL$(2,\mathbb{C})$ itself), while the $(A,-\bar{A})$ are not, corresponding to the fact that $\mathrm{L}_+^\uparrow$ is a subgroup of $\mathrm{L}_+(\mathbb{C})$, while $\mathrm{L}_+^\downarrow$ is not a group.

To verify that, note that if X represents a positive, timelike vector, then X is nonnegative definite. That is so because X, being Hermitean, has eigenvalues λ_1 and λ_2; and

$$\begin{aligned}\lambda_1+\lambda_2\;&=\;\mathrm{Tr}\,X\;=\;2\,x^0\;\geq 0\,,\\ \lambda_1\,\lambda_2\;&=\;\det X\;=\;x\cdot x\;\geq 0\,,\end{aligned} \tag{1.57}$$

implies that λ_1 and λ_2 are nonnegative, after a short calculation. Therefore $AXA^\dagger$ is nonnegative definite, too; and

$$\mathrm{Tr}\,AXA^\dagger\;=\;2\,\Lambda(A,\bar{A})^0{}_\mu\,x^\mu\;\geq\;0\,, \tag{1.58}$$

while

$$\mathrm{Tr}\,(-AXA^\dagger)\;=\;2\,\Lambda(A,-\bar{A})^0{}_\mu\,x^\mu\;\leq\;0\,. \tag{1.59}$$

This is equivalent to saying that $\Lambda(A,\bar{A})$ is orthochronous, and that $\Lambda(A,-\bar{A})$ is nonorthochronous. Indeed

$$
\begin{aligned}
\Lambda(I,-I) &= Y\,, \qquad \text{(easy to show)} \\
\Lambda(A,-\bar{A}) &= \Lambda(A,\bar{A})\,\Lambda(I,-I) = \Lambda(A,\bar{A})\,Y\,,
\end{aligned}
\tag{1.60}
$$

corresponding to $L_+^{\downarrow} = L_+^{\uparrow} Y$.

We conclude that $\Lambda(A,\bar{A})$ defines a two-to-one homomorphism of $\mathrm{SL}(2,\mathbb{C})$ onto $L_+^{\uparrow}$. That the homomorphism is *onto* follows from the analysis above if we grant that $\Lambda(A,B)$ is *onto* $L_+(\mathbb{C})$. We introduce the notation

$$
\Lambda(A) \;\equiv\; \Lambda(A,\bar{A})\,; \tag{1.61}
$$

(and often write $\Lambda = \Lambda(A)$), corresponding to

$$
(\Lambda x)\cdot\sigma \;=\; A\,x\cdot\sigma\,A^{\dagger}\,. \tag{1.62}
$$

Thus $\Lambda(A)$ defines a two-to-one homomorphism

$$
\mathrm{SL}(2,\mathbb{C}) \xrightarrow{\;\Lambda(A)\;} L_+^{\uparrow}\,.
$$

It is known that $\mathrm{SL}(2,\mathbb{C})$ is the universal covering group of $L_+^{\uparrow}$.

1.3.d Real Rotations

The subgroup of $L_+^{\uparrow}$ that leaves x^0 invariant for all real four-vectors $x \in \mathbb{R}^4$ is $O_+(3,\mathbb{R})$, and it corresponds to those $(A,\bar{A})$ satisfying

$$
A\,\sigma_0\,A^{\dagger} \;=\; \sigma_0\,, \qquad A\,A^{\dagger} \;=\; I\,. \tag{1.63}
$$

That is, A is unitary and unimodular, $A = U \in \mathrm{SU}_2$. The subgroup SU_2 of $\mathrm{SL}(2,\mathbb{C})$ is thus two-to-one homomorphic to $O_+(3,\mathbb{R})$, and it is known to be its covering group.

1.3.e Contragredient Real Lorentz Transformations

From now on the focus will be on real Lorentz transformations; and where confusion seems unlikely, we write simply $\Lambda = \Lambda(A)$. In terms of $\mathrm{SL}(2,\mathbb{C})$, the (real) Lorentz transformation reads

$$(\Lambda x)\cdot\sigma \;=\; A\,x\cdot\sigma\,A^\dagger\,; \tag{1.64}$$

and for (real) rotations, we have

$$(R\mathbf{x})\cdot\overrightarrow{\sigma} \;=\; U\,\mathbf{x}\cdot\overrightarrow{\sigma}\,U^\dagger\,, \qquad U\in \mathrm{SU}_2\,. \tag{1.65}$$

Most of the discussion so far has been based on the correspondence $x\leftrightarrow x\cdot\sigma$. It should be realized that we could have developed the theory via the alternative correspondence $x\leftrightarrow x\cdot\tilde\sigma$. We should then have gotten another homomorphism of $\mathrm{SL}(2,\mathbb{C})$ onto $\mathrm{L}_+^\uparrow$—let's call it $\tilde\Lambda(A)$:*

$$[\tilde\Lambda(A)\,x]\cdot\tilde\sigma \;=\; A\,x\cdot\tilde\sigma\,A^\dagger\,. \tag{1.66}$$

To find how Λ and $\tilde\Lambda$ are related to each other, recall (1.33) that

$$(x\cdot\tilde\sigma)^{-1} \;=\; \frac{x\cdot\sigma}{x\cdot x}\,. \tag{1.67}$$

Taking the inverse of both sides of (1.64), we get

$$[\tilde\Lambda(A)\,x]\cdot\sigma \;=\; A^{\dagger-1}\,x\cdot\sigma\,A^{-1} \;=\; [\Lambda(A^{\dagger-1})\,x]\cdot\sigma\,. \tag{1.68}$$

Hence

$$\tilde\Lambda(A) \;=\; \Lambda(A^{\dagger-1})\,. \tag{1.69}$$

We can show besides that $\tilde\Lambda = G\Lambda G$, because from (1.66) we can write

$$[G\,\tilde\Lambda(A)\,x]\cdot\sigma \;=\; A\,(Gx)\cdot\sigma\,A^\dagger\,; \tag{1.70}$$

and replacing x by Gx, we get

$$G\,\tilde\Lambda(A)\,G \;=\Lambda(A)\,. \tag{1.71}$$

The transformation law for $x\cdot\tilde\sigma$ becomes

$$(\Lambda x)\cdot\tilde\sigma \;=\; A^{\dagger-1}\,x\cdot\tilde\sigma\,A^{-1}\,, \tag{1.72}$$

where $\Lambda=\Lambda(A)$. When put this together with (1.62), it is evident that combinations such

* Although $\tilde\Lambda$ turns out to be the same as defined in Section 1.1.a, the two definitions should be treated as independent for the moment.

as $x\cdot\sigma\, y\cdot\tilde\sigma$ and $x\cdot\sigma\, y\cdot\tilde\sigma\, z\cdot\sigma$, etc., are covariant if the correct SL(2, ℂ) transformations are supplied; e.g.,

$$\begin{aligned}
\Lambda x\cdot\sigma\,\Lambda y\cdot\tilde\sigma &= A\, x\cdot\sigma\, y\cdot\tilde\sigma\, A^{-1}\,,\\
\Lambda x\cdot\sigma\,\Lambda y\cdot\tilde\sigma\,\Lambda z\cdot\sigma &= A\, x\cdot\sigma\, y\cdot\tilde\sigma\, z\cdot\sigma\, A^{\dagger}\,,\\
\Lambda x\cdot\tilde\sigma\,\Lambda y\cdot\sigma &= A^{\dagger -1}\, x\cdot\tilde\sigma\, y\cdot\sigma\, A^{\dagger}\,,\\
\Lambda x\cdot\tilde\sigma\,\Lambda y\cdot\sigma\,\Lambda z\cdot\tilde\sigma &= A^{\dagger -1}\, x\cdot\tilde\sigma\, y\cdot\sigma\, z\cdot\tilde\sigma\, A^{-1}\,.
\end{aligned}\tag{1.73}$$

Although the representations Λ and $\tilde\Lambda$ of $\mathrm{L}_+^\uparrow$ are similar to each other via G, the corresponding respresentations A and $A^{\dagger -1}$ of SL(2, ℂ) are not similar. That is, there is no 2×2 matrix M such that $A^{\dagger -1} = MAM^{-1}$ for all $A \in \mathrm{SL}(2,\mathbb{C})$. For if there were, we would have

$$M = A^\dagger\, M\, A\,;$$

and putting $M = m\cdot\sigma$ means the four-vector m obeys

$$m = \Lambda(A^\dagger)\, m$$

for all $A^\dagger$, and hence for all $\Lambda \in \mathrm{L}_+^\uparrow$. It is a basic fact that the only four-vector with that property is $m = 0$,* so $M = 0$, which is a contradiction.

Thus the mapping $A \mapsto A^{\dagger -1}$ is an automorphism of SL(2, ℂ) which cannot be realized by a similarity transformation. The inequivalence disappears if one restricts to the SU_2 subgroup, where the automorphism becomes trivial.

1.3.f Contragredient SL(2, ℂ) Transformations

We have seen that the automorphism $A \mapsto A^{\dagger -1}$ corresponds to $\Lambda \mapsto \tilde\Lambda$, where $\tilde\Lambda = \Lambda^{-1Tr}$ is the contragredient of Λ because of the defining properties of the Lorentz group. There is also an automorphism $A \mapsto A^{-1Tr}$ between elements of SL(2, ℂ) and their contragredients, which plays a special role in the theory of spinors.

* This fact is related to the irreducibility of the self representation of $\mathrm{L}_+^\uparrow$.

LEMMA. Let

$$\varepsilon \equiv i\,\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} . \tag{1.74}$$

Then

$$\varepsilon\,\sigma_\mu\,\varepsilon^{-1} = \tilde{\sigma}_\mu^{Tr} . \tag{1.75}$$

PROOF. Write down the four equations explicitly, and use the properties of the Pauli matrices (1.20) and (1.24). ▌

THEOREM. Let M by any complex 2×2 matrix. Then

$$\varepsilon\,M\,\varepsilon^{-1} = M^{-1Tr}\,\det M . \tag{1.76}$$

PROOF. Let $M = m\cdot\sigma$. Then from the lemma above,

$$\begin{aligned} \varepsilon\,M\,\varepsilon^{-1} &= m\cdot\tilde{\sigma}^{Tr} \\ &= [(m\cdot\sigma)^{-1}\,m\cdot m]^{Tr} \\ &= M^{-1Tr}\,\det M , \end{aligned} \tag{1.77}$$

where we have used (1.33) and (1.30). ▌

Even when $\det M = 0$, it is a fact that the statement in the theorem remains both well defined and true.*

This theorem gives another perspective on the automorphism of SL(2, $\mathbb{C}$) discussed in Section 1.3.f. Namely, although $A^{\dagger-1}$ is not similar to the A, it is similar to the complex conjugate $\bar{A}$:

$$\varepsilon\,A^{\dagger-1}\,\varepsilon^{-1} = \bar{A} . \tag{1.78}$$

In fact it is unitary equivalent, because ε is unitary (easy to check). The mapping $A \mapsto \bar{A}$ is also an automorphism of SL(2, $\mathbb{C}$), so the $\bar{A}$'s are another representation of SL(2, $\mathbb{C}$) which is not similar to the self representation, because it is similar to $A^{\dagger-1}$. It is known that up to unitary equivalence the self representation and the complex conjugate representation are the only two automorphisms of SL(2, $\mathbb{C}$).

* $M^{-1}\det M$ is the adjoint of M, in other words, the transpose of the matrix of minors.

1.3.g Discrete Transformations

The elements P, T, and Y of $\mathrm{L}(\mathbb{R})$ form a discrete subgroup, each member of which is its own inverse. They can also be applied as similarity transformations on elements of $\mathrm{L}_+^\uparrow$:

$$\begin{aligned} P\,\Lambda\,P^{-1} &= T\,\Lambda\,T^{-1} = \tilde{\Lambda}\,, \\ Y\,\Lambda\,Y^{-1} &= \Lambda\,, \end{aligned} \tag{1.79}$$

which in fact map $\mathrm{L}_+^\uparrow$ onto itself in a one-to-one fashion and preserve the group multiplication law. In other words, they induce automorphisms of $\mathrm{L}_+^\uparrow$, with P and T both giving the same thing, the contragredient (because $P = -T = G$), and Y giving the identity automorphism.

On the other hand, we saw a correspondence in Section 1.3.f between the automorphisms $A \mapsto A^{\dagger-1}$ of $\mathrm{SL}(2,\mathbb{C})$ and $\Lambda \mapsto \tilde{\Lambda}$ of $\mathrm{L}_+^\uparrow$. We should ask what this has to do with extending the homomorphism between $\mathrm{SL}(2,\mathbb{C})$ and $\mathrm{L}_+^\uparrow$ to include the discrete symmetries; in other words, what is the spinor representation of P, T, and Y?

The answer is not as simple as one might hope. The results in Section 1.3.f can also be used to show that there is no 2×2 matrix representation of P or T. If there were such a matrix M_P for P, for example, it would have to induce a similarity transformation $M_P A M_P^{-1} = A^{\dagger-1}$; and we have found that to be impossible. The situation changes when one considers the four-dimensional Dirac representation of $\mathrm{SL}(2,\mathbb{C})$, which extends naturally to include the discrete transformations as matrices.