LORENTZ TRANSFORMATIONS, ROTATIONS, AND BOOSTS

ARTHUR JAFFE

November 23, 2013

ABSTRACT. In these notes we study rotations in $\mathbb{R}^3$ and Lorentz transformations in $\mathbb{R}^4$. First we analyze the full group of Lorentz transformations and its four distinct, connected components. Then we focus on one subgroup, the restricted Lorentz transformations. This group contains the proper rotations of $\mathbb{R}^3$, and also the group of proper, orthochronous Lorentz transformations of $\mathbb{R}^4$. We investigate the correspondence between the space-time symmetries of the restricted Lorentz transformations acting on $\mathbb{R}^4$, on the one hand, and the group of 2×2 complex matrices with determinant one. This both gives insight into the structure of Lorentz transformations, and also into charge conjugation. The latter describes symmetry between spin-$\frac{1}{2}$ particles (like electrons, protons, neutrons, and quarks) and their anti-particles.

CONTENTS

I. Lorentz Transformations

I.1. **Points in $\mathbb{R}^4$ and Lorentz Transformations $\boldsymbol{\mathcal{L}}$.** Let us designate a point in $\mathbb{R}^4$ by the 4-vector with real coordinates $x = (x_0, x_1, x_2, x_3) = (ct, \vec{x}) \in \mathbb{R}^4$. In other words, we use units for which the four components of x all have dimension length, and we say that the four-vector x comprises a time component and a 3-vector spatial part. A homogeneous Lorentz transformation Λ is a 4×4 real matrix that acts on $x \in \mathbb{R}^4$ that preserves the Minkowski length $x_M^2 = x_0^2 - x_1^2 - x_2^2 - x_3^2$ of every 4-vector x. Let $\mathcal{L}$ denote the set of all such Lorentz transformation matrices.

More explicitly, let us denote a Lorentz transformation $x \mapsto x'$ by

$$x' = \Lambda x \ , \qquad \text{with} \quad x'_\mu = \sum_{\nu=0}^{3} \Lambda_{\mu\nu} x_\nu \ ,$$

with the property $x'^2_M = x^2_M$. The Minkowski square can be written in terms of the Minkowski-space metric

$$g = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} .$$

Denote by $\langle \ , \ \rangle$ the standard scalar product on Euclidean space, $\langle x, y\rangle = \sum_{\mu=0}^{3} x_\mu y_\mu$. We take the Minkowski scalar product to equal

$$\langle x, y\rangle_M = \langle x, gy\rangle = \sum_{\mu,\nu=0}^{3} x_\mu \, g_{\mu\nu} \, y_\nu = x_0 y_0 - \vec{x} \cdot \vec{y} \ . \tag{I.1}$$

Thus the Lorentz transformation Λ satisfies the relation

$$\langle \Lambda x, g\Lambda x\rangle = \langle x, gx\rangle \ , \ \text{ for all } \ x \ ,$$

or the matrix relation

$$\Lambda^{\mathrm{tr}} g \Lambda = g \ , \tag{I.2}$$

where Λ^{tr} is the transpose of the matrix Λ. The matrix Λ has 16 entries Λ_{ij}. There are 10 independent equations arising from (I.2), which is an equation for a symmetric matrix. Thus there are

$$6 = 16 - 10 \text{ independent real parameters} \tag{I.3}$$

that describe the possible matrices Λ.

A *multiplicative group* G is a set of elements that has three properties:

- There is an associative multiplication: $g_1, g_2 \in G$ ensures $g_1 g_2 \in G$, with $(g_1 g_2) g_3 = g_1 (g_2 g_3)$.
- G contains an identity (sometimes denoted Id, e, 1, I, or $\mathbb{1}$). If $g \in G$, then $g \,\mathrm{Id} = \mathrm{Id}\, g = g$.
- Every element $g \in G$ has an inverse g^{-1} in g such that $gg^{-1} = \mathrm{Id}$.

Proposition I.1. *The set Lorentz transform $\mathcal{L}$ form a multiplicative group. Every $\Lambda \in \mathcal{L}$ has* $\det \Lambda = \pm 1$. *The transpose Λ^{tr} of any $\Lambda \in \mathcal{L}$ is a Lorentz transformation.*

Proof. It is clear from (I.2) that if $\Lambda_1, \Lambda_2 \in \mathcal{L}$, then

$$(\Lambda_1\Lambda_2)^{\rm tr}\, g\, \Lambda_1\Lambda_2 = \Lambda_2^{\rm tr}(\Lambda_1^{\rm tr}\, g\, \Lambda_1)\Lambda_2 = \Lambda_2^{\rm tr}\, g\, \Lambda_2 = g\ . \tag{I.4}$$

Hence the product of two Lorentz transformations is another Lorentz transformation. Furthermore taking the determinant of (I.2), and using $\det(AB) = (\det A)(\det B)$, along with $\det g = -1$, shows that $(\det \Lambda^{\rm tr})(\det \Lambda) = 1$. But $\det \Lambda = \det \Lambda^{\rm tr}$, so $\det \Lambda = \pm 1$. Hence every Lorentz transformation matrix Λ has an inverse matrix Λ^{-1}.

As Λ preserves x_M^2, so does Λ^{-1}. We can also verify this fact algebraically, by using $(\Lambda^{\rm tr})^{-1} = (\Lambda^{-1})^{\rm tr}$, and observing,

$$g = \left(\Lambda^{-1}\right)^{\rm tr} \Lambda^{\rm tr}\, g\, \Lambda\Lambda^{-1} = \left(\Lambda^{-1}\right)^{\rm tr}\, g\, \Lambda^{-1}\ . \tag{I.5}$$

This is the identity of the form (I.2) that Λ^{-1} is a Lorentz transformation.

Also note that the identity matrix is a Lorentz transformation. So the Lorentz transformations form a multiplicative group. Finally the inverse of (I.2) ensures $\Lambda^{-1} g\, (\Lambda^{\rm tr})^{-1} = g$, or $g = \Lambda\, g\, \Lambda^{\rm tr}$, which shows that if Λ is a Lorentz transformation, then $\Lambda^{\rm tr}$ is a Lorentz transformation. □

I.2. **Components of $\mathcal{L}$.** We classify the matrices $\Lambda \in \mathcal{L}$, the set of all Lorentz transformations, into four distinct *connected components*.[1] We will show that there are four connected components of $\mathcal{L}$ that are determined by the sign of $\det \Lambda$ and the sign of Λ_{00}. Those transformations $\Lambda \in \mathcal{L}$ with $\det \Lambda = 1$ are caller *proper Lorentz transformations*, and one denotes the set of such transformations as $\mathcal{L}_+$. Those transformations $\Lambda \in \mathcal{L}$ with $\Lambda_{00} \geqslant 0$ are called *isochronous* Lorentz transformations, and one denotes the set of such transformations as $\mathcal{L}^{\uparrow}$.

Proposition I.2. *The Lorentz transformations $\mathcal{L}$ fall into four disconnected, disjoint components according to the sign of* $\det \Lambda = \pm 1$*, and the sign of* Λ_{00} *for which* $|\Lambda_{00}| \geqslant 1$.

Proof. We have seen in the proof of Proposition I.1 that $\det \Lambda = \pm 1$. As $\det \Lambda$ is a polynomial in the matrix elements Λ_{ij}, it depends continuously on these matrix elements. Hence $\mathcal{L}$ has disconnected components according to the sign of $\det \Lambda$.

The relation (I.2) also shows that $|\Lambda_{00}| \geqslant 1$. In fact the 00 matrix element of the identity (I.2) requires that

$$\Lambda_{00}^2 - \sum_{i=1}^{3} \Lambda_{i0}^2 = 1\ , \ \text{ namely }\ |\Lambda_{00}| = \left(1 + \sum_{i=1}^{3} \Lambda_{i0}^2\right)^{1/2} \geqslant 1\ . \tag{I.6}$$

In Proposition I.1 we saw that if $\Lambda' \in \mathcal{L}$, then $\Lambda'^{\,\rm tr} \in \mathcal{L}$. Thus from (I.6) we infer

$$\left|\Lambda_{00}'^{\,2}\right| = \left(1 + \sum_{j=1}^{3} \Lambda_{j0}'^{\,2}\right) \geqslant 1\ . \tag{I.7}$$

[1] One says that a set $\mathcal{L}_X \subset \mathcal{L}$ of Lorentz transformations is a connected component, if one can find a continuous trajectory of matrices between any two given $\Lambda_1, \Lambda_2 \in \mathcal{L}_X$. This means that for a parameter $0 \leqslant s \leqslant 1$, there is a family of Lorentz transformations $\Lambda(s) \in \mathcal{L}_X$, with matrix elements $\Lambda_{ij}(s)$, such that $\Lambda(s=0) = \Lambda_1$ and $\Lambda(s=1) = \Lambda_2$.

We now see that $\operatorname{sgn}((\Lambda\Lambda')_{00}) = \operatorname{sgn}\Lambda_{00} \operatorname{sgn}\Lambda'_{00}$. In fact matrix multiplication yields

$$(\Lambda\Lambda')_{00} = \Lambda_{00}\,\Lambda'_{00} + \sum_{j=1}^{3} \Lambda_{0j}\Lambda'_{j0}\ . \tag{I.8}$$

Using the Schwarz inequality

$$\left|\sum_{j=1}^{3} \Lambda_{0j}\Lambda'_{j0}\right| \leq \left(\sum_{i=1}^{3} \Lambda_{0i}^2\right)^{1/2} \left(\sum_{j=1}^{3} \Lambda'_{j0}\right)^{1/2}\ . \tag{I.9}$$

If Λ_{00} and Λ'_{00} have the same sign, the first term on the right of (I.8) is positive, and the bound in (I.9) yields (always for positive square roots),

$$(\Lambda\Lambda')_{00} \geqslant \left(1+\sum_{i=1}^{3} \Lambda_{0i}^2\right)^{1/2} \left(1+\sum_{j=1}^{3} \Lambda'^2_{j0}\right)^{1/2} - \left(\sum_{i=1}^{3} \Lambda_{0i}^2\right)^{1/2} \left(\sum_{j=1}^{3} \Lambda'_{j0}\right)^{1/2} \geqslant 0\ . \tag{I.10}$$

If Λ_{00} and Λ'_{00} have the opposite sign, one has similarly (with positive square roots)

$$(\Lambda\Lambda')_{00} \leqslant -\left(1+\sum_{i=1}^{3} \Lambda_{0i}^2\right)^{1/2} \left(1+\sum_{j=1}^{3} \Lambda'^2_{j0}\right)^{1/2} + \left(\sum_{i=1}^{3} \Lambda_{0i}^2\right)^{1/2} \left(\sum_{j=1}^{3} \Lambda'_{j0}\right)^{1/2} \leqslant 0\ . \tag{I.11}$$

As $|(\Lambda\Lambda')_{00}| \geqslant 1$ for any $\Lambda\Lambda' \in \mathcal{L}$, we infer that the components determined by Λ_{00} are disjoint. □

I.3. **The Components of $\mathcal{L}$.** We denote the four components of $\mathcal{L}$ by $\mathcal{L}_+^\uparrow$, $\mathcal{L}_-^\uparrow$, $\mathcal{L}_+^\downarrow$, and $\mathcal{L}_-^\downarrow$. Here the subscript $\pm$ denotes the sign of the determinant, and the superscript arrows denote the sign of Λ_{00} (arrow of time). The condition $\det\Lambda = \pm 1$ divides $\mathcal{L}$ into two disconnected components $\mathcal{L}_\pm$, called *proper* and *improper* Lorentz transformations. The condition $\operatorname{sgn}(\Lambda_{00}) = \pm 1$ divides $\mathcal{L}$ into two disconnected components, which one denotes $\mathcal{L}^\uparrow$ and $\mathcal{L}^\downarrow$, and calls *orthochronus* and *non-orthochronous*.

The fundamental Lorentz transformations which we study are the *restricted Lorentz group* $\mathcal{L}_+^\uparrow$. These are the Lorentz transformations that are both proper, $\det\Lambda = +1$, and orthochronous, $\Lambda_{00} \geqslant 1$. There are some elementary transformations in $\mathcal{L}$ that map one component into another, and which have special names:

- The parity transformation $P : (x_0, \vec{x}) \mapsto (x_0, -\vec{x})$.
- The time-reversal transformation $T : (x_0, \vec{x}) \mapsto (-x_0, \vec{x})$.
- The space-time-inversion transformation $PT : (x_0, \vec{x}) \mapsto (-x_0, -\vec{x})$.

The other three components of $\mathcal{L}$ arise from applying to $\mathcal{L}_+^\uparrow$ the discrete element listed in the third column of Table 1,

$$\mathcal{L}_-^\uparrow = P\mathcal{L}_+^\uparrow\ , \quad \mathcal{L}_-^\downarrow = T\mathcal{L}_+^\uparrow\ , \quad \mathcal{L}_+^\downarrow = PT\mathcal{L}_+^\uparrow\ . \tag{I.12}$$

The subset $\mathcal{L}^\uparrow = \mathcal{L}_+^\uparrow \cup \mathcal{L}_-^\uparrow$ is a group, like the subset $\mathcal{L}_+ = \mathcal{L}_+^\uparrow \cup \mathcal{L}_+^\downarrow$. Another subgroup of Lorentz transformations consists of Lorentz transformation matrices for which $\Lambda_{00}\det\Lambda \geqslant 1$ which is $\mathcal{L}_0 = \mathcal{L}_+^\uparrow \cup \mathcal{L}_-^\downarrow$. But the components $\mathcal{L}_-^\uparrow$ or $\mathcal{L}_+^\downarrow$, as well as the subsets $\mathcal{L}^\downarrow$ or $\mathcal{L}_-$ are not closed under multiplication, so they do not by themselves constitute groups.

Table 1. Components and Subgroups of the Lorentz Group $\mathcal{L}$

$\det \Lambda$	$\operatorname{sgn} \Lambda_{00}$	Discrete Element	Components of $\mathcal{L}$	Symbol
+1	+1	$I : x \mapsto x$	proper, isochronous	$\mathcal{L}_+^\uparrow$
-1	+1	$P : x \mapsto (x_0, -\vec{x})$	space (parity) inverting	$\mathcal{L}_-^\uparrow$
-1	-1	$T : x \mapsto (-x_0, \vec{x})$	time inversion	$\mathcal{L}_-^\downarrow$
+1	-1	$PT : x \mapsto -x$	proper, time reversing	$\mathcal{L}_+^\downarrow$
			Subgroups of $\mathcal{L}$ with entire components	
± 1	± 1		full Lorentz group	$\mathcal{L}$
+1	+1		restricted group (proper, isochronous group)	$\mathcal{L}_+^\uparrow$
± 1	+1		isochronous group	$\mathcal{L}^\uparrow$
+1	± 1		proper group	$\mathcal{L}_+$
+1 -1	+1 -1		another subgroup	$\mathcal{L}_0 = \mathcal{L}_+^\uparrow \cup \mathcal{L}_-^\downarrow$

Table 2. The Multiplication Table for the Components of the Lorentz Group $\mathcal{L}$

	$\mathcal{L}_+^\uparrow$	$\mathcal{L}_-^\uparrow$	$\mathcal{L}_+^\downarrow$	$\mathcal{L}_-^\downarrow$
$\mathcal{L}_+^\uparrow$	$\mathcal{L}_+^\uparrow$	$\mathcal{L}_-^\uparrow$	$\mathcal{L}_+^\downarrow$	$\mathcal{L}_-^\downarrow$
$\mathcal{L}_-^\uparrow$	$\mathcal{L}_-^\uparrow$	$\mathcal{L}_+^\uparrow$	$\mathcal{L}_-^\downarrow$	$\mathcal{L}_+^\downarrow$
$\mathcal{L}_+^\downarrow$	$\mathcal{L}_+^\downarrow$	$\mathcal{L}_+^\uparrow$	$\mathcal{L}_+^\uparrow$	$\mathcal{L}_-^\uparrow$
$\mathcal{L}_-^\downarrow$	$\mathcal{L}_-^\downarrow$	$\mathcal{L}_+^\downarrow$	$\mathcal{L}_-^\uparrow$	$\mathcal{L}_+^\uparrow$

II. Restricted Lorentz Transformations

We have reduced the analysis of a general Lorentz transformation $\Lambda \in \mathcal{L}$ to the analysis of a restricted Lorentz transformation $\Lambda \in \mathcal{L}_+^\uparrow$ with $\det \Lambda = 1$ and $\Lambda_{00} \geqslant 1$, along with the study of the discrete transformations P, T and PT. In this section we analyze the restricted Lorentz transformations in detail. We find that they can be factored into a proper rotation R times a proper, orthochronous boost B.

II.1. **Proper Rotations.** A restricted Lorentz transformation $\Lambda = R$ in $L_+^\uparrow$ is said to be a *proper rotation*, if it leaves the time unchanged, namely $\Lambda_{00} = 1$. By definition $\det R = 1$. Note that the restriction (I.6), along with the fact established in Proposition I.1 that R^{tr} is a Lorentz transformation, means that a pure rotation has the form

$$R = \begin{pmatrix} 1 & 0 \\ 0 & \mathcal{R} \end{pmatrix} . \tag{II.1}$$

We use the letter $\mathcal{R}$ to denote the 3×3 orthogonal matrix with determinant 1 that implements the rotation three-vectors $\vec{x}$. Technically, $\mathcal{R}\in SO(3)$, the group of real, orthogonal, 3×3 matrices with determinant one.

Each pure rotation matrix $\mathcal{R}\in SO(3)$ is specified by an axis, namely a unit vector $\vec{n}$ in 3-space, and an angle θ of rotation about this axis. One writes $\mathcal{R}=\mathcal{R}(\vec{n},\theta)$. The rotation leaves $\vec{n}$ unchanged, and acts in the plane orthogonal to $\vec{n}$. For example, a rotation by angle θ about the third axis $\vec{n}=\vec{e}_3$ is given by the matrix

$$\mathcal{R}(\vec{e}_3,\theta)=\begin{pmatrix}\cos\theta & -\sin\theta & 0\\ \sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{pmatrix}\ , \tag{II.2}$$

or imbedded in a transformation on space-time,

$$R(\vec{e}_3,\theta)=\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & \cos\theta & -\sin\theta & 0\\ 0 & \sin\theta & \cos\theta & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\ . \tag{II.3}$$

Here the sign of the angle θ is a convention: one can interpret the rotation as rotating the coordinate system (one sign) or rotating the space in a fixed coordinate system (the other sign). The former is called a *passive* transformation, while the latter is called an *active* transformation.

II.2. **Pure Lorentz Boost:** A restricted Lorentz transformation $\Lambda\in L_+^\uparrow$ is a *pure boost* in the direction $\vec{n}$ (here $\vec{n}$ is a unit vector in 3-space), if it leaves unchanged any vectors in 3-space in the plane orthogonal to $\vec{n}$. Such a pure boost in the direction $\vec{n}$ depends on one more real parameter $\chi\in\mathbb{R}$ that determines the magnitude of the boost. By choosing the direction of the boost to be $\pm\vec{n}$, we can restrict the parameter χ to be non-negative, $0\leqslant\chi$.

For example the pure Lorentz boost along the first coordinate axis is $\Lambda=B(\vec{e}_1,\chi)$, where $\vec{e}_1$ denotes a unit vector for the Cartesian direction x_1, and with parameter χ. This Lorentz transformation is given by the real, symmetric matrix

$$B(\vec{e}_1,\chi)=\begin{pmatrix}\cosh\chi & \sinh\chi & 0 & 0\\ \sinh\chi & \cosh\chi & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\ . \tag{II.4}$$

Clearly $B=B(\vec{e}_1,\chi)$ satisfies the defining relation $B^{\rm tr}gB=g$ of a Lorentz transformation, and also $\det B=+1$. Furthermore $B_{00}=\cosh\chi\geqslant 1$. Thus $B\in L_+^\uparrow$. One has the transformed point

$$\boxed{x_0'=x_0\cosh\chi+x_1\sinh\chi\ ,\ \text{ and }\ x_1'=x_0\sinh\chi+x_1\cosh\chi\ ,\quad x_2'=x_2\ ,\ \text{ and }\ x_3'=x_3}\ . \tag{II.5}$$

It is interesting to use another parameterization. Define $\beta = \tanh\chi \in [0,1)$, and so using the relation $\cosh^2\chi - \sinh^2\chi = 1$, we see that

$$\cosh\chi = \frac{1}{\sqrt{1-\beta^2}} \;, \quad \text{and} \quad \sinh\chi = \frac{\beta}{\sqrt{1-\beta^2}} \;. \tag{II.6}$$

So in the example,

$$B(\vec{e}_1,\chi) = \begin{pmatrix} \frac{1}{\sqrt{1-\beta^2}} & \frac{\beta}{\sqrt{1-\beta^2}} & 0 & 0 \\ \frac{\beta}{\sqrt{1-\beta^2}} & \frac{1}{\sqrt{1-\beta^2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{II.7}$$

Thus in this example we can write $x' = B(\vec{e}_1,\chi)x$, expressed above as (II.5), in the familiar form

$$\boxed{x_0' = \frac{x_0 + x_1\beta}{\sqrt{1-\beta^2}} \;, \quad \text{and} \quad x_1' = \frac{\beta x_0 + x_1}{\sqrt{1-\beta^2}} \;, \quad x_2' = x_2 \;, \quad \text{and} \quad x_3' = x_3} \;. \tag{II.8}$$

This is a standard form of the pure Lorentz boost along the first axis, where $x_0 = ct$ (to give x_0 and x_1 the same dimension) and $\beta = v/c$ is the dimensionless velocity of the boost. Here the velocity is characterized by $0 \leqslant v < c$. (Of course one could also consider negative v or β if one wished.) Note that ${x_0'}^2 - {x_1'}^2 = x_0^2 - x_1^2$.

II.3. **The Structure of Restricted Lorentz Transformations.** The following describes the structure of a restricted Lorentz transformation in $L_+^\uparrow$. We state these properties now in a proposition. *We show that this proposition is true toward the end of these notes.* Meanwhile we develop some properties relating points on $\mathbb{R}^4$ to hermitian matrices, which we need to analyze Lorentz transformations in a simple way.

Proposition II.1 (**Restricted Lorentz Transformations**). *Every restricted Lorentz transformation $\Lambda \in L_+^\uparrow$ has a unique decomposition as a product of a pure rotation R followed by a pure boost B,*

$$\Lambda = BR \;. \tag{II.9}$$

The rotation has the form (II.1) *with $\mathcal{R} \in SO(3)$; the boost is a symmetric matrix $B \in SO(1,3)_+$.*

Remark II.2. One could write this decomposition of a restricted transformation in the reverse order $\Lambda = R\widetilde{B}$, with a pure boost $\widetilde{B} = R^{-1}BR$ along a different direction. Alternatively there is a pair of pure rotations R_1, R_2, and a pure boost $B(\vec{e}_1,\chi)$ of the form (II.4), such that

$$\Lambda = R_1 B(\vec{e}_1,\chi) R_2 \;. \tag{II.10}$$

This decomposition is also unique, unless the Lorentz transformation Λ is a pure rotation.

III. 2 × 2 Matrices and Points in $\mathbb{R}^4$

The restricted Lorentz transformations have a special relation to the group of 2×2 matrices with determinant $+1$, namely the group $SL(2,\mathbb{C})$. We develop this connection here. The first step is to understand a mapping between $\mathbb{R}^4$ and 2×2 hermitian matrices.

III.1. $\mathbb{R}^4$ **and** $\boldsymbol{H}_2$**.** There is a 1-1 transformation between points in $\mathbb{R}^4$ and the space of 2×2 hermitian matrices that we denote by $\boldsymbol{H}_2$. We consider the map

$$x=(x_0,\vec{x})=(x_0,x_1,x_2,x_3)\quad\longleftrightarrow\quad \widehat{x}=\begin{pmatrix} x_0+x_3 & x_1-ix_2\\ x_1+ix_2 & x_0-x_3\end{pmatrix}\ . \tag{III.1}$$

We use this correspondence in order to analyze the structure of a general Lorentz transformation.

It is natural to define a scalar product on 2×2 matrices as

$$\langle A,B\rangle=\frac{1}{2}\,\mathrm{Tr}(A^*B)\ . \tag{III.2}$$

The trace of a matrix, denoted Tr, is the sum of the diagonal entries. It is convenient to introduce four hermitian matrices σ_μ for $\mu=0,1,2,3$, as a basis for $\boldsymbol{H}_2$. Let

$$\sigma_0=\begin{pmatrix}1&0\\0&1\end{pmatrix}\ ,\quad \sigma_1=\begin{pmatrix}0&1\\1&0\end{pmatrix}\ ,\quad \sigma_2=\begin{pmatrix}0&-i\\i&0\end{pmatrix}\ ,\quad \sigma_3=\begin{pmatrix}1&0\\0&-1\end{pmatrix}\ . \tag{III.3}$$

Note that one can write the 4-vector $\sigma=(I,\vec{\sigma})$. The zero component σ_0 is the identity matrix; the three-vector with components σ_i for $i=1,2,3$ are equal to the 2×2 Pauli matrices. Then $\sigma_1\sigma_2=i\sigma_3=-\sigma_2\sigma_1$, and similarly for cyclic permutation of $1,2,3$. The matrices σ_μ are orthonormal in this scalar product,

$$\langle\sigma_\mu,\sigma_\nu\rangle=\delta_{\mu\nu}\ ,\ \text{ for }\ \mu,\nu=0,1,2,3\ .$$

In other words, the σ_μ are an orthonormal basis for the space of 2×2 hermitian matrices. We summarize the properties of $\widehat{x}$ that follow immediately from this observation:

Proposition III.1. *Consider the transformation $x\mapsto\widehat{x}$ from $\mathbb{R}^4$ to $\boldsymbol{H}_2$ defined by*

$$\widehat{x}=x_0\sigma_0+x_1\sigma_1+x_2\sigma_2+x_3\sigma_3=\begin{pmatrix} x_0+x_3 & x_1-ix_2\\ x_1+ix_2 & x_0-x_3\end{pmatrix}\ . \tag{III.4}$$

The corresponding inverse transformation $\widehat{x}\mapsto x$ from $\boldsymbol{H}_2$ to $\mathbb{R}^4$ is

$$x_\mu=\langle\sigma_\mu,\widehat{x}\rangle=\frac{1}{2}\,\mathrm{Tr}\left(\sigma_\mu\widehat{x}\right)\ . \tag{III.5}$$

Both transformations are linear and 1 to 1.

Proof. The fact that (III.4) and (III.1) agree follows from our choice of σ_μ. The inverse transformation is a consequence of the orthonormal property of the σ_μ's in the scalar product $\langle\ ,\ \rangle$. Finally for $x,y\in\mathbb{R}^4$ and $\lambda\in\mathbb{R}$, the linearity of the correspondence between x and $\widehat{x}$ follows from

$$(\widehat{x+\lambda y})=\widehat{x}+\lambda\widehat{y}\ ,\quad\text{and}\quad (x+\lambda y)_\mu=\left\langle\sigma_\mu,\widehat{x+\lambda y}\right\rangle=\langle\sigma_\mu,\widehat{x}\rangle+\lambda\,\langle\sigma_\mu,\widehat{y}\rangle=x_\mu+\lambda y_\mu\ . \tag{III.6}$$

Thus $x\leftrightarrow\widehat{x}$ is linear and 1-to-1. ∎

III.2. **Determinants and Minkowski Geometry.** The link between the mappings of (III.4)–(III.5) and Minkowski geometry comes from observing that the determinant of the hermitian matrix $\widehat{x}$ is the Minkowski length squared of the four-vector $x = (x_0, \vec{x})$. In particular defining $x_M^2 = x_0^2 - \vec{x}^{\,2}$, one has

$$\det \widehat{x} = (x_0 + x_3)(x_0 - x_3) - |x_1 - ix_2|^2 = x_0^2 - \vec{x}^{\,2} = x_M^2\ .$$

Now we study transformations $\widehat{x} \mapsto \widehat{x}'$ that are linear and that preserve the determinant of $\widehat{x}$.

III.3. **Irreducible Sets of Matrices.**

Definition III.2. *A set of matrices $\mathfrak{A}$ is irreducible, if any matrix C that commutes with every matrix in $\mathfrak{A}$ must be a multiple of the identity matrix.*

Proposition III.3. *Any two of the three Pauli matrices $\{\sigma_1, \sigma_2, \sigma_3\}$ are an irreducible set of* 2×2 *matrices.*

Proof. Take $C = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ for arbitrary complex numbers a, b, c, d. Explicitly

$$C\sigma_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} b & a \\ d & c \end{pmatrix}\ , \quad \text{and} \quad \sigma_1 C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ a & b \end{pmatrix}\ .$$

Hence $C\sigma_1 = \sigma_1 C$ means that $b = c$ and $a = d$. Likewise

$$C\sigma_2 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} ib & -ia \\ id & -ic \end{pmatrix}\ , \quad \text{and} \quad \sigma_2 C = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -ic & -id \\ ia & ib \end{pmatrix}\ .$$

Hence $C\sigma_2 = \sigma_2 C$ means that $b = -c$ and $a = d$. And finally

$$C\sigma_3 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} a & -b \\ c & -d \end{pmatrix}\ , \quad \text{while} \quad \sigma_3 C = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ -c & -d \end{pmatrix}\ .$$

Hence $C\sigma_3 = \sigma_3 C$ means that $b = c = 0$.

Inspecting the consequences of any two of the three conditions $C\sigma_j = \sigma_j C$, we infer $b = c = 0$ and $a = d$. In other words, C commuting with any two Pauli matrices shows that

$$C = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = a\, I\ .$$

Thus any two of the Pauli matrices form an irreducible set of 2×2 matrices. □

III.4. **Unitary Matrices are Exponentials of Anti-Hermitian Matrices.** The matrix H is hermitian, $H = H^*$, if and only if the matrix $A = iH$ is anti-hermitian, $A = -A^*$.

Proposition III.4. *Unitary matrices have the following properties:*

(1) *Every unitary $N \times N$ matrix U can be written $U = e^{iH}$, where H is hermitian.*
(2) *In case $N = 2$, the matrix H can be written $H = \sum_{j=0}^{3} \lambda_j \sigma_j$ for real constants λ_j.*
(3) *In case $N = 2$ and* $\det U = 1$, *the matrix H can be written $H = \sum_{j=1}^{3} \lambda_j \sigma_j$, namely $\lambda_0 = 0$.*

Proof. Every unitary matrix is normal, so there is another unitary V so that $V^*UV = D$ is diagonal. The diagonal entries $D_{jj} = d_j$ of D are the eigenvalues of U, and they have absolute value 1. Therefore $d_j = e^{i\lambda_j}$ for some real λ_j. Hence $D = e^{i\Lambda}$ where Λ is the diagonal hermitian matrix with real entries $\Lambda_{ij} = \lambda_j\delta_{ij}$. Hence $H = V\Lambda V^*$ is hermitian and $U = VDV^* = Ve^{i\Lambda}V^* = e^{iV\Lambda V^*} = e^{iH}$ has the form claimed in (1).

In case $N = 2$, we know that the σ_j for $j = 0, 1, 2, 3$ are a basis for hermitian matrices, that is orthonormal in the scalar product (III.2). Therefore the given H has a unique expansion $H = \sum_{j=0}^{3} \lambda_j\sigma_j$. In case that $\det U = 1$ as well, one has

$$\det U = \det(V^*DV) = \det(VV^*D) = \det D = \det\left(e^{i\Lambda}\right) = e^{i\mathrm{Tr}\Lambda} = e^{2i\lambda_0} .$$

Thus one needs to choose $\lambda_0 = n\pi$ for integer n. But all these choices yield the same U, so we choose $n = 0$. □

III.5. **Strictly-Positive Hermitian Matrices are Exponentials of Hermitian Matrices.** A strictly-positive $N \times N$ hermitian matrix is one with strictly positive eigenvalues, $h_j > 0$, for $j = 1, \ldots, N$.

Proposition III.5. *A strictly positive hermitian matrix has the following properties:*

(1) *Every $N \times N$ hermitian matrix H can be written $H = e^K$, where K is hermitian.*
(2) *In case $N = 2$, the matrix K can be written $K = \sum_{j=0}^{3} \lambda_j\sigma_j$ for real constants λ_j.*
(3) *In case $N = 2$ and $\det H = 1$, the matrix K can be written $K = \sum_{j=1}^{3} \lambda_j\sigma_j$, namely $\lambda_0 = 0$.*

Proof. The proof is very similar to the proof of Proposition III.4. One significant difference is that for $h_j > 0$, we can find real λ_j for which $h_j = e^{\lambda_j}$. Also $\det H = e^{2\lambda_0}$; so $\det H = 1$ ensures $\lambda_0 = 0$. □

IV. The Group $SL(2, \mathbb{C})$ Acting on H_2

Let A denote an element of the *two-dimensional special linear group*, namely the set of 2×2 matrices with complex entries and with determinant 1. As A has non-vanishing determinant, it has a matrix inverse. Each matrix $A \in SL(2, \mathbb{C})$ defines a linear transformation of hermitian matrices,

$$\widehat{x} \quad \mapsto \quad \widehat{x}\,' = A\,\widehat{x}\,A^* , \tag{IV.1}$$

that preserves the determinant. Clearly both A and $-A$ determine the same transformation $\widehat{x} \mapsto \widehat{x}'$. Linearity follows from

$$(\widehat{x} + \lambda\widehat{y})' = A\,(\widehat{x} + \lambda\widehat{y})\,A^* = A\widehat{x}A^* + \lambda A\widehat{y}A^* = \widehat{x}\,' + \lambda\widehat{y}\,' .$$

Also $\det\widehat{x}' = \det(A\,\widehat{x}\,A^*) = (\det A)(\det\widehat{x})(\det A^*) = \det\widehat{x}$.

Proposition IV.1. *Any linear transformation taking hermitian, 2×2 matrices $\widehat{x}$ to hermitian 2×2 matrices $\widehat{x}'$, with $\det\widehat{x} = \det\widehat{x}'$, and with $x_0x_0' \geqslant 0$, can be written in the form* (IV.1) *for some matrix $A \in SL(2, \mathbb{C})$. Furthermore the transformation $\widehat{x} \mapsto \widehat{x}'$ uniquely determines A, up to its overall sign.*

Proof. Since the linear, determinant-preserving transformations $\widehat{x} \mapsto \widehat{x}'$ are in 1 to 1 correspondence with Lorentz transformations, we know from (I.3) that there is a 6 real-parameter family of such transformations. An arbitrary 2×2 matrix has 4 complex or 8 real parameters. The restriction that $\det A = 1$ gives two equations, one for the real and one for the imaginary part of the determinant. That leaves $6 = 8-2$ real parameters for matrices in $SL(2,\mathbb{C})$. Therefore matrices $A \in SL(2,\mathbb{C})$ have the correct number of independent parameters to describe all Lorentz transformations.

Just to cover all bases, let us consider a more general transformation of the form $\widehat{x}' = A\widehat{x}B$, for both $A, B \in SL(2,\mathbb{C})$. Since $\widehat{x}'$ must be hermitian for all $\widehat{x}$, one must have $A\widehat{x}B = B^*\widehat{x}A^*$ for all $\widehat{x}$. This is the case if and only if $B^{*\,-1}A\widehat{x} = \widehat{x}A^*B^{-1}$ for all $\widehat{x}$. Taking $\widehat{x} = I$ shows that $T = B^{*\,-1}A$ must be hermitian. It then follows that $T\widehat{x} = \widehat{x}T$ for all $\widehat{x}$. But the matrices (σ_1, σ_2) are an irreducible set of matrices. They equal $\widehat{x}$ for the choices $x = (0,1,0,0)$ and $x = (0,0,1,0)$ respectively. Thus T commutes with an irreducible set and must be a multiple of the identity, $T = \lambda I$, with $\lambda \in \mathbb{R}$ as $T = T^*$. Also $\det T = 1$, so $\lambda^2 = 1$ and $\lambda = \pm 1$. This is equivalent to $B = \pm A^*$, and to

$$\widehat{x}' = \pm A\widehat{x}A^* \,. \tag{IV.2}$$

The $+$ sign in (IV.2) gives (IV.1). On the other hand suppose that for the $-$ sign, one has $\widehat{x}' = \widehat{\Lambda x}$. Then

$$\Lambda_{00} = -\left\langle \sigma_0, A\sigma_0A^* \right\rangle = -\frac{1}{2}\mathrm{Tr}(AA^*) < 0 \,. \tag{IV.3}$$

In this case Λ reverses the sign of the time, so it is ruled out by the assumption that $x_0x_0' \geqslant 0$. Note on the other hand that for $\widehat{x}' = A\widehat{x}A^*$, the same argument shows that $\Lambda_{00} = \frac{1}{2}\mathrm{Tr}(AA^*) > 0$, so $x_0x_0' \geqslant 0$.

Finally let us suppose that A and B are two different $SL(2,\mathbb{C})$ matrices that yield the same transformation $\widehat{x} \mapsto \widehat{x}'$. Then $A\widehat{x}A^* = B\widehat{x}B^*$ for all $\widehat{x}$. In other words $B^{-1}A\widehat{x}(B^{-1}A)^* = \widehat{x}$ for all $\widehat{x}$. Taking $\widehat{x} = I$ shows that $B^{-1}A$ is unitary, and so $B^{-1}A$ commutes with $\widehat{x}$ for all $\widehat{x}$. Again choosing $\widehat{x}$ to be σ_1 and σ_2 shows that $B^{-1}A = \lambda I$ where λ is a constant, or $A = \lambda B$. Since both A and B have determinant 1, we infer that $\lambda^2 = 1$ and $\lambda = \pm 1$. Hence A is determined uniquely up to its overall sign. □

Corollary IV.2. *There is a 2 to 1 correspondence between matrices $A \in SL(2,\mathbb{C})$ and restricted Lorentz transformations $\Lambda \in \mathcal{L}_+^\uparrow$, given by the representation*

$$\widehat{\Lambda x} = A\,\widehat{x}\,A^* \,. \tag{IV.4}$$

The matrices $\Lambda = \Lambda(A) = \Lambda(-A)$ are a representation of $SL(2,\mathbb{C})$, namely $\Lambda(AB) = \Lambda(A)\,\Lambda(B)$. The matrix elements of $\Lambda(A)$ are given by

$$\Lambda(A)_{\mu\nu} = \left\langle \sigma_\mu, A\,\sigma_\nu\,A^* \right\rangle = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_\mu\,A\,\sigma_\nu\,A^*\right) \,, \quad \textit{and} \quad \Lambda(A)_{\mu\nu} = \overline{\Lambda(A)_{\mu\nu}} \,, \tag{IV.5}$$

and are real. If $A = U$ is unitary, then $\Lambda(U) = R$ has the form (II.1) with $\mathcal{R} \in SO(3)$, a proper rotation. It has $\Lambda(U)_{00} = 1$, and $\Lambda(U)_{0i} = \Lambda(U)_{i0} = 0$ for $i = 1,2,3$. On the other hand, if $A = H$ is hermitian, then $\Lambda(H)$ is symmetric and equal to a pure Lorentz boost.

Proof. We have shown in Proposition IV.1 that the map (IV.4) gives a restricted Lorentz transformation, and that for a given restricted Lorentz transformation A is unique up to its overall sign.

In order to show that $A \mapsto \Lambda(A)$ is a representation, note that for every x,

$$(\Lambda(AB)x)^{\wedge} = AB\,\widehat{x}\,B^*A^* = A\,\widehat{\Lambda(B)x}\,A^* = (\Lambda(A)\Lambda(B)x)^{\wedge}\ .$$

Hence for any x, one can invert the map to hermitian matrices to show $\Lambda(AB)x = \Lambda(A)\,\Lambda(B)x$. As a consequence, the identity $\Lambda(AB) = \Lambda(A)\,\Lambda(B)$ also holds.

In order to see that the Lorentz transformation $\Lambda(A)$ is an element of $L_+^{\uparrow}$, we argue that both $\det \Lambda(A)$ and $\Lambda(A)_{00}$ are continuous functions of the matrix A. For every $A \in SL(2,\mathbb{C})$ we can find a continuous path $A(\alpha)$ from $A(0) = I$ to $A(1) = A$, which we construct as follows. The matrix A has a unique polar decomposition $A = HU$, where $0 < H$ is a positive hermitian matrix, and U is a unitary matrix. Every hermitian H and unitary U has an orthonormal basis of eigenvectors, so we can raise H and U to an arbitrary fractional power $\alpha \in [0,1]$. Define

$$A(\alpha) = H^{\alpha}U^{\alpha}\ ,\ \text{ for which }\ A(0) = I\ ,\ \text{ and }\ A(1) = A\ .$$

As $\Lambda^{\rm tr} g\Lambda = g$, we infer that $\det \Lambda^2 = 1$, so $\det \Lambda = \pm 1$. Also $\Lambda(A(0)) = I$, and $\det \Lambda(A(\alpha))$ is a continuous function of α; so $\det \Lambda(A) = \det \Lambda(A(1)) = 1$. Likewise $\Lambda(A(\alpha))_{00}^2 \geqslant 1$ is continuous in α, and $\Lambda(A(\alpha))_{00} = 1$ for $\alpha = 0$. Hence $\Lambda(A)_{00} \geqslant 1$. Taken together, these two facts show that $\Lambda(A) \in L_+^{\uparrow}$.

We can compute $\Lambda(A)$ from

$$\sum_{\nu=0}^{3} \Lambda(A)_{\mu\nu}\,x_\nu = (\Lambda(A)x)_\mu = \left\langle \sigma_\mu, \widehat{\Lambda(A)x} \right\rangle = \langle \sigma_\mu, A\,\widehat{x}\,A^* \rangle = \sum_{\nu=0}^{3} \langle \sigma_\mu, A\,\sigma_\nu\,A^* \rangle\ x_\nu\ . \tag{IV.6}$$

The values of the matrix elements $\Lambda(A)_{\mu\nu}$ are the coefficients of x_ν in the relation (IV.6). Thus $\Lambda(A)_{\mu\nu} = \langle \sigma_\mu, A\,\sigma_\nu\,A^* \rangle = \frac{1}{2}\,{\rm Tr}(\sigma_\mu A\,\sigma_\nu\,A^*)$. Also

$$\overline{\Lambda(A)_{\mu\nu}} = \frac{1}{2}\,\overline{{\rm Tr}(\sigma_\mu\,A\,\sigma_\nu\,A^*)} = \frac{1}{2}\,{\rm Tr}\left(A\,\sigma_\nu\,A^*\,\sigma_\mu\right) = \frac{1}{2}\,{\rm Tr}\left(\sigma_\mu\,A\,\sigma_\nu\,A^*\right) = \Lambda(A)_{\mu\nu}\ ,$$

so the matrix elements of $\Lambda(A)$ are real.

As $\Lambda(A)$ is a representation, $\Lambda(A^{-1}) = \Lambda(A)^{-1}$. In case that $A = U$ is unitary, $\Lambda(U)_{00} = \frac{1}{2}{\rm Tr}(I) = 1$. Also $\Lambda(U)_{0i} = \frac{1}{2}{\rm Tr}(U\sigma_i U^*) = \frac{1}{2}{\rm Tr}(\sigma_i) = 0$, and similarly $\Lambda(U)_{i0} = 0$. Furthermore

$$\begin{aligned}\Lambda(U)_{\mu\nu} &= \frac{1}{2}\,{\rm Tr}\left(\sigma_\mu\,U\sigma_\nu\,U^*\right) = \frac{1}{2}\,{\rm Tr}\left(\sigma_\mu\,U^{-1\,*}\sigma_\nu\,U^{-1}\right) = \frac{1}{2}\,{\rm Tr}\left(\sigma_\nu\,U^{-1}\,\sigma_\mu\,U^{-1\,*}\right) = \Lambda(U^{-1})_{\nu\mu} \\ &= \left(\Lambda(U)^{-1}\right)_{\nu\mu}\ .\end{aligned} \tag{IV.7}$$

Hence $\Lambda(U) = R(U)$ is orthogonal.

On the other hand, if $A = H$ is hermitian, then

$$\Lambda(H)_{\mu\nu} = \frac{1}{2}\,{\rm Tr}(\sigma_\mu\,H\,\sigma_\nu\,H) = \frac{1}{2}\,{\rm Tr}(H\,\sigma_\nu\,H\,\sigma_\mu) = \frac{1}{2}\,{\rm Tr}(\sigma_\nu\,H\,\sigma_\mu\,H) = \Lambda(H)_{\nu\mu}\ ,$$

so $\Lambda(H)$ is symmetric. □

V. Structure of Restricted Lorentz Transformations

We now identify the transformations $\Lambda(U)$ and $\Lambda(H)$ arising from unitary or self-adjoint matrices A as rotations and boosts. Also we decompose an arbitrary restricted Lorentz transformation into a product of a rotation and a boost, and this decomposition is unique.

First note that every matrix $A \in SL(2,\mathbb{C})$ has a polar decomposition into a strictly-positive hermitian matrix times a unitary matrix,

$$A = HU \;, \quad \text{where} \quad 0 < H \;, \quad \text{and} \quad UU^* = I \;. \tag{V.1}$$

One takes H as the positive square root of the hermitian matrix AA^*, namely $H = (AA^*)^{1/2}$, and defines $U = H^{-1}A = (AA^*)^{-1/2}A$. One then sees that U is unitary, for

$$UU^* = H^{-1}AA^*H^{-1} = (AA^*)^{-1/2}\, AA^*\, (AA^*)^{-1/2} = I \;. \tag{V.2}$$

Note that H and U are uniquely determined. We summarize this as:

Proposition V.1. *We therefore have a unique decomposition of $\Lambda(A)$ into*

$$\boxed{\Lambda(A) = \Lambda(HU) = \Lambda(H)\,\Lambda(U)}\;, \quad \textit{where} \quad \boxed{0 < H = H^* \;, \quad UU^* = I}\;. \tag{V.3}$$

Along the properties established in Corollary IV.2, this completes the proof of Proposition II.1. We now find out in detail what are the transformations $\Lambda(H)$ and $\Lambda(U)$.

Proposition V.2 (Identification of $\boldsymbol{\Lambda(U)}$ and $\boldsymbol{\Lambda(H)}$). *If $U = e^{-i\frac{\theta}{2}\vec{n}\cdot\vec{\sigma}}$, then $\Lambda(U) = R(\vec{n},\theta)$ rotates by angle θ about the axis $\vec{n}$. If $H = e^{\frac{\chi}{2}\vec{n}\cdot\vec{\sigma}}$, then $\Lambda(H) = B(\vec{n},\chi)$ is a pure boost along the axis $\vec{n}$ by velocity $\frac{v}{c} = \beta = \tanh\chi$.*

Proof. In order to identify the rotation arising from $U = e^{-i\frac{\theta}{2}\vec{n}\cdot\vec{\sigma}}$, note that U commutes with the matrix $\widehat{n} = \sum_{j=1}^{3} n_j\sigma_j$. Thus $U\widehat{n}\,U^* = \widehat{n}$, so $\widehat{n}' = \widehat{n}$, and consequently $\Lambda(U)$ leaves $\vec{n}$ unchanged. In other words it is a rotation about the axis $\vec{n}$.

In order to analyze the angle by which one rotates, it is sufficient to choose $\vec{n} = \vec{e}_3$. Then we claim that

$$\Lambda(U) = \Lambda(e^{-i\frac{\theta}{2}\,\sigma_3}) = \Lambda(e^{-i\frac{\theta}{2}\,\sigma_3}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \;, \tag{V.4}$$

which is just the rotation by angle θ about the third axis found in (II.3). We compute the matrix elements of $\Lambda(U)$ and show that they equal the matrix elements in (V.4). We have already shown in Corollary IV.2 that $\Lambda(U)_{00} = 1$ and $\Lambda(U)_{0i} = \Lambda(U)_{i0} = 0$, for $i = 1,2,3$. We also have

$$\Lambda(U)_{11} = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_1 e^{-i\frac{\theta}{2}\,\sigma_3}\sigma_1 e^{i\frac{\theta}{2}\,\sigma_3}\right) = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_1^2 e^{i\theta\,\sigma_3}\right) = \frac{1}{2}\,\mathrm{Tr}\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta}\end{pmatrix} = \frac{1}{2}\,\left(e^{i\theta} + e^{-i\theta}\right) = \cos\theta \;,$$

$$\begin{aligned}\Lambda(U)_{12} &= \frac{1}{2}\,\mathrm{Tr}\left(\sigma_1 e^{-i\frac{\theta}{2}\,\sigma_3}\sigma_2 e^{i\frac{\theta}{2}\,\sigma_3}\right) = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_1\sigma_2 e^{i\theta\,\sigma_3}\right) = \frac{1}{2}\,\mathrm{Tr}\left(i\sigma_3 e^{i\theta\,\sigma_3}\right)\\ &= i\frac{1}{2}\,\mathrm{Tr}\begin{pmatrix} e^{i\theta} & 0\\ 0 & -e^{-i\theta}\end{pmatrix} = -\frac{1}{2i}\left(e^{i\theta}-e^{-i\theta}\right) = -\sin\theta\ ,\end{aligned}$$

$$\Lambda(U)_{13} = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_1 e^{-i\frac{\theta}{2}\,\sigma_3}\sigma_3 e^{i\frac{\theta}{2}\,\sigma_3}\right) = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_1\sigma_3\right) = \frac{1}{2}\,\mathrm{Tr}\left(-i\sigma_2\right) = 0\ ,$$

$$\Lambda(U)_{21} = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_2 e^{-i\frac{\theta}{2}\,\sigma_3}\sigma_1 e^{i\frac{\theta}{2}\,\sigma_3}\right) = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_2\sigma_1 e^{i\theta\,\sigma_3}\right) = -\Lambda(U)_{12} = \sin\theta\ ,$$

$$\Lambda(U)_{22} = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_2 e^{-i\frac{\theta}{2}\,\sigma_3}\sigma_2 e^{i\frac{\theta}{2}\,\sigma_3}\right) = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_2^2 e^{i\theta\,\sigma_3}\right) = \frac{1}{2}\,\mathrm{Tr}\begin{pmatrix} e^{i\theta} & 0\\ 0 & e^{-i\theta}\end{pmatrix} = \cos\theta\ ,$$

$$\Lambda(U)_{23} = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_2 e^{-i\frac{\theta}{2}\,\sigma_3}\sigma_3 e^{i\frac{\theta}{2}\,\sigma_3}\right) = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_2\sigma_3\right) = \frac{1}{2}\,\mathrm{Tr}\left(i\sigma_1\right) = 0\ ,$$

$$\begin{aligned}\Lambda(U)_{31} &= \frac{1}{2}\,\mathrm{Tr}\left(\sigma_3 e^{-i\frac{\theta}{2}\,\sigma_3}\sigma_1 e^{i\frac{\theta}{2}\,\sigma_3}\right) = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_3\sigma_1 e^{i\theta\sigma_3}\right) = \frac{1}{2}\,\mathrm{Tr}\left(i\sigma_2 e^{i\theta\sigma_3}\right)\\ &= \frac{1}{2}\,\mathrm{Tr}\left(\begin{pmatrix} 0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix} e^{i\theta} & 0\\ 0 & e^{-i\theta}\end{pmatrix}\right) = \frac{1}{2}\,\mathrm{Tr}\begin{pmatrix} 0 & e^{-i\theta}\\ -e^{i\theta} & 0\end{pmatrix} = 0\ ,\end{aligned}$$

$$\begin{aligned}\Lambda(U)_{32} &= \frac{1}{2}\,\mathrm{Tr}\left(\sigma_3 e^{-i\frac{\theta}{2}\,\sigma_3}\sigma_2 e^{i\frac{\theta}{2}\,\sigma_3}\right) = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_3\sigma_2 e^{i\theta\sigma_3}\right) = \frac{1}{2}\,\mathrm{Tr}\left(-i\sigma_1 e^{i\theta\sigma_3}\right)\\ &= \frac{1}{2}\,\mathrm{Tr}\left(\begin{pmatrix} 0 & -i\\ -i & 0\end{pmatrix}\begin{pmatrix} e^{i\theta} & 0\\ 0 & e^{-i\theta}\end{pmatrix}\right) = \frac{1}{2}\,\mathrm{Tr}\begin{pmatrix} 0 & -ie^{-i\theta}\\ -ie^{i\theta} & 0\end{pmatrix} = 0\ ,\end{aligned}$$

and

$$\Lambda(U)_{33} = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_3 e^{-i\frac{\theta}{2}\,\sigma_3}\sigma_3 e^{i\frac{\theta}{2}\,\sigma_3}\right) = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_3^2\right) = \frac{1}{2}\,\mathrm{Tr}(I) = 1\ .$$

Putting together all these matrix elements shows that the rotation $\Lambda(U)$ just equals (II.3), as claimed.

Similarly we can work out the matrix $\Lambda(H)$, for

$$H = e^{\frac{1}{2}\,\vec{n}\cdot\vec{\sigma}\chi} = e^{\frac{1}{2}\,\vec{n}\cdot\vec{\sigma}\chi}\ , \tag{V.5}$$

where $\vec{n}$ is a unit vector and χ is a real parameter. As in the case of the rotation, we include the factor $\frac{1}{2}$ in the exponent, so that everything works out nicely.

For $i, j = 1, 2, 3$, we have

$$\sigma_i\sigma_j = -\sigma_j\sigma_i + 2\delta_{ij} I\ ,\ \text{ for }\ i, j = 1, 2, 3\ .$$

Assume that $y = (0, \vec{y})$ where $\vec{y} \in \mathbb{R}^3$ is orthogonal to $\vec{n}$, namely $\vec{y}\cdot\vec{n} = 0$. Then

$$(\vec{y}\cdot\vec{\sigma})\ (\vec{n}\cdot\vec{\sigma}) = -(\vec{n}\cdot\vec{\sigma})\ (\vec{y}\cdot\vec{\sigma})\ ,$$

and it follows from (V.5) that

$$\vec{y}\cdot\vec{\sigma}\,H = H^{-1}\,\vec{y}\cdot\vec{\sigma}\ .$$

Then

$$\widehat{\Lambda(H)y} = H\vec{y}\cdot\vec{\sigma}H = \vec{y}\cdot\vec{\sigma} = \widehat{y}\ . \tag{V.6}$$

So $\Lambda(H)$ leaves $\vec{y}\perp\vec{n}$ unchanged. This is the definition of a pure Lorentz boost along $\vec{n}$. We denote this transformation $B(H)$.

Let us compute the components of $B(H)$ for H or the form (V.5) and with $\vec{n}=e_1$. We claim that in this case,

$$B(H) = \begin{pmatrix} \cosh\chi & \sinh\chi & 0 & 0 \\ \sinh\chi & \cosh\chi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\ . \tag{V.7}$$

In other words, we claim that $B(H)$ is exactly the boost $B(\vec{e}_1,\chi)$ given in the example (II.4).

In the case that $A=H$ is hermitian, we have shown in Corollary IV.2 that the corresponding Lorentz boost $B(H)$ is always a real, symmetric matrix, so in order to determine $B(H)$, we only need to find $B(H)_{\mu\nu}$ for $\mu\leqslant\nu$. These components are

$$B(H)_{00} = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_0\, e^{\frac{\chi}{2}\,\sigma_1}\,\sigma_0\, e^{\frac{\chi}{2}\,\sigma_1}\right) = \frac{1}{2}\,\mathrm{Tr}\,(e^{\chi\,\sigma_1}) = \frac{1}{2}\,\mathrm{Tr}\begin{pmatrix}\cosh\chi & \sinh\chi \\ \sinh\chi & \cosh\chi\end{pmatrix} = \cosh\chi\ ,$$

$$B(H)_{01} = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_0\, e^{\frac{\chi}{2}\,\sigma_1}\,\sigma_1\, e^{\frac{\chi}{2}\,\sigma_1}\right) = \frac{1}{2}\,\mathrm{Tr}\,(\sigma_1\, e^{\chi\,\sigma_1}) = \frac{1}{2}\,\mathrm{Tr}\begin{pmatrix}\sinh\chi & \cosh\chi \\ \cosh\chi & \sinh\chi\end{pmatrix} = \sinh\chi\ ,$$

$$B(H)_{02} = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_0\, e^{\frac{\chi}{2}\,\sigma_1}\,\sigma_2\, e^{\frac{\chi}{2}\,\sigma_1}\right) = \frac{1}{2}\,\mathrm{Tr}\,(\sigma_2) = 0\ ,$$

and

$$B(H)_{03} = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_0\, e^{\frac{\chi}{2}\,\sigma_1}\,\sigma_3\, e^{\frac{\chi}{2}\,\sigma_3}\right) = \frac{1}{2}\,\mathrm{Tr}\,(\sigma_3) = 0\ .$$

Also

$$B(H)_{11} = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_1\, e^{\frac{\chi}{2}\,\sigma_1}\,\sigma_1\, e^{\frac{\chi}{2}\,\sigma_1}\right) = \frac{1}{2}\,\mathrm{Tr}\,(e^{\chi\,\sigma_1}) = \cosh\chi\ ,$$

$$B(H)_{12} = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_1\, e^{\frac{\chi}{2}\,\sigma_1}\,\sigma_2\, e^{\frac{\chi}{2}\,\sigma_1}\right) = \frac{1}{2}\,\mathrm{Tr}\,(\sigma_1\,\sigma_2) = \frac{1}{2}\,\mathrm{Tr}\,(i\sigma_3) = 0\ ,$$

$$B(H)_{13} = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_1\, e^{\frac{\chi}{2}\,\sigma_1}\,\sigma_3\, e^{\frac{\chi}{2}\,\sigma_1}\right) = \frac{1}{2}\,\mathrm{Tr}\,(\sigma_1\,\sigma_3) = \frac{1}{2}\,\mathrm{Tr}\,(-i\sigma_2) = 0\ ,$$

$$B(H)_{22} = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_2\, e^{\frac{\chi}{2}\,\sigma_1}\,\sigma_2\, e^{\frac{\chi}{2}\,\sigma_1}\right) = \frac{1}{2}\,\mathrm{Tr}\,(I) = 1\ ,$$

$$B(H)_{23} = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_2\, e^{\frac{\chi}{2}\,\sigma_1}\,\sigma_3\, e^{\frac{\chi}{2}\,\sigma_1}\right) = \frac{1}{2}\,\mathrm{Tr}\,(\sigma_2\,\sigma_3) = \frac{1}{2}\,\mathrm{Tr}\,(i\sigma_1) = 0\ ,$$

and

$$B(H)_{33} = \frac{1}{2}\,\mathrm{Tr}\left(\sigma_3\, e^{\frac{\chi}{2}\,\sigma_1}\,\sigma_3\, e^{\frac{\chi}{2}\,\sigma_1}\right) = \frac{1}{2}\,\mathrm{Tr}\,(I) = 1\ .$$

So these all agree with (V.7). □

VI. Complex-Conjugate Representations in $SU(2)$ and $SL(2,\mathbb{C})$

By the complex conjugate of the matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we mean the matrix $\overline{A} = \begin{pmatrix} \overline{a} & \overline{b} \\ \overline{c} & \overline{d} \end{pmatrix}$, where $\overline{a}$ denotes the complex conjugate of the matrix element a.

VI.1. The Group $SU(2)$.

We remark that for the unitary matrices $U \in SU(2)$, the transformation

$$U \mapsto U' = \sigma_2 U \sigma_2 \tag{VI.1}$$

maps any U to its complex conjugate matrix $\overline{U}$. To verify this, note that

$$\sigma_2\sigma_1\sigma_2 = -\sigma_1 = -\overline{\sigma_1}\ , \quad \sigma_2\sigma_2\sigma_2 = \sigma_2 = -\overline{\sigma_2}\ , \quad \sigma_2\sigma_3\sigma_2 = -\sigma_3 = -\overline{\sigma_3}\ . \tag{VI.2}$$

In summary, for $j = 1, 2, 3$, and for $i = \sqrt{-1}$,

$$\sigma_2\left(i\sigma_j\right)\sigma_2 = \overline{\left(i\sigma_j\right)}\ . \tag{VI.3}$$

Using the fact that σ_2 is unitary and self adjoint, it follows that

$$\sigma_2\left(i\sigma_j\right)^n\sigma_2 = \overline{\left(i\sigma_j\right)}^n\ . \tag{VI.4}$$

Writing $U = e^{-i\theta\vec{n}\cdot\vec{\sigma}}$, and using $\sigma_2 = \sigma_2^*$, we see that

$$\sigma_2\, U\sigma_2^* = \sigma_2\, U\sigma_2 = \overline{U}\ . \tag{VI.5}$$

Now consider two representations of the group $SU(2)$ by 2×2 unitary matrices. In the first case we take the representation $U \mapsto U$ given by the identity function. Let us call this the *standard* representation of $SU(2)$. Multiplication in this representation of the group is just the ordinary multiplication of matrices, that defines the group, $U_1U_2 = U_3$.

In the second case consider taking the complex conjugate of each matrix, $U \mapsto \overline{U}$. In order to check that this is a representation of $SU(2)$, note that for any matrices A and B, it is true that $\overline{AB} = \overline{A}\,\overline{B}$. Thus it follows that $\overline{U_1}\,\overline{U_2} = \overline{U_3}$, and the multiplication table for the complex conjugate matrices $\overline{U}$ is the same as the multiplication table for the matrices U. Hence the complex conjugate matrices yield another representation of the group $SU(2)$. One calls this representation of $SU(2)$ the *complex-conjugate representation*.

What we have shown in (VI.5) is that the complex conjugate representation is unitarily equivalent to the standard representation, and the unitary equivalence is implemented by the matrix σ_2,

$$\sigma_2U_1U_2\sigma_2^* = \sigma_2U_3\sigma_2^*\ , \quad \text{means that} \quad \overline{U_1}\,\overline{U_2} = \overline{U_3}\ . \tag{VI.6}$$

VI.2. The Group $SL(2,\mathbb{C})$.

In the group $SL(2,\mathbb{C})$ one can show as above that $A \mapsto \overline{A}$ gives a representation of the group. If $A_1A_2 = A_3$ then $\overline{A_1}\,\overline{A_2} = \overline{A_3}$. This is called the complex-conjugate representation. However now the situation with $SL(2,\mathbb{C})$ is different from the situation with the group of unitary matrices $SU(2)$. The complex conjugate representation $\overline{SL(2,\mathbb{C})}$ of $SL(2,\mathbb{C})$ is **not** unitarily equivalent to the standard representation.

There are many ways to see if two matrices X and Y are unitarily equivalent, namely if $Y = UXU^*$ for a unitary matrix U. A necessary condition for this to be the case is that

$$\mathrm{Tr}(Y) = \mathrm{Tr}(UXU^*) = \mathrm{Tr}(U^*UX) = \mathrm{Tr}(X)\ . \tag{VI.7}$$

Here we use "cyclicity of the trace," namely $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. Also $\mathrm{Tr}(\overline{Y}) = \overline{Tr(Y)}$. Therefore if X and $\overline{X}$ are unitarily equivalent, then $\mathrm{Tr}(X) = \overline{\mathrm{Tr}(X)}$ must be real.

Hence one way to show that a general matrix $A \in SL(2,\mathbb{C})$ is not unitarily equivalent to its complex conjugate is to find an $SL(2,\mathbb{C})$ matrix whose trace is complex. In fact, one can take the diagonal 2×2 matrix[2]

$$A = e^{\chi\sigma_3} e^{i\theta\sigma_3} = \begin{pmatrix} e^{\chi+i\theta} & 0 \\ 0 & e^{-\chi-i\theta} \end{pmatrix} \in SL(2,\mathbb{C}) \ , \quad \text{with} \quad \chi, \theta \in \mathbb{R} \ . \tag{VI.8}$$

Then

$$\mathrm{Tr}(A) = 2\cosh\chi\, \cos\theta + 2i \sinh\chi\, \sin\theta \ . \tag{VI.9}$$

This is real only if $\chi = 0$ or $\theta = n\pi$. So if we choose χ, θ so $\sinh\chi\, \cos\theta \neq 0$, then A is not unitarily equivalent to $\overline{A}$.

Remark VI.1 (Particle/Anti-Particle Symmetry). In quantum theory, complex conjugation is the symmetry that has the physical interpretation of transforming particle states to corresponding anti-particle states. (This is a slight oversimplification.)

Furthermore the two-dimensional matrices in $SU(2)$ or $SL(2,\mathbb{C})$ are involved in the interpretation of states of spin-$\frac{1}{2}$ particles, such as electrons and positrons, or protons and anti-protons. So one can interpret the difference between rotations and Lorentz transformations as saying that rotational symmetry does not mix spin-$\frac{1}{2}$ fermions from their anti-particles. However relativistic symmetry can mix the two.

P.A.M. Dirac discovered this fact, when he found his relativistic wave equation for the electron, the equation we know as the "Dirac equation." The Dirac equation is an equation for a 4-component wave function. In order to obtain this equation, one combines the 2-dimensional standard representation of $SL(2,\mathbb{C})$ (which is the familiar Pauli equation involving the Pauli matrices in non-relativistic quantum theory) with the 2-dimensional complex-conjugate representation, $\overline{SL(2,\mathbb{C})}$. Putting them together they make up the 4 components of the relativistic Dirac wave function. The Weyl equation is a two-component relativistic wave equation that describes particles or anti-particles. The Dirac equation is a four-component relativistic wave equation that describes both.

[2]Our example does not need the trace as a test, for we just observe that the set of eigenvalues of A, namely the diagonal entries, do not coincide with those of $\overline{A}$. However it is instructive to look at *invariants* of A under unitary transformations, such as $\mathrm{Tr}(A)$ and $\det(A)$ which must be the same for unitary equivalent matrices. These serve as useful tests when A is not diagonal.