LORENTZ TRANSFORMATIONS AS 2X2 MATRICES

Link to: physicspages home page.

To leave a comment or report an error, please use the auxiliary blog.

References: W. Greiner & J. Reinhardt, *Field Quantization*, Springer-Verlag (1996), Chapter 2, Section 2.4.

Arthur Jaffe, *Lorentz transformations, rotations and boosts*, online notes available (at time of writing, Sep 2016) here.

Continuing our examination of general Lorentz transformations, recall that a Lorentz transformation can be represented by a 4×4 matrix Λ which preserves the Minkowski length $x_\mu x^\mu$ of all four-vectors x. This leads to the condition

$$\Lambda^T g\Lambda = g \tag{1}$$

where g is the flat-space Minkowski metric

$$g=\left[\begin{array}{cccc}1 & 0 & 0 & 0\\0 & -1 & 0 & 0\\0 & 0 & -1 & 0\\0 & 0 & 0 & -1\end{array}\right] \tag{2}$$

It turns out that we can map any 4-vector x to a 2×2 Hermitian matrix $\widehat{x}$ defined as

$$\widehat{x}\equiv\left[\begin{array}{cc}x_0+x_3 & x_1-ix_2\\x_1+ix_2 & x_0-x_3\end{array}\right] \tag{3}$$

[Recall that a Hermitian matrix H is equal to the complex conjugate of its transpose:

$$H=\left(H^T\right)^*\equiv H^\dagger \tag{4}$$

Also note that Jaffe uses an unconventional notation for the Hermitian conjugate, as he uses a superscript * rather that a superscript †. This can be confusing since usually a superscript * indicates just complex conjugate, without the transpose. I'll use the more usual superscript † for Hermitian conjugate here.]

Although we're used to the scalar product of two vectors, it is also useful to define the scalar product of two matrices as

$$\left\langle A,B\right\rangle \equiv\frac{1}{2}\mathrm{Tr}\left(A^\dagger B\right) \tag{5}$$

where 'Tr' means the trace of a matrix, which is the sum of its diagonal elements. Note that the scalar product of $\widehat{x}$ with itself is

$$\langle\widehat{x},\widehat{x}\rangle = \frac{1}{2}\mathrm{Tr}\begin{bmatrix} x_0+x_3 & x_1-ix_2 \\ x_1+ix_2 & x_0-x_3 \end{bmatrix}\begin{bmatrix} x_0+x_3 & x_1-ix_2 \\ x_1+ix_2 & x_0-x_3 \end{bmatrix} \tag{6}$$

$$\frac{1}{2}\left[(x_0+x_3)^2+2(x_1-ix_2)(x_1+ix_2)+(x_0-x_3)^2\right] \tag{7}$$

$$= x_0^2+x_1^2+x_2^2+x_3^2 \tag{8}$$

The determinant of $\widehat{x}$ is

$$\det\widehat{x} = (x_0+x_3)(x_0-x_3)-(x_1-ix_2)(x_1+ix_2) \tag{9}$$

$$= x_0^2-x_1^2-x_2^2-x_3^2 \tag{10}$$

$$= x_\mu x^\mu \tag{11}$$

Thus $\det\widehat{x}$ is the Minkowski length squared.
From 3, we observe that we can write $\widehat{x}$ as a sum:

$$\widehat{x} = \sum_{\mu=0}^{4} x_\mu \sigma_\mu \tag{12}$$

where the σ_μ are four Hermitian matrices:

$$\sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \tag{13}$$

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{14}$$

$$\sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \tag{15}$$

$$\sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{16}$$

The last three are the Pauli spin matricesthat we met when looking at spin-$\frac{1}{2}$ in quantum mechanics.
The σ_μ are orthonormal under the scalar product operation, as we can verify by direct calculation. For example

$$\begin{aligned}
\langle\sigma_2,\sigma_3\rangle &= \frac{1}{2}\mathrm{Tr}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} && (17)\\
&= \frac{1}{2}(0+0) && (18)\\
&= 0 && (19)
\end{aligned}$$

And:

$$\begin{aligned}
\langle\sigma_2,\sigma_2\rangle &= \frac{1}{2}\mathrm{Tr}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} && (20)\\
&= \frac{1}{2}(1+1) && (21)\\
&= 1 && (22)
\end{aligned}$$

The other products work out similarly, so we have

$$\left\langle\sigma_\mu,\sigma_\nu\right\rangle = \delta_{\mu\nu} \tag{23}$$

We can work out the inverse transformation to 3 by taking the scalar product of 12 with σ_ν:

$$\begin{aligned}
\langle\sigma_\nu,\widehat{x}\rangle &= \sum_{\mu=0}^{4} x_\mu\left\langle\sigma_\nu,\sigma_\mu\right\rangle && (24)\\
&= \sum_{\mu=0}^{4} x_\mu\delta_{\nu\mu} && (25)\\
&= x_\nu && (26)
\end{aligned}$$

Now a few more theorems that will be useful later.

Irreducible Sets of Matrices. A set of matrices $\mathfrak{U}$ is called *irreducible* if the only matrix C that commutes with every matrix in $\mathfrak{U}$ is the identity matrix I (or a multiple of I). Any two of the three Pauli matrices σ_i, $i = 1,2,3$ above form an irreducible set of 2×2 Hermitian matrices. This can be shown by direct calculation, which Jaffe does in detail in his article. For example, if we define C to be some arbitrary matrix

$$C = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{27}$$

where a,b,c,d are complex numbers, then

$$C\sigma_1 = \begin{bmatrix} b & a \\ d & c \end{bmatrix} \tag{28}$$

$$\sigma_1 C = \begin{bmatrix} c & d \\ a & b \end{bmatrix} \tag{29}$$

If C is to commute with σ_1, we must therefore require $b = c$ and $a = d$.
Similarly, for σ_2 we have

$$C\sigma_2 = \begin{bmatrix} ib & -ia \\ id & -ic \end{bmatrix} \tag{30}$$

$$\sigma_2 C = \begin{bmatrix} -ic & -id \\ ia & ib \end{bmatrix} \tag{31}$$

so that $C\sigma_2 = \sigma_2 C$ requires $b = -c$ and $a = d$.
And for σ_3:

$$C\sigma_3 = \begin{bmatrix} a & -b \\ c & -d \end{bmatrix} \tag{32}$$

$$\sigma_3 C = \begin{bmatrix} a & b \\ -c & -d \end{bmatrix} \tag{33}$$

so that $C\sigma_3 = \sigma_3 C$ requires $b = -b$ and $c = -c$, so $b = c = 0$ (no conditions can be inferred for a or d).

If we form a set $\mathfrak{U}$ containing σ_3 and one of σ_1 or σ_2, we see that $b = c = 0$ and $a = d$, so C is a multiple of I. If we form $\mathfrak{U}$ from σ_1 and σ_2 we again have $a = d$, but we must have simultaneously $b = c$ and $b = -c$ which can be true only if $b = c = 0$, so again C is a multiple of I.

Unitary Matrices. A unitary matrix is one whose Hermitian conjugate is its inverse, so that $U^\dagger = U^{-1}$. Some properties of unitary matrices are given on the Wikipedia page, so we'll just use those without going through the proofs. First, a unitary matrix is *normal*, which means that $U^\dagger U = UU^\dagger$ (this actually follows from the condition $U^\dagger = U^{-1}$). Second, there is another unitary matrix V which diagonalizes U, that is

$$V^\dagger U V = D \tag{34}$$

where D is a diagonal, unitary matrix.
Third,

$$|\det U| = 1 \tag{35}$$

(The determinant can be complex, but has magnitude 1.)

From this it follows that $|\det D| = 1$ and since D is unitary and diagonal, each diagonal element d_j of D must satisfy $\left|d_j\right| = 1$. (Remember that d_j could be a complex number.) That means that $d_j = e^{i\lambda_j}$ for some *real* number λ_j, so we can write

$$D = e^{i\Lambda} \tag{36}$$

where Λ is a diagonal hermitian matrix containing only real elements, non-zero along its diagonal: $\Lambda_{ij} = \lambda_j \delta_{ij}$. As usual, the exponential of a matrix is interpreted in terms of its power series, so that

$$e^{i\Lambda} = 1 + i\Lambda + \frac{(i\Lambda)^2}{2!} + \frac{(i\Lambda)^3}{3!} + \ldots \tag{37}$$

For a diagonal matrix Λ with diagonal elements $\Lambda_{jj} = \lambda_j$, the diagonal elements of Λ^n are just $\Lambda^n_{jj} = \lambda^n_j$.

From 34, we have

$$\begin{aligned} U &= VDV^{\dagger} && (38)\\ &= Ve^{i\Lambda}V^{\dagger} && (39) \end{aligned}$$

Now we also have, since $VV^{\dagger} = I$

$$\begin{aligned} V\Lambda^n V^{\dagger} &= V\Lambda\left(VV^{\dagger}\right)\Lambda\left(VV^{\dagger}\right)\ldots\Lambda V^{\dagger} && (40)\\ &= \left(V\Lambda V^{\dagger}\right)^n && (41) \end{aligned}$$

Therefore, from 37

$$\begin{aligned} U &= Ve^{i\Lambda}V^{\dagger} && (42)\\ &= e^{iV\Lambda V^{\dagger}} && (43)\\ &\equiv e^{iH} && (44) \end{aligned}$$

where $H = V\Lambda V^{\dagger}$ is another Hermitian matrix. In other words, we can always write a unitary matrix as the exponential of a Hermitian matrix.

In the case where H is a 2×2 matrix, we can write it in terms of the σ_μ matrices above as

$$H = \sum_{\mu=0}^{3} a_\mu \sigma_\mu \tag{45}$$

where the a_μ are real, since the diagonal elements of a Hermitian matrix must be real. This follows because the σ_μ form an orthonormal basis for the 2×2 Hermitian matrices. [For some reason, Jaffe refers to the $a_\mu$as

λ_μ which is confusing since he has used λ_μ as the diagonal elements of Λ above, and they're not the same thing.]

If $\det U = +1$, then

$$\begin{aligned} \det U &= \det\left(VDV^\dagger\right) && (46)\\ &= \det\left(VV^\dagger D\right) && (47)\\ &= \det D && (48)\\ &= \det e^{i\Lambda} && (49) \end{aligned}$$

The second line follows because the determinant of a product of matrices is the product of the determinants, so we can rearrange the multiplication order. To evaluate the last line, we observe that for a diagonal matrix Λ, using 37 and applying the result to each diagonal element

$$e^{i\Lambda} = \begin{bmatrix} e^{i\Lambda_{11}} & 0 \\ 0 & e^{i\Lambda_{22}} \end{bmatrix} \qquad (50)$$

Therefore

$$\det e^{i\Lambda} = e^{i(\Lambda_{11}+\Lambda_{22})} = e^{i\mathrm{Tr}\Lambda} \qquad (51)$$

[By the way, the relation $\det e^A = e^{\mathrm{Tr}A}$ is actually true for any square matrix A, and is a corollary of *Jacobi's formula.*]

We can now use the cyclic property of the trace (another matrix algebra theroem) which says that for 3 matrices A, B, C,

$$\mathrm{Tr}(ABC) = \mathrm{Tr}(CAB) = \mathrm{Tr}(BCA) \qquad (52)$$

This gives us

$$\mathrm{Tr}H = \mathrm{Tr}\left(V\Lambda V^\dagger\right) = \mathrm{Tr}\left(V^\dagger V\Lambda\right) = \mathrm{Tr}\Lambda \qquad (53)$$

Finally, from 45 and the fact that the traces of the σ_i are all zero for $i = 1, 2, 3$, and $\mathrm{Tr}\sigma_0 = 2$, we have

$$\det U = \det e^{i\Lambda} = e^{i\mathrm{Tr}H} = e^{2ia_0} = 1 \qquad (54)$$

Thus $a_0 = n\pi$ for some integer n, but as all values of n give the same original unitary matrix U, we can choose $n = 0$ so that $a_0 = 0$ and

$$H = \sum_{\mu=1}^{3} a_\mu \sigma_\mu \qquad (55)$$

PINGBACKS

Pingback: Lorentz transformations and the special linear group SL(2,C)