

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

$$(1) \quad \Lambda^T g \Lambda = g$$

where g is the flat-space Minkowski metric

$$(2) \quad g = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

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

$$(3) \quad \hat{x} \equiv \begin{bmatrix} x_0 + x_3 & x_1 - ix_2 \\ x_1 + ix_2 & x_0 - x_3 \end{bmatrix}$$

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

$$(4) \quad H = (H^T)^* \equiv H^\dagger$$

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

$$(5) \quad \langle A, B \rangle \equiv \frac{1}{2} \text{Tr} (A^\dagger B)$$

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

$$(6) \quad \langle \hat{x}, \hat{x} \rangle = \frac{1}{2} \text{Tr} \begin{bmatrix} x_0 + x_3 & x_1 - ix_2 \\ x_1 + ix_2 & x_0 - x_3 \end{bmatrix} \begin{bmatrix} x_0 + x_3 & x_1 - ix_2 \\ x_1 + ix_2 & x_0 - x_3 \end{bmatrix}$$

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

$$(8) \quad = x_0^2 + x_1^2 + x_2^2 + x_3^2$$

The determinant of \hat{x} is

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

$$(10) \quad = x_0^2 - x_1^2 - x_2^2 - x_3^2$$

$$(11) \quad = x_\mu x^\mu$$

Thus $\det \hat{x}$ is the Minkowski length squared.

From 3, we observe that we can write \hat{x} as a sum:

$$(12) \quad \hat{x} = \sum_{\mu=0}^4 x_\mu \sigma_\mu$$

where the σ_μ are four Hermitian matrices:

$$(13) \quad \sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

$$(14) \quad \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$(15) \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

$$(16) \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

The last three are the Pauli spin matrices that 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

$$(17) \quad \langle \sigma_2, \sigma_3 \rangle = \frac{1}{2} \text{Tr} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$(18) \quad = \frac{1}{2} (0 + 0)$$

$$(19) \quad = 0$$

And:

$$(20) \quad \langle \sigma_2, \sigma_2 \rangle = \frac{1}{2} \text{Tr} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

$$(21) \quad = \frac{1}{2} (1 + 1)$$

$$(22) \quad = 1$$

The other products work out similarly, so we have

$$(23) \quad \langle \sigma_\mu, \sigma_\nu \rangle = \delta_{\mu\nu}$$

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

$$(24) \quad \langle \sigma_\nu, \hat{x} \rangle = \sum_{\mu=0}^4 x_\mu \langle \sigma_\nu, \sigma_\mu \rangle$$

$$(25) \quad = \sum_{\mu=0}^4 x_\mu \delta_{\nu\mu}$$

$$(26) \quad = x_\nu$$

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

$$(27) \quad C = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

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

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

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

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

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

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

so that $C\sigma_2 = \sigma_2 C$ requires $b = -c$ and $a = d$.

And for σ_3 :

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

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

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 $U^\dagger U = U U^\dagger$ (this actually follows from the condition $U^\dagger = U^{-1}$). Second, there is another unitary matrix V which diagonalizes U , that is

$$(34) \quad V^\dagger U V = D$$

where D is a diagonal, unitary matrix.

Third,

$$(35) \quad |\det U| = 1$$

(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 $|d_j| = 1$. (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

$$(36) \quad D = e^{i\Lambda}$$

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

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

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

From 34, we have

$$(38) \quad U = VDV^\dagger$$

$$(39) \quad = Ve^{i\Lambda}V^\dagger$$

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

$$(40) \quad V\Lambda^n V^\dagger = V\Lambda(VV^\dagger)\Lambda(VV^\dagger)\dots\Lambda V^\dagger$$

$$(41) \quad = (V\Lambda V^\dagger)^n$$

Therefore, from 37

$$(42) \quad U = Ve^{i\Lambda}V^\dagger$$

$$(43) \quad = e^{iV\Lambda V^\dagger}$$

$$(44) \quad \equiv e^{iH}$$

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

$$(45) \quad H = \sum_{\mu=0}^3 a_\mu \sigma_\mu$$

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_μ 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

$$(46) \quad \det U = \det(VDV^\dagger)$$

$$(47) \quad = \det(VV^\dagger D)$$

$$(48) \quad = \det D$$

$$(49) \quad = \det e^{i\Lambda}$$

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

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

Therefore

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

[By the way, the relation $\det e^A = e^{\text{Tr}A}$ 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 theorem) which says that for 3 matrices A, B, C ,

$$(52) \quad \text{Tr}(ABC) = \text{Tr}(CAB) = \text{Tr}(BCA)$$

This gives us

$$(53) \quad \text{Tr}H = \text{Tr}(V\Lambda V^\dagger) = \text{Tr}(V^\dagger V\Lambda) = \text{Tr}\Lambda$$

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

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

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

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

PINGBACKS

Pingback: Lorentz transformations and the special linear group $SL(2, \mathbb{C})$