As we know from last semester, rotations about an arbitrary axis $\hat{n}$ can be written in exponential form,

$$R(\hat{n},\theta) = e^{\theta \hat{n}\cdot\vec{J}}$$

where $\vec{J}$ is a "vector" of matrices, given explicitly by

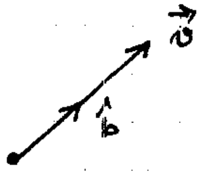
$$J_1 = \left(\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

$$J_2 = \left(\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right)$$

$$J_3 = \left(\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

These are the same as last semester, except that now we are appending an extra row and column of zeros to take care of the time component. The spatial part of these matrices is the same as last semester (see Eq. 9.22).

Similarly, boosts along an arbitrary direction $\hat{b}$ can be written in exponential form. Let the velocity of the boost be written $\vec{v} = v\hat{b}$ where $-\infty < v < \infty$ is the velocity component along $\hat{b}$.



Then the boost $B(\hat{b},\lambda)$, where λ is the rapidity, can be written

$$B(\hat{b},\lambda) = e^{\lambda \hat{b}\cdot\vec{\mathcal{K}}},$$

where $\vec{\mathcal{K}}$ is another (in addition to $\vec{J}$) "vector" of matrices, given by

$$\mathcal{K}_1 = \left(\begin{array}{c|ccc} 0 & 1 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$\mathcal{K}_2 = \left(\begin{array}{c|ccc} 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$\mathcal{K}_3 = \left(\begin{array}{c|ccc} 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right)$$

You can easily check that $e^{\lambda\hat{b}\cdot\vec{\mathcal{K}}}$ gives the right answer when $\hat{b} = \hat{x}, \hat{y}$ or $\hat{z}$, and the general case is not too hard, either.

The 6 matrices $\vec{J}, \vec{\mathcal{K}}$ constitute the <u>Lie algebra</u> of the Lorentz group. They satisfy characteristic commutation relations.

Return now to the Dirac equation. To understand how it transforms under Lorentz transformations, we must understand how the Dirac wave function transforms. Before getting into this, let's review how other wave functions transform under ordinary rotations.

First, for the wave function of a nonrelativistic spin-0 particle, we have the transformation law

$$\psi(\vec{r}) \xrightarrow{R} \psi(R^{-1}\vec{r}),$$

where $R \in SO(3)$. This means that if $\psi(\vec{r})$ is an old wave function and $\psi'(\vec{r})$ the new one, after a rotation has been applied, then $\psi'(\vec{r}) = \psi(R^{-1}\vec{r})$. We must use R^{-1} instead of R to get the active transformation law. See Sec. 13.2.

Next, for a NR particle of spin s, the wave fn. is a $(2s+1)$-component spinor, that transforms according to

$$\psi(\vec{r}) = \begin{pmatrix} \psi_s(\vec{r}) \\ \vdots \\ \psi_{-s}(\vec{r}) \end{pmatrix} \xrightarrow{R} D^s(R)\,\psi(R^{-1}\vec{r}),$$

where $D^s(R)$ is the rotation D-matrix for the value $j=s$ (a $(2s+1)\times(2s+1)$ matrix; see Sec. 11.10). The new wave fn. at point $\vec{r}$ is the old one at point $R^{-1}\vec{r}$, multiplied by the spinor rotn matrix $D^s(R)$. This was a homework problem last semester.

Note that $D^s(R)$ forms a representation of the rotations,

$$D^s(R_1)\,D^s(R_2) = D^s(R_1R_2).$$

(But see further remarks below.).

In particular, in the case $s=1/2$, ψ is a 2-component spinor, and D^s is the 2×2 rotation matrix that in axis-angle form is

$$D^{1/2}(\hat{n},\theta) = U(\hat{n},\theta) = e^{-i\frac{\theta}{2}\hat{n}\cdot\vec{\sigma}} = \cos\frac{\theta}{2} - i(\hat{n}\cdot\vec{\sigma})\sin\frac{\theta}{2}$$

(different notations for the spin-1/2 rotation matrix). In this case $U(R)$ is really a double-valued representation of the classical rotations ($R \to \pm U(R)$), so the representation law might better be written

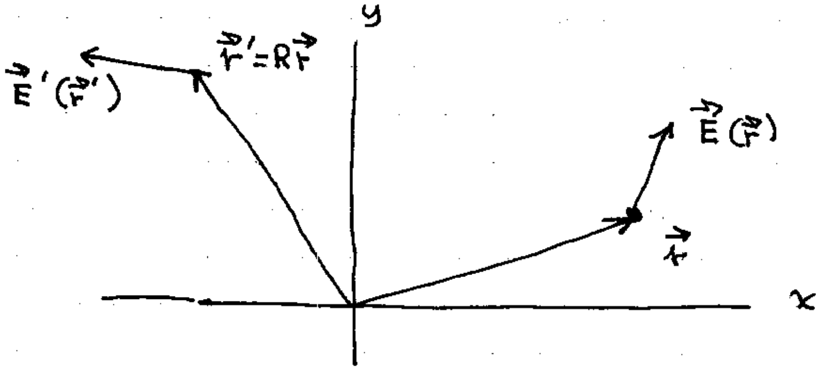
$$U(R_1)\,U(R_2) = \pm\, U(R_1R_2).$$

In any case, the spinor transforms according to

$$\psi(\vec{r}) = \begin{pmatrix} \psi_+(\vec{r}) \\ \psi_-(\vec{r}) \end{pmatrix} \xrightarrow{R(\hat{n},\theta)} U(\hat{n},\theta)\,\psi(R^{-1}\vec{r}).$$

We have seen that Dirac spinors (the upper 2 components) become Pauli spinors in the NR limit, so this transformation law must come out of the transformation law for Dirac spinors that we are about to derive.

Finally, let's examine how a classical vector field such as an electric field $\vec{E}(\vec{r})$ transforms under a rotation. A picture shows what the rotated field $\vec{E}'$ evaluated at the rotated point $\vec{r}' = R\vec{r}$ must be:



The electric field itself must be rotated along with the point at which it is evaluated, that is, we must have $\vec{E}'(\vec{r}') = R\,\vec{E}(\vec{r})$. Now replacing $\vec{r} \to R^{-1}\vec{r}$, we get

$$\vec{E}'(\vec{r}) = R\,\vec{E}(R^{-1}\vec{r}).$$

In all these transformation laws, the new field at point $\vec{r}$ is the old field at the inverse rotated point $R^{-1}\vec{r}$, multiplied by whatever type of rotation is required by the value of the field (scalar, spinor, classical vector, etc).

Now we can guess how a Dirac spinor must transform under Lorentz transformations. Let $x = (ct, \vec{x})$ stand for the complete space-time dependence. Then we must have

$$\psi(x) = \begin{pmatrix} \psi_1(x) \\ \vdots \\ \psi_4(x) \end{pmatrix} \xrightarrow{\Lambda} D(\Lambda)\,\psi(\Lambda^{-1}x)$$

where $D(\Lambda)$ is a 4×4 matrix parameterized by Λ acting on spinors that brings about the effect of Lorentz transformations on the spin part of the wave function. These matrices must satisfy the representation law,

$$D(\Lambda_1)D(\Lambda_2) = D(\Lambda_1\Lambda_2),$$

in order that the application of two Lorentz transformations Λ_1, Λ_2 in succession to a Dirac wave fn (Λ_1 first, Λ_2 second) has the same effect as applying $\Lambda = \Lambda_1\Lambda_2$. More precisely, in view of what happens with NR Pauli spinors (a limiting case of the Dirac spinors) we expect that this representation will be double valued, so that

$$D(\Lambda_1)D(\Lambda_2) = \pm D(\Lambda_1\Lambda_2)$$

We will see that this turns out to be the case.

It's a guess that this transformation law $\psi(x) \to D(\Lambda)\psi(\Lambda^{-1}x)$ is correct, but it gives us something to start with. $D(\Lambda)$ is some as-yet-to-be-determined, 4×4 spinor representation of the classical Lorentz group. For the time being, we are thinking of proper L.T.'s only.

Let's concentrate on the space-time part of the transformation law for now to see if it makes physical sense. Suppose we have an eigenfunction of energy and momentum,

$$\psi(x) \sim e^{i(\vec{p}\cdot\vec{x} - Et)/\hbar},$$

where $\sim$ means that we will ignore the spin part of ψ and concentrate on the spatial part. The phase can be put into covariant form by using $p^\mu = (E/c, \vec{p})$, $x^\mu = (ct, \vec{x})$, so that

$$\psi(x) \sim e^{-i p_\mu x^\mu/\hbar}.$$

The phase of the wave is a Lorentz invariant, something that holds for classical light waves, too. This makes sense, since the phase is something that can be measured (at least for a classical light wave) at a given space-time point, and the answer should be independent of Lorentz frame.

Under a L.T. this wave goes over into

$$D(\Lambda)\,\psi(\Lambda^{-1}x) \sim e^{-i p_\mu (\Lambda^{-1})^\mu{}_\nu x^\nu/\hbar}.$$

Let's look at the new phase, and use $(\Lambda^{-1})^\mu{}_\nu = g^{\mu\alpha}\Lambda^\beta{}_\alpha g_{\beta\nu}$:

$$p_\mu(\Lambda^{-1})^\mu{}_\nu x^\nu = p_\mu g^{\mu\alpha}\Lambda^\beta{}_\alpha g_{\beta\nu} x^\nu = p^\alpha \Lambda^\beta{}_\alpha x_\beta$$

$$= p'^\beta x_\beta = p'_\mu x^\mu,$$

where $p'^\mu = \Lambda^\mu{}_\nu p^\nu$. We see that if p^μ is the energy-momentum 4-vector of the original plane wave (its energy and momentum eigenvalues), then the energy-momentum 4-vector of the ~~~~ Lorentz transformed wave fn. is $p' = \Lambda p$, exactly what we would expect under an active transformation. So the $\psi(\Lambda^{-1}x)$ part of the transformation law (the space-time part) makes sense.

What is left is the spin part, ie, the matrices $D(\Lambda)$. We obtain a condition on these by requiring that the Dirac eqn be covariant. ~~~~ For this it suffices to work with the free-particle Dirac eqn. Suppose $\psi(x)$ is a free particle solution, i.e. it satisfies

$$i\hbar \gamma^\mu \frac{\partial \psi(x)}{\partial x^\mu} = mc\, \psi(x).$$

Let us demand that the Lorentz transformed wave fn. $\psi'(x) = D(\Lambda)\psi(\Lambda^{-1}x)$ also satisfy the free particle Dirac eqn, i.e. we demand that Lorentz transformations map free particle solutions into other free particle solutions. Then we have

$$i\hbar \gamma^\mu D(\Lambda) \frac{\partial}{\partial x^\mu} \psi(\Lambda^{-1}x) = mc\, D(\Lambda)\psi(\Lambda^{-1}x).$$

In this formula, space-time indices are indicated explicitly, e.g. $\gamma^\mu \frac{\partial}{\partial x^\mu}$, but spinor indices are not. For example, ψ is a 4-component column spinor and $D(\Lambda)$ a 4×4 matrix, and matrix multiplication is implied. We have pulled the $\frac{\partial}{\partial x^\mu}$ past $D(\Lambda)$ since $D(\Lambda)$ depends on Λ but not on x.

Let us write

$$y^\mu = (\Lambda^{-1})^\mu{}_{\cdot\nu}\, x^\nu$$

and multiply the eqn above by $D(\Lambda)^{-1}$ to clear the $D(\Lambda)$ on the RHS. Then we get

$$i\hbar\, D(\Lambda)^{-1}\gamma^\mu D(\Lambda)\, \frac{\partial\psi(y)}{\partial x^\mu} = mc\,\psi(y).$$

But since ψ satisfies the Dirac eqn, we have

$$i\hbar\,\gamma^\mu \frac{\partial\psi(y)}{\partial y^\mu} = mc\,\psi(y)$$

(y is just a dummy variable), and we also have

$$\frac{\partial\psi(y)}{\partial x^\mu} = \frac{\partial\psi(y)}{\partial y^\nu}\underbrace{\frac{\partial y^\nu}{\partial x^\mu}}_{(\Lambda^{-1})^\nu{}_{\cdot\mu}}.$$

So, cancelling $i\hbar$, we get

$$D(\Lambda)^{-1}\gamma^\mu D(\Lambda)\, \frac{\partial\psi(y)}{\partial y^\nu}(\Lambda^{-1})^\nu{}_{\cdot\mu} = i\hbar\,\gamma^\nu \frac{\partial\psi(y)}{\partial y^\nu}.$$

But $\partial\psi(y)/\partial y^\nu$ is arbitrary, so

$$D(\Lambda)^{-1}\gamma^\mu D(\Lambda)\,(\Lambda^{-1})^\nu{}_{\cdot\mu} = \gamma^\nu,$$

or, multiplying by Λ to clear the Λ^{-1} on the LHS,

$$\boxed{D(\Lambda)^{-1}\, \gamma^\mu\, D(\Lambda) = \Lambda^\mu{}_\nu\, \gamma^\nu} \, .$$

This is the equation that the as-yet-unknown representation $D(\Lambda)$ must satisfy, in order that free particle solutions be mapped into other free particle solutions by Lorentz transformations. If we had worked with the passive instead of active point of view, this would be the condition that the free particle Dirac equation should have the same form in all Lorentz frames. We must use this equation to determine $D(\Lambda)$.

Notice the similarity of the eqn. above to the transformation law for vector operators under rotations. If $\vec{V}$ is a vector operator with components V_i, then

$$U(R)^\dagger V_i\, U(R) = \sum_j R_{ij} V_j \, .$$

Here $U(R)$ is the rotation operator acting on the same ket space that $\vec{V}$ acts on. This is the definition of a vector operator. By analogy, we see that the equation above says that γ^μ transforms as a 4-vector under Lorentz transformation. This is the justification for the covariant notation used with γ^μ. It implies (as we will see) that scalar products like $\not{A} = \gamma^\mu A_\mu$ are actually Lorentz invariant. Later we will see, in addition to vector operators like γ^μ, also tensor operators of various ranks.

Our strategy will be to use the vector transformation law for γ^μ to determine $D(\Lambda)$, first in the case of infinitesimal transformations. Notice that if this law is true for two L.T.'s Λ_1 and Λ_2,

$$D(\Lambda_1)^{-1}\gamma^\mu D(\Lambda_1) = (\Lambda_1)^\mu_{\cdot\nu}\gamma^\nu$$

$$\text{and } D(\Lambda_2)^{-1}\gamma^\mu D(\Lambda_2) = (\Lambda_2)^\mu_{\cdot\nu}\gamma^\nu,$$

then it is true for the product $\Lambda = \Lambda_1\Lambda_2$, since

$$D(\Lambda)^{-1}\gamma^\mu D(\Lambda) = D(\Lambda_1\Lambda_2)^{-1}\gamma^\mu D(\Lambda_1\Lambda_2)$$

$$= D(\Lambda_2)^{-1} D(\Lambda_1)^{-1}\gamma^\mu D(\Lambda_1)D(\Lambda_2)$$

$$= D(\Lambda_2)^{-1}(\Lambda_1)^\mu_{\cdot\nu}\gamma^\nu D(\Lambda_2)$$

$$= (\Lambda_1)^\mu_{\cdot\nu}\; D(\Lambda_2)^{-1}\gamma^\nu D(\Lambda_2)$$

$$= (\Lambda_1)^\mu_{\cdot\nu}\;(\Lambda_2)^\nu_{\cdot\sigma}\gamma^\sigma = (\Lambda_1\Lambda_2)^\mu_{\cdot\sigma}\gamma^\sigma$$

$$= \Lambda^\mu_{\cdot\sigma}\gamma^\sigma.$$

So, if it is true for infinitesmal transformations, then it is true for any Lorentz transformation that can be built up as products of infinitesimals, which includes all proper Lorentz transformations. This relies on the fact that $D(\Lambda)$ forms a representation of the Lorentz group.

Let us put $D(\Lambda)$ on hold and look at classical, infinitesimal Lorentz transformations. Suppose $\Lambda = I + C$ where C = "correction" is a small matrix. Then

$$(I+C)^t g \, (I+C) = g,$$

or

$$C^t g + gC = 0,$$

or

$$C^t g = (gC)^t = -gC,$$

or $\quad gC =$ antisymmetric.

Compare this to what we had with purely spatial rotations. If $R^tR = I$ and $R = I + A$ (A = small correction), then A = antisym. Now introduce the basis $\mathcal{J}_i$ $(i=1,2,3)$ of antisymmetric 3×3 matrices, so that A is a linear combination of these with expansion coefficients we call θ_i, $i=1,2,3$:

$$A = \sum_i \theta_i \mathcal{J}_i = \vec{\theta}\cdot\vec{\mathcal{J}},$$

$$R = I + \vec{\theta}\cdot\vec{\mathcal{J}} \qquad (\text{infinitesimal rot}^n,\ \vec{\theta} = \text{small}).$$

The axis and angle of this rotation are given by $\vec{\theta} = \hat{n}\theta$, ie.

$$\theta = |\vec{\theta}|, \qquad \hat{n} = \frac{\vec{\theta}}{\theta}.$$

For Lorentz transformations, let us introduce a basis of 4×4 antisymmetric matrices, call them $\mathcal{B}^{(\alpha\beta)}$ ($\mathcal{B}$ = "basis"). Here are some examples of the $\mathcal{B}^{(\alpha\beta)}$, which give the idea:

$$B^{(01)} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B^{(12)} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

etc. In general, $B^{(\alpha\beta)} = 0$ if $\alpha = \beta$, and otherwise $B^{(\alpha\beta)}$ has a 1 in position $(\alpha\beta)$ and a -1 in position $(\beta\alpha)$ (and 0 everywhere else). This rule leads to the general formula,

$$\left(B^{(\alpha\beta)}\right)_{\mu\nu} = \delta^\alpha_\mu \delta^\beta_\nu - \delta^\alpha_\nu \delta^\beta_\mu$$

The parens around $(\alpha\beta)$ indicate that these are not component indices, but rather labels of the matrix. The components (in the formula above) are $\mu\nu$. The parens are just for clarity. Similarly we could write $\gamma^{(\mu)}$ on the Dirac matrices to indicate that μ is a label of the matrix, not a component index, but we won't because there is probably no danger of confusion.

~~Now since $B^{\alpha\beta}$ is antisymmetric, it~~

The $B^{(\alpha\beta)}$ matrices have the following properties:

$$B^{(\alpha\beta)\,t} = -B^{(\alpha\beta)} \qquad \text{(they are anti-symmetric)}$$

$$B^{(\beta\alpha)} = -B^{(\alpha\beta)}$$

There are only 6 independent $B^{(\alpha\beta)}$, obtained if we restrict the indices to $\alpha < \beta$.

Now since gC is antisymmetric, we can expand it as a linear combination of the $B^{(\alpha\beta)}$,

$$gC = \sum_{\alpha<\beta} \theta_{\alpha\beta}\, B^{(\alpha\beta)}$$

where $\theta_{\alpha\beta}$ are the expansion coefficients (6 of them). These are infinitesimal since C is infinitesimal. If we extend the definition of $\theta_{\alpha\beta}$ to the case $\alpha>\beta$ by requiring $\theta_{\alpha\beta} = -\theta_{\beta\alpha}$, then

$$gC = \frac{1}{2}\,\theta_{\alpha\beta}\, B^{(\alpha\beta)},$$

where we use the usual summation convention (summing over all values of α,β).

Now multiply by g^{-1} to solve for C, and define

$$\mathcal{J}^{(\alpha\beta)} = g^{-1} B^{(\alpha\beta)},$$

so that

$$C = \frac{1}{2}\,\theta_{\alpha\beta}\, \mathcal{J}^{(\alpha\beta)}$$

and the infinitesimal Lorentz transformation becomes

$$\Lambda = I + \frac{1}{2}\,\theta_{\alpha\beta}\, \mathcal{J}^{(\alpha\beta)}.$$

It is specified by the 6 independent small parameters $\theta_{\alpha\beta}$, the same number of parameters needed to specify an arbitrary Lorentz transformation.

The components of the $\mathcal{J}^{(\alpha\beta)}$ matrices are easy to work out. They are

$$\left(\mathcal{J}^{(\alpha\beta)}\right)^{\mu}_{\ \cdot\nu} = g^{\mu\sigma}\left(B^{(\alpha\beta)}\right)_{\sigma\nu}$$

$$= g^{\mu\sigma}\left(\delta^{\alpha}_{\sigma}\delta^{\beta}_{\nu} - \delta^{\alpha}_{\nu}\delta^{\beta}_{\sigma}\right)$$

$$= g^{\mu\alpha}\delta^{\beta}_{\nu} - g^{\mu\beta}\delta^{\alpha}_{\nu}.$$

As matrices, here are a few of the $\mathcal{J}^{(\alpha\beta)}$:

$$\mathcal{J}^{(01)} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \mathcal{K}_1$$

$$\mathcal{J}^{(12)} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \mathcal{J}_3$$

etc. It's not hard to see that

$$\mathcal{J}^{(0i)} = \mathcal{K}_i$$

$$\mathcal{J}^{(ij)} = \varepsilon_{ijk}\mathcal{J}_k$$

Thus the matrices $\mathcal{J}^{(\alpha\beta)}$ constitute another version of the Lie algebra of the Lorentz group (in addition to $\vec{\mathcal{J}}, \vec{\mathcal{K}}$).

By specializing the coefficients $\theta_{\alpha\beta}$ we get either pure rotations or pure boosts. Look at rotations first. Suppose $\theta_{0i}=0$, so

$$\Lambda = I + \tfrac{1}{2}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)} = I + \tfrac{1}{2}\theta_{ij}\mathcal{J}^{(ij)}$$

$$= I + \tfrac{1}{2}\theta_{ij}\,\varepsilon_{ijk}\mathcal{J}_k$$

$$= I + \theta_k\mathcal{J}_k = I + \vec{\theta}\cdot\vec{\mathcal{J}},$$

where

$$\theta_k = \tfrac{1}{2}\varepsilon_{ijk}\theta_{ij} \quad \text{or} \quad \theta_{ij} = \varepsilon_{ijk}\theta_k.$$

This is the usual formula for infinitesimal rotations, and it shows how the θ_i (related to the axis and angle of the rotation by $\vec{\theta}=\hat{n}\theta$) are contained in the space-space part of $\theta_{\alpha\beta}$, i.e., in θ_{ij}. In summary, if

$$\theta_{\alpha\beta} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \theta_3 & -\theta_2 \\ 0 & -\theta_3 & 0 & \theta_1 \\ 0 & \theta_2 & -\theta_1 & 0 \end{pmatrix},$$

then $\Lambda = I + \vec{\theta}\cdot\vec{\mathcal{J}} = R(\hat{n},\theta) \quad (\theta \ll 1). \quad (\theta = |\vec{\theta}|).$

Now look at pure boosts. If $\theta_{ij}=0$, then the only nonzero $\theta_{\alpha\beta}$ are θ^{0i} and $\theta^{i0} = -\theta^{0i}$, so

$$\Lambda = I + \tfrac{1}{2}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)} = I + \tfrac{1}{2}\left(\underbrace{\theta_{0i}\mathcal{J}^{(0i)} + \theta_{i0}\mathcal{J}^{(i0)}}_{\text{equal}}\right)$$

$$= I + \theta_{0i}\mathcal{J}^{(0i)} = I + \lambda_i\mathcal{K}_i = I + \vec{\lambda}\cdot\vec{\mathcal{K}}$$

where we define $\lambda_i = \theta_{0i}$ and we use $\mathcal{J}^{(0i)} = \mathcal{K}_i$. If we now write $\hat{b} = \dfrac{\vec{\lambda}}{\lambda}$, $\lambda = |\vec{\lambda}|$, then this is $I + \lambda\hat{b}\cdot\vec{\mathcal{K}}$, which is an infinitesimal boost $B(\hat{b},\lambda) = I + \lambda \hat{b}\cdot\vec{\mathcal{K}}$. In summary, if

$$\theta_{\alpha\beta} = \begin{pmatrix} 0 & \lambda_1 & \lambda_2 & \lambda_3 \\ -\lambda_1 & 0 & 0 & 0 \\ -\lambda_2 & 0 & 0 & 0 \\ -\lambda_3 & 0 & 0 & 0 \end{pmatrix},$$

then $\Lambda = I + \frac{1}{2}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)} = I + \vec{\lambda}\cdot\vec{\mathcal{K}} = I + \lambda\hat{b}\cdot\vec{\mathcal{K}}$.

We have exponential forms for pure rotations and pure boosts when θ or λ is not small. These are

$$R(\hat{n},\theta) = e^{\theta\hat{n}\cdot\vec{\mathcal{J}}} = e^{\vec{\theta}\cdot\vec{\mathcal{J}}}$$

$$B(\hat{b},\lambda) = e^{\lambda\hat{b}\cdot\vec{\mathcal{K}}} = e^{\vec{\lambda}\cdot\vec{\mathcal{K}}}.$$

It looks like the general infinitesimal Lorentz transformation $I + \frac{1}{2}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)}$ is the first terms of an exponential series $e^{\frac{1}{2}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)}}$ when $\theta_{\alpha\beta}$ are not small. This would be an exponential form for a Lorentz transformation containing both rotations and boosts. This is correct, the matrix $e^{\frac{1}{2}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)}}$ is a Lorentz transformation for all values of $\theta_{\alpha\beta}$, not only small ones.

To prove this let

$$M(s) = \exp\left[\frac{s}{2}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)}\right]$$

where s is a parameter. We will prove that $M(s)$ is a Lorentz transformation for all values of s. Consider the s-dependent matrix

$$F(s) = M(s)^t \, g \, M(s).$$

Note that

$$\frac{dM(s)}{ds} = \left[\tfrac{1}{2}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)}\right] M(s)$$

so

$$\frac{dF}{ds} = M(s)^t \left[\tfrac{1}{2}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)t}\right] g \; M(s)$$

$$+ M(s)^t g \left[\tfrac{1}{2}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)}\right] M(s)$$

$$= \tfrac{1}{2}\theta_{\alpha\beta}\, M(s)^t \left[\mathcal{J}^{(\alpha\beta)t} g + g\,\mathcal{J}^{(\alpha\beta)}\right] M(s).$$

But $\mathcal{J}^{(\alpha\beta)} = g^{-1} B^{(\alpha\beta)}$, so

$$= B^{(\alpha\beta)t} g^{-1} g + g g^{-1} B^{(\alpha\beta)} = B^{(\alpha\beta)t} + B^{(\alpha\beta)} = 0$$

since $B^{(\alpha\beta)}$ is antisymmetric. Thus $\frac{dF}{ds} = 0$, and $F(s) = F(0)$ $= M(0)^t g\, M(0) = g$ since $M(0) = I$. So

$$M(s)^t g \, M(s) = g,$$

and $M(s)$ is a Lorentz transformation.

Setting $s = 1$, we conclude that

$$\Lambda(\theta_{\alpha\beta}) = e^{\frac{1}{2}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)}}$$

is a Lorentz transformation for all values of $\theta_{\alpha\beta}$. This is a parameterization of Lorentz transformations by 6 parameters

$\theta_{\alpha\beta}$.

Now return to $D(\Lambda)$ which is a 4×4 spinor matrix depending on Λ. But since Λ depends on $\theta_{\alpha\beta}$, we can think of D as depending on $\theta_{\alpha\beta}$, too, $D(\theta_{\alpha\beta})$. To get the form for infinitesimal Lorentz transformations, let $\theta_{\alpha\beta}$ be small, and expand D to first order:

$$D(\Lambda) = D(\theta_{\alpha\beta}) = 1 + \sum_{\alpha<\beta} \theta_{\alpha\beta} \frac{\partial D}{\partial \theta_{\alpha\beta}}(0).$$

We only sum over $\alpha<\beta$ since only $\theta_{\alpha\beta}$ for $\alpha<\beta$ are independent. Now define

$$\frac{\partial D}{\partial \theta_{\alpha\beta}}(0) = -\frac{i}{2}\sigma^{\alpha\beta}$$

for $\alpha<\beta$, and then define $\sigma^{\alpha\beta} = -\sigma^{\beta\alpha}$ for $\alpha>\beta$. The $-i/2$ is conventional, but is intended to make the answers come out in familiar form for rotations on NR spinors, which we know about already. Then extending the sum over all α,β, we have

$$D(\Lambda) = 1 - \frac{i}{4}\theta_{\alpha\beta}\sigma^{\alpha\beta},$$

an infinitesimal spinor transformation valid when $\theta_{\alpha\beta} \ll 1$, and corresponding to the classical Lorentz transformation

$$\Lambda = I + \frac{1}{2}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)}.$$

We don't know what $\sigma^{\alpha\beta}$ is yet, but it is some collection of 6 Dirac matrices (an antisymmetric tensor of 4×4 Dirac matrices) that are constants (independent of $\theta_{\alpha\beta}$).

To find what $\sigma^{\alpha\beta}$ are, we use the vector transformation law on γ^μ, $D(\Lambda)^{-1}\gamma^\mu D(\Lambda) = \Lambda^\mu{}_\nu \gamma^\nu$, specializing to infinitesimal transformations. This gives (note $D(\Lambda)^{-1} = 1 + \frac{i}{4}\theta_{\alpha\beta}\sigma^{\alpha\beta}$):

$$\left(1 + \frac{i}{4}\theta_{\alpha\beta}\sigma^{\alpha\beta}\right)\gamma^\mu\left(1 - \frac{i}{4}\theta_{\alpha\beta}\sigma^{\alpha\beta}\right) = \left[I + \frac{1}{2}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)}\right]^\mu{}_\nu \gamma^\nu$$

or $$\gamma^\mu + \frac{i}{4}\theta_{\alpha\beta}\left[\sigma^{\alpha\beta},\gamma^\mu\right] = \left[\delta^\mu_\nu + \frac{1}{2}\theta_{\alpha\beta}\left(g^{\mu\alpha}\delta^\beta_\nu - g^{\mu\beta}\delta^\alpha_\nu\right)\right]\gamma^\nu$$

$$= \gamma^\mu + \frac{1}{2}\theta_{\alpha\beta}\left(g^{\mu\alpha}\gamma^\beta - g^{\mu\beta}\gamma^\alpha\right)$$

or, since $\theta_{\alpha\beta}$ is contracted against something antisymmetric on both sides, and since $\theta_{\alpha\beta}$ is arbitrary (but antisymmetric), we can cancel the $\theta_{\alpha\beta}$, and we get

$$\frac{i}{4}\left[\sigma^{\alpha\beta},\gamma^\mu\right] = \frac{1}{2}\left(g^{\mu\alpha}\gamma^\beta - g^{\mu\beta}\gamma^\alpha\right).$$

This equation must be solved for $\sigma^{\alpha\beta}$.

We guess that $\sigma^{\alpha\beta}$ can be expressed in terms of the γ^μ matrices. This is because ~~the [illegible]~~ Dirac's construction of his equation gave us a 4-vector γ^μ, but not any other relativistic tensor operator, so we guess that all others can be constructed out of γ^μ. Since $\sigma^{\alpha\beta}$ is antisymmetric in $\alpha\beta$, we guess that it must be proportional to the commutator $[\gamma^\alpha,\gamma^\beta]$, since this is also antisymmetric and would transform as a second rank tensor. That is, we guess

$$\sigma^{\alpha\beta} = k[\gamma^\alpha, \gamma^\beta] \qquad \text{some constant } k.$$

Then we have

$$\frac{ik}{4}[[\gamma^\alpha,\gamma^\beta],\gamma^\mu] \overset{?}{=} \frac{1}{2}(g^{\mu\alpha}\gamma^\beta - g^{\mu\beta}\gamma^\alpha).$$

To check this, first note that

$$[\gamma^\alpha,\gamma^\beta] = \gamma^\alpha\gamma^\beta - \gamma^\beta\gamma^\alpha$$

$$\hookrightarrow + \gamma^\alpha\gamma^\beta - 2g^{\alpha\beta},$$

where we use the anticommutator $\{\gamma^\alpha,\gamma^\beta\} = 2g^{\alpha\beta}$. So,

$$[\gamma^\alpha,\gamma^\beta] = 2(\gamma^\alpha\gamma^\beta - 2g^{\alpha\beta}),$$

and

$$[[\gamma^\alpha,\gamma^\beta],\gamma^\mu] = 2[\gamma^\alpha\gamma^\beta - 2g^{\alpha\beta},\gamma^\mu] = 2[\gamma^\alpha\gamma^\beta,\gamma^\mu]$$

since $g^{\alpha\beta}$ (times the identity matrix) commutes with anything.

But this is

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

$$\hookrightarrow + \gamma^\alpha\gamma^\mu\gamma^\beta - 2g^{\mu\alpha}\gamma^\beta$$

$$\hookrightarrow - \gamma^\alpha\gamma^\beta\gamma^\mu + 2g^{\mu\beta}\gamma^\alpha.$$

The cubic terms cancel, and

$$[[\gamma^\alpha,\gamma^\beta],\gamma^\mu] = -4(g^{\mu\alpha}\gamma^\beta - g^{\mu\beta}\gamma^\alpha).$$

The identity above is satisfied if $k = i/2$, so we get

$$\boxed{\sigma^{\alpha\beta} = \frac{i}{2}\left[\gamma^\alpha, \gamma^\beta\right]}$$

and

$$D(\Lambda) = 1 - \frac{i}{4}\theta_{\alpha\beta}\sigma^{\alpha\beta} \qquad \text{when } \theta_{\alpha\beta} \ll 1$$

and

$$\Lambda = I + \frac{1}{2}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)}$$

← representation for $D(\Lambda)$ when Λ is infinites.

As we have seen, when $\theta_{\alpha\beta}$ is not small, $\Lambda(\theta_{\alpha\beta}) = e^{\frac{1}{2}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)}}$, of which the first 2 terms of the Taylor series are shown above. The corresponding $D(\Lambda)$ matrix can also be written in exponential form; it is

$$D(\Lambda(\theta_{\alpha\beta})) \equiv D(\theta_{\alpha\beta}) = e^{-\frac{i}{4}\theta_{\alpha\beta}\sigma^{\alpha\beta}}.$$

This can be proved as follows. (The idea is to build up finite transformations from infinitesimal ones, using the representation law $D(\Lambda_1)D(\Lambda_2) = D(\Lambda_1\Lambda_2)$.) Suppose $\theta_{\alpha\beta}$ is not small, but N (an integer) is so large that $\theta_{\alpha\beta}/N$ is small. Then

$$\Lambda(\theta_{\alpha\beta}) = e^{\frac{1}{2}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)}} = \left[e^{\frac{1}{2N}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)}}\right]^N$$

$$\approx \left[I + \frac{\theta_{\alpha\beta}}{2N}\mathcal{J}^{(\alpha\beta)}\right]^N.$$

In fact, we have the exact limit,

$$\lim_{N\to\infty}\left[I + \frac{1}{2N}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)}\right]^N = \exp\left[\frac{\theta_{\alpha\beta}}{2}\mathcal{J}^{(\alpha\beta)}\right] = \Lambda(\theta_{\alpha\beta}).$$

This is ~~a~~ a matrix version of the limit from elementary calculus,

$$\lim_{N\to\infty} \left(1 + \frac{x}{N}\right)^N = e^x .$$

Now since each factor $I + \frac{1}{2N} \theta_{\alpha\beta} \mathcal{J}^{(\alpha\beta)}$ is infinitesimal, it corresponds to the infinitesimal $D(\Lambda)$ matrix $1 - \frac{i}{4N} \theta_{\alpha\beta} \sigma^{\alpha\beta}$. So,

$$\Lambda(\theta_{\alpha\beta}) = \lim_{N\to\infty} \left[I + \frac{1}{2N} \theta_{\alpha\beta} \mathcal{J}^{(\alpha\beta)} \right]^N$$

$$\longrightarrow \lim_{N\to\infty} \left(1 - \frac{i}{4N} \theta_{\alpha\beta} \sigma^{\alpha\beta}\right)^N = e^{-\frac{i}{4}\theta_{\alpha\beta}\sigma^{\alpha\beta}} = D(\Lambda)$$
$$= D(\theta_{\alpha\beta}).$$

In summary, for finite Lorentz transformations we have the association,

$$\boxed{\begin{aligned} &\Lambda(\theta_{\alpha\beta}) = \exp\left(\frac{1}{2}\theta_{\alpha\beta}\mathcal{J}^{(\alpha\beta)}\right) \\ &\to D(\Lambda) = D(\theta_{\alpha\beta}) = \exp\left(-\frac{i}{4}\theta_{\alpha\beta}\sigma^{\alpha\beta}\right) \end{aligned}}$$

From this we can compute the $D(\Lambda)$ matrices for any combination of rotations and boosts.