Dirac Matrices and Lorentz Spinors

Background: In 3D, the spinor $j = \frac{1}{2}$ representation of the Spin(3) rotation group is constructed from the Pauli matrices $\sigma^x$, $\sigma^y$, and $\sigma^z$, which obey both commutation and anticommutation relations

$$[\sigma^i, \sigma^j] = 2i\epsilon^{ijk}\sigma^k \quad \text{and} \quad \{\sigma^i, \sigma^j\} = 2\delta^{ij} \times 1_{2 \times 2}.$$  \hspace{1cm} (1)$$

Consequently, the spin matrices

$$S = -\frac{i}{2} \sigma \times \sigma = \frac{1}{2} \sigma$$  \hspace{1cm} (2)$$

commute with each other like angular momenta, $[S^i, S^j] = i\epsilon^{ijk}S^k$, so they represent the generators of the rotation group. In this spinor representation, the finite rotations $R(\phi, n)$ are represented by

$$M(R) = \exp(-i\phi n \cdot S),$$  \hspace{1cm} (3)$$

while the spin matrices themselves transform into each other as components of a 3–vector,

$$M^{-1}(R)S^i M(R) = R^{ij}S^j.$$  \hspace{1cm} (4)$$

In this note, I shall generalize this construction to the Dirac spinor representation of the Lorentz symmetry Spin(3,1).

The Dirac Matrices $\gamma^\mu$ generalize the anti-commutation properties of the Pauli matrices $\sigma^i$ to the $3 + 1$ Minkowski dimensions:

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu} \times 1_{4 \times 4}.$$  \hspace{1cm} (5)$$

The $\gamma^\mu$ are $4 \times 4$ matrices, but there are several different conventions for their specific form. In my class I shall follow the same convention as the Peskin & Schroeder textbook, namely
the Weyl convention where in $2 \times 2$ block notations

$$\gamma^0 = \begin{pmatrix} 0 & 1_{2\times2} \\ 1_{2\times2} & 0 \end{pmatrix}, \quad \vec{\gamma} = \begin{pmatrix} 0 & +\vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix}. \quad (6)$$

Note that the $\gamma^0$ matrix is hermitian while the $\gamma^1$, $\gamma^2$, and $\gamma^3$ matrices are anti-hermitian. Apart from that, the specific forms of the matrices are not important, the Physics follows from the anti-commutation relations (5).

**The Lorentz spin matrices** generalize $S = -\frac{i}{2} \sigma \times \sigma$ rather than $S = \frac{1}{2} \sigma$. In 4D, the vector product becomes the antisymmetric tensor product, so we define

$$S^{\mu\nu} = -S^{\nu\mu} \overset{\text{def}}{=} \frac{i}{4} [\gamma^\mu, \gamma^\nu]. \quad (7)$$

Thanks to the anti-commutation relations (5) for the $\gamma^\mu$ matrices, the $S^{\mu\nu}$ obey the commutation relations of the Lorentz generators $\hat{J}^{\mu\nu} = -\hat{J}^{\nu\mu}$. Moreover, the commutation relations of the spin matrices $S^{\mu\nu}$ with the Dirac matrices $\gamma^\mu$ are similar to the commutation relations of the $\hat{J}^{\mu\nu}$ with a Lorentz vector such as $\hat{P}^\mu$.

**Lemma:**

$$[\gamma^\lambda, S^{\mu\nu}] = i g^{\lambda\mu} \gamma^\nu - i g^{\lambda\nu} \gamma^\mu. \quad (8)$$

**Proof:** Combining the definition (7) of the spin matrices as commutators with the anti-commutation relations (5), we have

$$\gamma^\mu \gamma^\nu = \frac{1}{2} \{\gamma^\mu, \gamma^\nu\} + \frac{1}{2} [\gamma^\mu, \gamma^\nu] = g^{\mu\nu} \times 1_{4\times4} - 2i S^{\mu\nu}. \quad (9)$$

Since the unit matrix commutes with everything, we have

$$[X, S^{\mu\nu}] = \frac{i}{2} [X, \gamma^\mu \gamma^\nu] \quad \text{for any matrix } X, \quad (10)$$

and the commutator on the RHS may often be obtained from the Leibniz rules for the commutators or anticommutators:


$$\{A, BC\} = [A, B]C + B\{A, C\} = \{A, B\}C - B[A, C]. \quad (11)$$
In particular,
\[
[\gamma^\lambda, \gamma^\mu \gamma^\nu] = \{\gamma^\lambda, \gamma^\mu\} \gamma^\nu - \gamma^\mu \{\gamma^\lambda, \gamma^\nu\} = 2g^{\lambda\mu} \gamma^\nu - 2g^{\lambda\nu} \gamma^\mu
\]  
(12)
and hence
\[
[\gamma^\lambda, S^{\mu\nu}] = \frac{i}{2} [\gamma^\lambda, \gamma^\mu \gamma^\nu] = ig^{\lambda\mu} \gamma^\nu - ig^{\lambda\nu} \gamma^\mu.
\]  
(13)

*Quod erat demonstrandum.*

**Theorem:** The $S^{\mu\nu}$ matrices commute with each other like Lorentz generators,
\[
[S^{\kappa\lambda}, S^{\mu\nu}] = ig^{\lambda\mu} S^{\kappa\nu} - ig^{\lambda\nu} S^{\kappa\mu} - ig^{\kappa\mu} S^{\lambda\nu} + ig^{\kappa\nu} S^{\lambda\mu}.
\]  
(14)

**Proof:** Again, we use the Leibniz rule and eq. (9):
\[
[S^{\kappa\lambda}, S^{\mu\nu}] = \gamma^\kappa [\gamma^\lambda, S^{\mu\nu}] + [\gamma^\kappa, S^{\mu\nu}] \gamma^\lambda
\]
\[
= \gamma^\kappa (ig^{\lambda\mu} \gamma^\nu - ig^{\lambda\nu} \gamma^\mu) + (ig^{\kappa\mu} \gamma^\nu - ig^{\kappa\nu} \gamma^\mu) \gamma^\lambda
\]
\[
= ig^{\lambda\mu} (\gamma^\kappa \gamma^\lambda = g^{\kappa\nu} - 2iS^{\kappa\nu}) - ig^{\lambda\nu} (\gamma^\kappa \gamma^\lambda = g^{\kappa\mu} - 2iS^{\kappa\mu})
\]
\[
+ ig^{\kappa\mu} (\gamma^\nu \gamma^\lambda = g^{\lambda\nu} + 2iS^{\lambda\nu}) - ig^{\kappa\nu} (\gamma^\mu \gamma^\lambda = g^{\lambda\mu} + 2iS^{\lambda\mu})
\]
\[
= 2g^{\lambda\mu} S^{\kappa\nu} - 2g^{\lambda\nu} S^{\kappa\mu} - 2g^{\kappa\mu} S^{\lambda\nu} + 2g^{\kappa\nu} S^{\lambda\mu}
\]
(15)
since all the \(\pm ig^{\cdots} g^{-\cdot}\) cancel each other, hence
\[
[S^{\kappa\lambda}, S^{\mu\nu}] = \frac{i}{2} [\gamma^\kappa \gamma^\lambda, S^{\mu\nu}] = ig^{\lambda\mu} S^{\kappa\nu} - ig^{\lambda\nu} S^{\kappa\mu} - ig^{\kappa\mu} S^{\lambda\nu} + ig^{\kappa\nu} S^{\lambda\mu}.
\]  
(16)

*Quod erat demonstrandum.*

In light of this theorem, the $S^{\mu\nu}$ matrices represent the Lorentz generators \(J^{\mu\nu}\) in the 4-component spinor multiplet.
Finite Lorentz transforms:
Any continuous Lorentz transform — a rotation, or a boost, or a product of a boost and a rotation — obtains from exponentiating an infinitesimal symmetry

\[ X'\mu = X\mu + \epsilon^{\mu\nu}X_{\nu} \]  

(17)

where the infinitesimal \( \epsilon^{\mu\nu} \) matrix is antisymmetric when both indices are raised (or both lowered), \( \epsilon^{\mu\nu} = -\epsilon^{\nu\mu} \). Thus, the \( L_{\mu}^\nu \) matrix of any continuous Lorentz transform is a matrix exponential

\[ L_{\nu}^\mu = \exp(\Theta)_{\nu}^\mu \equiv \delta_{\nu}^\mu + \Theta_{\nu}^\mu + \frac{1}{2}\Theta_{\lambda}^\mu \Theta_{\nu}^\lambda + \frac{1}{6}\Theta_{\lambda}^\mu \Theta_{\kappa}^\lambda \Theta_{\nu}^\kappa + \cdots \]  

(18)

of some matrix \( \Theta \) that becomes antisymmetric when both of its indices are raised or lowered, \( \Theta^{\mu\nu} = -\Theta^{\nu\mu} \). Note however that in the matrix exponential (18), the first index of \( \Theta \) is raised while the second index is lowered, so the antisymmetry condition becomes \( (g \Theta)^T = -(g \Theta) \) instead of \( \Theta^T = -\Theta \).

The Dirac spinor representation of the finite Lorentz transform (18) is the \( 4 \times 4 \) matrix

\[ M_D(L) = \exp(-i\theta_{\alpha\beta}S^{\alpha\beta}). \]  

(19)

The group law for such matrices

\[ \forall L_1, L_2 \in SO^+(3,1), \quad M_D(L_2 L_1) = M_D(L_2)M_D(L_1) \]  

(20)

follows automatically from the \( S^{\mu\nu} \) satisfying the commutation relations (14) of the Lorentz generators, so I am not going to prove it. Instead, let me show that when the Dirac matrices \( \gamma^{\mu} \) are sandwiched between the \( M_D(L) \) and its inverse, they transform into each other as components of a Lorentz 4–vector,

\[ M_D^{-1}(L)\gamma^{\mu}M_D(L) = L_{\nu}^\mu\gamma^{\nu}. \]  

(21)

This formula makes the Dirac equation transform covariantly under the Lorentz transforms.
Proof: In light of the exponential form (19) of the matrix $M_D(L)$ representing a finite Lorentz transform in the Dirac spinor multiplet, let’s use the multiple commutator formula (AKA the Hadamard Lemma): for any 2 matrices $F$ and $H$,

$$\exp(-F)H\exp(+F) = H + [H, F] + \frac{1}{2} [[H, F], F] + \frac{1}{6} [[[H, F], F], F] + \cdots.$$ (22)

In particular, let $H = \gamma^\mu$ while $F = -\frac{i}{2} \Theta_{\alpha\beta} S^{\alpha\beta}$ so that $M_D(L) = \exp(+F)$ and $M_D^{-1}(L) = \exp(-F)$. Consequently,

$$M_D^{-1}(L)\gamma^\mu M_D(L) = \gamma^\mu + [\gamma^\mu, F] + \frac{1}{2} [[\gamma^\mu, F], F] + \frac{1}{6} [[[\gamma^\mu, F], F], F] + \cdots.$$ (23)

where all the multiple commutators turn out to be linear combinations of the Dirac matrices. Indeed, the single commutator here is

$$[\gamma^\mu, F] = -\frac{i}{2} \Theta_{\alpha\beta} [\gamma^\mu, S^{\alpha\beta}] = -\frac{1}{2} \Theta_{\alpha\beta} (g^{\mu\alpha} \gamma^\beta - g^{\mu\beta} \gamma^\alpha) = \Theta_{\alpha\beta} g^{\mu\alpha} \gamma^\beta = \Theta^\mu_{\lambda} \gamma^\lambda,$$ (24)

while the multiple commutators follow by iterating this formula:

$$[[\gamma^\mu, F], F] = \Theta^\mu_{\lambda} [\gamma^\lambda, F] = \Theta^\mu_{\lambda} \Theta^\lambda_{\nu} \gamma^\nu, \quad [[[\gamma^\mu, F], F], F] = \Theta^\mu_{\lambda} \Theta^\lambda_{\rho} \Theta^\rho_{\nu} \gamma^\nu, \ldots.$$ (25)

Combining all these commutators as in eq. (23), we obtain

$$M_D^{-1}\gamma^\mu M_D = \gamma^\mu + [\gamma^\mu, F] + \frac{1}{2} [[\gamma^\mu, F], F] + \frac{1}{6} [[[\gamma^\mu, F], F], F] + \cdots$$

$$= \gamma^\mu + \Theta^\mu_{\nu} \gamma^\nu + \frac{1}{2} \Theta^\mu_{\lambda} \Theta^\lambda_{\nu} \gamma^\nu + \frac{1}{6} \Theta^\mu_{\lambda} \Theta^\lambda_{\rho} \Theta^\rho_{\nu} \gamma^\nu + \cdots$$

$$= \left(\delta^\mu_{\nu} + \Theta^\mu_{\nu} + \frac{1}{2} \Theta^\mu_{\lambda} \Theta^\lambda_{\nu} + \frac{1}{6} \Theta^\mu_{\lambda} \Theta^\lambda_{\rho} \Theta^\rho_{\nu} + \cdots\right)\gamma^\nu$$

$$\equiv L^\mu_{\nu} \gamma^\nu.$$ (26)

*Quod erat demonstrandum.*
Dirac Equation and Dirac Spinor Fields

History:
Originally, the Klein–Gordon equation was thought to be the relativistic version of the Schrödinger equation — that is, an equation for the wave function $\psi(x, t)$ for one relativistic particle. But pretty soon this interpretation run into trouble with bad probabilities (negative, or $> 1$) when a particle travels through high potential barriers or deep potential wells. There were also troubles with relativistic causality, and a few other things.

Paul Adrien Maurice Dirac had thought that the source of all those troubles was the ugly form of relativistic Hamiltonian $\hat{H} = \sqrt{\hat{p}^2 + m^2}$ in the coordinate basis, and that he could solve all the problems with the Klein-Gordon equation by rewriting the Hamiltonian as a first-order differential operator

$$\hat{H} = \hat{p} \cdot \vec{\alpha} + m \beta \implies \text{Dirac equation } i \frac{\partial \psi}{\partial t} = -i \vec{\alpha} \cdot \nabla \psi + m \beta \psi \quad (27)$$

where $\alpha_1, \alpha_2, \alpha_3, \beta$ are matrices acting on a multi-component wave function. Specifically, all four of these matrices are Hermitian, square to 1, and anticommute with each other,

$$\{\alpha_i, \alpha_j\} = 2 \delta_{ij}, \quad \{\alpha_i, \beta\} = 0, \quad \beta^2 = 1. \quad (28)$$

Consequently

$$(\vec{\alpha} \cdot \hat{p})^2 = \alpha_i \alpha_j \times \hat{p}_i \hat{p}_j = \frac{1}{2} \{\alpha_i, \alpha_j\} \times \hat{p}_i \hat{p}_j = \delta_{ij} \times \hat{p}_i \hat{p}_j = \hat{p}^2, \quad (29)$$

and therefore

$$\hat{H}_{\text{Dirac}}^2 = (\vec{\alpha} \cdot \hat{p} + \beta m)^2 = (\vec{\alpha} \cdot \hat{p})^2 + \{\alpha_i, \beta\} \times \hat{p}_i m + \beta^2 \times m^2 = \hat{p}^2 + 0 + m^2. \quad (30)$$

This, the Dirac Hamiltonian squares to $\hat{p}^2 + m^2$, as it should for the relativistic particle.

The Dirac equation (27) turned out to be a much better description of a relativistic electron (which has spin $= \frac{1}{2}$) than the Klein–Gordon equation. However, it did not resolve the troubles with relativistic causality or bad probabilities for electrons going through big potential differences $e \Delta \Phi > 2m_e c^2$. Those problems are not solvable in the context of a relativistic single-particle quantum mechanics but only in the quantum field theory.
Modern point of view:

Today, we interpret the Dirac equation as the equation of motion for a Dirac spinor field \( \Psi(x) \), comprising 4 complex component fields \( \Psi_\alpha(x) \) arranged in a column vector

\[
\Psi(x) = \begin{pmatrix}
\Psi_1(x) \\
\Psi_2(x) \\
\Psi_3(x) \\
\Psi_4(x)
\end{pmatrix},
\]

and transforming under the continuous Lorentz symmetries \( x'\mu = L_\nu x^\nu \) according to

\[
\Psi'(x') = M_D(L)\Psi(x).
\]

The classical Euler–Lagrange equation of motion for the spinor field is the Dirac equation

\[
i \frac{\partial}{\partial t} \Psi + i\vec{\alpha} \cdot \nabla \Psi - m\beta \Psi = 0.
\]

To recast this equation in a Lorentz-covariant form, let

\[
\beta = \gamma^0, \quad \alpha^i = \gamma^0 \gamma^i;
\]

it is easy to see that if the \( \gamma^\mu \) matrices obey the anticommutation relations (5) then the \( \vec{\alpha} \) and \( \beta \) matrices obey the relations (28) and vice versa. Now let’s multiply the whole LHS of the Dirac equation (33) by the \( \beta = \gamma^0 \):

\[
0 = \gamma^0 \left( i\partial_0 + i\gamma^0 \vec{\gamma} \cdot \nabla - m\gamma^0 \right) \Psi(x) = \left( i\gamma^0 \partial_0 + i\gamma^i \partial_i - m \right) \Psi(x),
\]

and hence

\[
\left( i\gamma^\mu \partial_\mu - m \right) \Psi(x) = 0.
\]

As expected from \( \hat{H}_{\text{Dirac}}^2 = \hat{p}^2 + m^2 \), the Dirac equation for the spinor field implies the Klein–Gordon equation for each component \( \Psi_\alpha(x) \). Indeed, if \( \Psi(x) \) obey the Dirac equation,
then obviously
\[(−iγ^\nu\partial_\nu − m) \times (iγ^\mu\partial_\mu − m)Ψ(x) = 0, \quad (37)\]
buts the differential operator on the LHS is equal to the Klein–Gordon \(m^2 + \partial^2\) times a unit matrix:
\[(−iγ^\nu\partial_\nu − m)(iγ^\mu\partial_\mu − m) = m^2 + γ^\nuγ^\mu\partial_\nu\partial_\mu = m^2 + γ^\nuγ^\mu\partial_\nu\partial_\mu. \quad (38)\]

**The Dirac equation** (36) **transforms covariantly under the Lorentz symmetries** — its LHS transforms exactly like the spinor field itself.

**Proof**: Note that since the Lorentz symmetries involve the \(x^\mu\) coordinates as well as the spinor field components, the LHS of the Dirac equation becomes
\[(iγ^\mu\partial'_\mu − m)Ψ'(x') \quad (39)\]
where
\[\partial'_\mu \equiv \frac{\partial}{\partial x'_\mu} = \frac{\partial x^\nu}{\partial x'_\mu} \times \frac{\partial}{\partial x^\nu} = (L^{-1})^\nu_\mu × \partial_\nu. \quad (40)\]
Consequently,
\[\partial'_\mu Ψ'(x') = (L^{-1})^\nu_\mu × M_D(L) \partial_\nu Ψ(x) \quad (41)\]
and hence
\[γ^\mu\partial'_\mu Ψ'(x') = (L^{-1})^\nu_\mu × γ^\mu M_D(L) \partial_\nu Ψ(x). \quad (42)\]
But according to eq. (23),
\[M_D^{-1}(L)γ^\mu M_D(L) = L^\mu_\nu γ^\nu \quad \Rightarrow \quad γ^\mu M_D(L) = L^\mu_\nu × M_D(L)γ^\nu \quad \Rightarrow \quad (L^{-1})^\nu_\mu × γ^\mu M_D(L) = M_D(L)γ^\nu, \quad (43)\]
so
\[γ^\mu\partial'_\mu Ψ'(x') = M_D(L) × γ^\nu \partial_\nu Ψ(x). \quad (44)\]
Altogether,
\[(iγ^\mu\partial_\mu − m)Ψ(x) \xrightarrow{\text{Lorentz}} (iγ^\mu\partial'_\mu − m)Ψ'(x') = M_D(L) × (iγ^\mu\partial_\mu − m)Ψ(x), \quad (45)\]
which proves the covariance of the Dirac equation. *Quod erat demonstrandum.*
Dirac Lagrangian

The Dirac equation is a first-order differential equation, so to obtain it as an Euler–Lagrange equation, we need a Lagrangian which is linear rather than quadratic in the spinor field’s derivatives. Thus, we want

\[ \mathcal{L} = \overline{\Psi} \times \left( i \gamma^\mu \partial_\mu - m \right) \Psi \]  

(46)

where \( \overline{\Psi}(x) \) is some kind of a conjugate field to the \( \Psi(x) \). Since \( \Psi \) is a complex field, we treat \( \Psi \) and \( \overline{\Psi} \) as linearly-independent from each other, so the Euler–Lagrange equation for the \( \overline{\Psi} \) immediately gives us the Dirac equation for the \( \Psi(x) \) field,

\[ 0 = \frac{\partial \mathcal{L}}{\partial \Psi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \Psi)} = (i \gamma^\nu \partial_\nu - m) \Psi - \partial_\mu (0). \]  

(47)

To keep the action \( S = \int d^4x \mathcal{L} \) Lorentz-invariant, the Lagrangian (46) should transform as a Lorentz scalar, \( \mathcal{L}'(x') = \mathcal{L}(x) \). In light of eq. (19) for the \( \Psi(x) \) field and covariance (45) of the Dirac equation, the conjugate field \( \overline{\Psi}(x) \) should transform according to

\[ \overline{\Psi}'(x') = \overline{\Psi}(x) \times M_D^{-1}(L) \implies \mathcal{L}'(x') = \mathcal{L}(x). \]  

(48)

Note that the \( M_D(L) \) matrix is generally not unitary, so the inverse matrix \( M_D^{-1}(L) \) in eq. (48) is different from the hermitian conjugate \( M_D^\dagger(L) \). Consequently, the conjugate field \( \overline{\Psi}(x) \) cannot be identified with the hermitian conjugate field \( \Psi^\dagger(x) \), since the latter transforms to

\[ \Psi^\dagger(x') = \Psi^\dagger(x) \times M_D^\dagger(L) \neq \Psi^\dagger(x) \times M_D^{-1}(L). \]  

(49)

Instead of the hermitian conjugate, we are going to use the Dirac conjugate spinor, see below.

**Dirac conjugates:**

Let \( \Psi \) be a 4-component Dirac spinor and \( \Gamma \) be any \( 4 \times 4 \) matrix; we *define* their Dirac conjugates according to

\[ \overline{\Psi} = \Psi^\dagger \times \gamma^0, \quad \overline{\Gamma} = \gamma^0 \times \Gamma^\dagger \times \gamma^0. \]  

(50)

Thanks to \( \gamma^0 \gamma^0 = 1 \), the Dirac conjugates behave similarly to hermitian conjugates or transposed matrices:
- For a product of 2 matrices, \( (\Gamma_1 \times \Gamma_2) = \Gamma_2 \times \Gamma_1 \).
- Likewise, for a matrix and a spinor, \( (\Gamma \times \Psi) = \overline{\Psi} \times \Gamma \).
- The Dirac conjugate of a complex number is its complex conjugate, \( (c \times 1) = c^* \times 1 \).
- For any two spinors \( \Psi_1 \) and \( \Psi_2 \) and any matrix \( \Gamma \), \( \Psi_1 \Gamma \Psi_2 = (\overline{\Psi}_2 \Gamma \Psi_1)^* \).

  - The Dirac spinor fields are fermionic, so they anticommute with each other, even in the classical limit. Nevertheless, \( (\Psi_\alpha \Psi_\beta)^\dagger = +\Psi_\beta \Psi_\alpha \), and therefore for any matrix \( \Gamma \), \( \overline{\Psi}_1 \Gamma \Psi_2 = + (\overline{\Psi}_2 \Gamma \Psi_1)^* \).

The point of the Dirac conjugation (50) is that it works similarly for all four Dirac matrices \( \gamma^\mu \),

\[
\overline{\gamma}^\mu = +\gamma^\mu. \tag{51}
\]

**Proof:** For \( \mu = 0 \), the \( \gamma^0 \) is hermitian, hence

\[
\overline{\gamma}^0 = \gamma^0 (\gamma^0)^\dagger \gamma^0 = +\gamma^0 \gamma^0 \gamma^0 = +\gamma^0. \tag{52}
\]

For \( \mu = i = 1, 2, 3 \), the \( \gamma^i \) are anti-hermitian and also anticommute with the \( \gamma^0 \), hence

\[
\overline{\gamma}^i = \gamma^0 (\gamma^i)^\dagger \gamma^0 = -\gamma^0 \gamma^i \gamma^0 = +\gamma^0 \gamma^0 \gamma^i = +\gamma^i. \tag{53}
\]

**Corollary:** The Dirac conjugate of the matrix

\[
M_D(L) = \exp(-\frac{i}{2} \Theta_{\mu\nu} S^{\mu\nu}) \tag{19}
\]

representing any continuous Lorentz symmetry \( L = \exp(\Theta) \) is the inverse matrix

\[
\overline{M}_D(L) = M_D^{-1}(L) = \exp(+\frac{i}{2} \Theta_{\mu\nu} S^{\mu\nu}). \tag{54}
\]

**Proof:** Let

\[
X = -\frac{i}{2} \Theta_{\mu\nu} S^{\mu\nu} = +\frac{1}{8} \Theta_{\mu\nu} [\gamma^\mu, \gamma^\nu] = +\frac{1}{3} \Theta_{\mu\nu} \gamma^\mu \gamma^\nu \tag{55}
\]

for some real antisymmetric Lorentz parameters \( \Theta_{\mu\nu} = -\Theta_{\nu\mu} \). The Dirac conjugate of the
\( X \) matrix is
\[
\overline{X} = \frac{1}{4} \Theta_{\mu\nu} \gamma^\mu \gamma^\nu = \frac{1}{4} \Theta^{\mu} \gamma_\mu \gamma^\nu = \frac{1}{4} \Theta_{\nu\mu} \gamma^\nu \gamma^\mu = -\frac{1}{4} \Theta_{\mu\nu} \gamma^\mu \gamma^\nu = -X. \quad (56)
\]
Consequently,
\[
\overline{X}^2 = \overline{X} \times \overline{X} = +X^2, \quad \overline{X}^3 = \overline{X} \times \overline{X}^2 = \overline{X}^2 \times \overline{X} = -X^3, \quad \ldots, \quad \overline{X}^n = (-X)^n,
\]
and hence
\[
\exp(\overline{X}) = \sum_n \frac{1}{n!} \overline{X}^n = \sum_n \frac{1}{n!} (-X)^n = \exp(-X). \quad (58)
\]
In light of eq. (55), this means
\[
\exp(-\frac{i}{2} \Theta_{\mu\nu} S_{\mu\nu}) = \exp(+\frac{i}{2} \Theta_{\mu\nu} S_{\mu\nu}),
\]
that is,
\[
\overline{M}_D(L) = M_D^{-1}(L). \quad (60)
\]
*Quod erat demonstrandum.*

**Back to the Dirac Lagrangian:**
Thanks to the theorem (60), the conjugate field \( \overline{\Psi}(x) \) in the Lagrangian (46) is simply the Dirac conjugate of the Dirac spinor field \( \Psi(x) \),
\[
\overline{\Psi}(x) = \Psi^\dagger(x) \times \gamma^0, \quad (61)
\]
which transforms under Lorentz symmetries as
\[
\overline{\Psi}'(x') = \overline{\Psi}'(x') = \overline{M}_D(L) \times \overline{\Psi}(x) = \overline{\Psi}(x) \times \overline{M}_D(x) = \overline{\Psi}(x) \times M_D^{-1}(L). \quad (62)
\]
Consequently, the Dirac Lagrangian
\[
\mathcal{L} = \overline{\Psi} \times (i \gamma^\mu \partial_\mu - m) \Psi = \Psi^\dagger \gamma^0 \times (i \gamma^\mu \partial_\mu - m) \Psi \quad (46)
\]
is a Lorentz scalar and the action is Lorentz invariant.
Hamiltonian for the Dirac Field

Canonical quantization of the Dirac spinor field $\Psi(x)$ — just like any other field — begins with the classical Hamiltonian formalism. Let’s start with the canonical conjugate fields,

$$\Pi_\alpha = \frac{\partial \mathcal{L}}{\partial (\partial_0 \Psi_\alpha)} = (i\overline{\Psi}\gamma^0)_\alpha = i\Psi^\dagger_\alpha$$

— the canonical conjugate to the Dirac spinor field $\Psi(x)$ is simply its hermitian conjugate $\Psi^\dagger(x)$. This is similar to what we had for the non-relativistic field, and it happens for the same reason — the Lagrangian which is linear in the time derivative.

In the non-relativistic field theory, the conjugacy relation (63) in the classical theory lead to the equal-time commutation relations in the quantum theory,

$$[\hat{\psi}(x,t), \hat{\psi}(y,t)] = 0, \quad [\hat{\psi}^\dagger(x,t), \hat{\psi}^\dagger(y,t)] = 0, \quad [\hat{\psi}(x,t), \hat{\psi}^\dagger(y,t)] = \delta^{(3)}(x - y). \quad (64)$$

However, the Dirac spinor field describes spin = $\frac{1}{2}$ particles — like electrons, protons, or neutrons — which are fermions rather than bosons. So instead of the commutations relations (64), the spinor fields obey the equal-time anti-commutation relations

$$\begin{align*}
\{\hat{\Psi}_\alpha(x,t), \hat{\Psi}_\beta(y,t)\} &= 0, \\
\{\hat{\Psi}^\dagger_\alpha(x,t), \hat{\Psi}^\dagger_\beta(y,t)\} &= 0, \\
\{\hat{\Psi}_\alpha(x,t), \hat{\Psi}^\dagger_\beta(y,t)\} &= \delta_{\alpha\beta}\delta^{(3)}(x - y). \quad (65)
\end{align*}$$

Next, the classical Hamiltonian obtains as

$$\begin{align*}
H &= \int d^3x \mathcal{H}(x), \\
\mathcal{H} &= i\Psi^\dagger \partial_0 \Psi - \mathcal{L} \\
&= i\Psi^\dagger \partial_0 \Psi - \Psi^\dagger \gamma^0(\gamma^0\partial_0 + \gamma^0 \cdot \nabla - m) \Psi \\
&= \Psi^\dagger(-i\gamma^0 \gamma^0 \cdot \nabla + \gamma^0 m) \Psi
\end{align*} \quad (66)$$

where the terms involving the time derivative $\partial_0$ cancel out. Consequently, the Hamiltonian
operator of the quantum field theory is

\[ \hat{H} = \int d^3x \hat{\Psi}^\dagger(x)(-i\gamma^0 \vec{\gamma} \cdot \nabla + \gamma^0 m) \hat{\Psi}(x). \]  \tag{67}

Note that the derivative operator \((-i\gamma^0 \vec{\gamma} \cdot \nabla + \gamma^0 m)\) in this formula is precisely the 1-particle Dirac Hamiltonian (27). This is very similar to what we had for the quantum non-relativistic fields,

\[ \hat{H} = \int d^3x \hat{\psi}^\dagger(x) \left( -\frac{1}{2M} \nabla^2 + V(x) \right) \hat{\psi}(x), \]  \tag{68}

except for a different differential operator, Schrödinger instead of Dirac.

In the Heisenberg picture, the quantum Dirac field obeys the Dirac equation. To see how this works, we start with the Heisenberg equation

\[ i \frac{\partial}{\partial t} \hat{\Psi}_\alpha(x,t) = [\hat{\Psi}_\alpha(x,t), \hat{H}] = \int d^3y [\hat{\Psi}_\alpha(x,t), \hat{H}(y,t)], \]  \tag{69}

and then evaluate the last commutator using the anti-commutation relations (65) and the Leibniz rules (11). Indeed, let’s use the Leibniz rule

\[ [A, BC] = \{A, B\} C - B\{A, C\} \]  \tag{70}

for

\begin{align*}
A &= \hat{\Psi}_\alpha(x,t), \\
B &= \hat{\Psi}_\beta^\dagger(y,t), \\
C &= (-i\gamma^0 \vec{\gamma} \cdot \nabla + \gamma^0 m)_{\beta\gamma} \hat{\Psi}_\gamma(y,t),
\end{align*}  \tag{71}

so that \(BC = \hat{H}(y,t)\). For the \(A, B, C\) at hand,

\[ \{A, B\} = \delta_{\alpha\beta}\delta^{(3)}(x-y) \]  \tag{72}

while

\[ \{A, C\} = (-i\gamma^0 \vec{\gamma} \cdot \nabla_y + \gamma^0 m)_{\beta\gamma} \{\hat{\Psi}_\alpha(x,t), \hat{\Psi}_\gamma(y,t)\} = (\text{diff.op.}) \times 0 = 0. \]  \tag{73}
Consequently

\[
\left[ \hat{\Psi}_\alpha(x, t), \hat{\mathcal{H}}(y, t) \right] \equiv [A, BC] = \{A, B\} \times C - B \times \{A, C\}
\]

\[= \delta_{\alpha\beta}\delta^{(3)}(x - y) \times (-i\gamma^0 \vec{\gamma} \cdot \nabla + \gamma^0 m)_{\beta\gamma} \hat{\Psi}_\gamma(y, t) - 0,\]

(74)

hence

\[
\left[ \hat{\Psi}_\alpha(x, t), \hat{H} \right] = \int d^3 y \delta^{(3)}(x - y) \times (-i\gamma^0 \vec{\gamma} \cdot \nabla + \gamma^0 m)_{\alpha\gamma} \hat{\Psi}_\gamma(y, t)
\]

\[= (-i\gamma^0 \vec{\gamma} \cdot \nabla + \gamma^0 m)_{\alpha\gamma} \hat{\Psi}_\gamma(x, t),\]

(75)

and therefore

\[i\partial_0 \hat{\Psi}(x, t) = (-i\gamma^0 \vec{\gamma} \cdot \nabla + \gamma^0 m)\hat{\Psi}(x, t).\]

(76)

Or if you prefer,

\[(i\gamma^\mu \partial_\mu - m)\hat{\Psi}(x) = 0.\]

(77)