Introducing the Dirac Equation

Bill Celmaster

February 23, 2021

1 An extremely brief history

Physicists knew that the Klein-Gordon equation was Lorentz invariant:

$$\left(\partial_0^2 - \partial_1^2 - \partial_2^2 - \partial_3^2 + m^2\right)\phi(x).$$

However, Dirac wanted (for various reasons) wanted a Lorentz-invariant equation linear in derivatives. If there weren't all those terms, you could take the square-root.

If one only knew how to take a square root, then it would seem obvious that the following expression would also be Lorentz invariant.

$$\sqrt{(\partial_0^2 - \partial_1^2 - \partial_2^2 - \partial_3^2 + m^2)}\phi(x).$$

Dirac figured out a way to take the square root (sort of).

2 Outline

- Defining and explaining the Dirac equation for 2-spinors and 4-spinors
- Showing that the Dirac equation for 2-spinors is Lorentz invariant and for 4-spinors is also parity invariant
- Deriving the Dirac equation from a Lagrangian
- Deriving the Lagrangian from symmetry representation theory
  - Go over solutions to the Lorentz symmetry exercises
  - Show how these can be used to hypothesize a Lagrangian
- Quantizing the Lagrangian theory – anticommutation relations

- Solving the Dirac equation
  - Particles and antiparticles

3 Defining and explaining the Dirac equation

I will follow Lancaster chapter 36 but in a different order to better connect (later) to the formal symmetry theory we've been looking at.

We will constantly use the Pauli spin matrices so here they are to remind you.

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1}$$

and write $\boldsymbol{\sigma}$ as the 3-tuple $(\sigma_1, \sigma_2, \sigma_3)$.

Note: I will use the term "spinor". A spinor is just a vector which transforms a certain way under a representation of the Lorentz group. In this section, we don't care about the transformation properties so just think of spinors as vectors.

3.1 2-spinors

The 2-spinor equations don't have mass terms. There are 2 distinct Dirac[1] equations for 2-spinors (or if you prefer, 2-vectors).

- $i(\mathbf{I}_2\partial_0 + \boldsymbol{\sigma} \cdot \boldsymbol{\nabla})\psi_R(x) = 0$.

  Unpack this and then compare to Lancaster eq. (36.14).

  - Recall that $\partial_0 \equiv \frac{\partial}{\partial x^0}$, and $\boldsymbol{\nabla}$ is the vector $(\partial_1, \partial_2, \partial_3) = (\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3})$.
  - The symbol $\mathbf{I}_2$ means the matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. [Sometimes that symbol is dropped for convenience, resulting in an equation that looks like $i(\partial_0 - \boldsymbol{\sigma} \cdot \boldsymbol{\nabla})\psi_R(x) = 0$.]
  - Another notational convention (which I don't love) is to replace the symbol $i\partial_0$ by $\hat{p}_0$ and $-i\boldsymbol{\nabla}$ by $\hat{\mathbf{p}}$. Then our equation looks like $(\hat{p}_0 - \boldsymbol{\sigma} \cdot \hat{\mathbf{p}})\psi_R(x) = 0$ (which is exactly Lancaster eq. (36.14)).

[1] I refer to all these kinds of equations as "Dirac" equations, although Dirac's original equation was only for 4-spinors, sometimes known as Dirac spinors. These 2-spinor equations are often known as Weyl equations and the 2-spinors are known as Weyl spinors.

- Write everything out in components.

$$i\left(\begin{pmatrix}\partial_0 & 0\\ 0 & \partial_0\end{pmatrix}+\begin{pmatrix}0 & \partial_1\\ \partial_1 & 0\end{pmatrix}+\begin{pmatrix}0 & -i\partial_2\\ i\partial_2 & 0\end{pmatrix}+\begin{pmatrix}\partial_3 & 0\\ 0 & -\partial_3\end{pmatrix}\right)\begin{pmatrix}\psi_R^1(x)\\ \psi_R^2(x)\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}.$$

- Multiply out the matrices to get

$$\begin{pmatrix}(i\partial_0+i\partial_3)\psi_R^1(x)+(i\partial_1+\partial_2)\psi_R^2(x)\\ (i\partial_0-i\partial_3)\psi_R^2(x)+(i\partial_1-\partial_2)\psi_R^1(x)\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}. \qquad (2)$$

These are two linear differential equations in two unknowns.

- $i(\mathbf{I}_2\partial_0-\boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\psi_L(x)=0.$

  Just like the prior equation, we can expand it out to get

$$i\left(\begin{pmatrix}\partial_0 & 0\\ 0 & \partial_0\end{pmatrix}-\begin{pmatrix}0 & \partial_1\\ \partial_1 & 0\end{pmatrix}-\begin{pmatrix}0 & -i\partial_2\\ i\partial_2 & 0\end{pmatrix}-\begin{pmatrix}\partial_3 & 0\\ 0 & -\partial_3\end{pmatrix}\right)\begin{pmatrix}\psi_L^1(x)\\ \psi_L^2(x)\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}.$$

  and then multiplying out the matrices, obtaining

$$\begin{pmatrix}(i\partial_0-i\partial_3)\psi_L^1(x)-(i\partial_1+\partial_2)\psi_L^2(x)\\ (i\partial_0+i\partial_3)\psi_L^2(x)-(i\partial_1-\partial_2)\psi_L^1(x)\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}. \qquad (3)$$

3.2 4-spinors

The simplest 4-spinor equation will look like 2 copies of the 2-spinor equations, and will also lack mass terms. We will also consider a 4-spinor equation with a mass term that mixes ψ_L and ψ_R. This equation is know as the "massive Dirac equation". In the next section, we'll investigate the question of why mass terms shouldn't be added to the 2-spinor equations.

- The simplest 4-spinor equation.

$$\begin{pmatrix}i(\mathbf{I}_2\partial_0+\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}) & 0\\ 0 & i(\mathbf{I}_2\partial_0-\boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\end{pmatrix}\begin{pmatrix}\psi_R(x)\\ \psi_L(x)\end{pmatrix}=\mathbf{0}. \qquad (4)$$

  Each block of the matrix is 2 x 2 and each entry in the vector is a 2-spinor. Therefore the matrix is 4 x 4 and the vector is of length 4. This 4-spinor equation has exactly the same content as the two 2-spinor equations in the last section so we think of it as *reducible.*

- 4-spinor equation with mass term.

$$\begin{pmatrix}i(\mathbf{I}_2\partial_0+\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}) & -m\mathbf{I}_2\\ -m\mathbf{I}_2 & i(\mathbf{I}_2\partial_0-\boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\end{pmatrix}\begin{pmatrix}\psi_R(x)\\ \psi_L(x)\end{pmatrix}=\mathbf{0}. \qquad (5)$$

Multiply out, we get

$$\begin{pmatrix} i(\mathbf{I}_2\partial_0 + \boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\psi_R(x) - m\psi_L(x) \\ i(\mathbf{I}_2\partial_0 - \boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\psi_L(x) - m\psi_R(x) \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}. \tag{6}$$

We see that the top and bottom are two separate equations which mix ψ_L and ψ_R. Compare to Lancaster eq. (36.14)

This is a version of the massive Dirac equation. However, people have invented notation to make things look neater.

- First rewritten version of the Dirac equation.

$$i\left[\begin{pmatrix} \mathbf{I}_2 & 0 \\ 0 & \mathbf{I}_2 \end{pmatrix}\partial_0 + \begin{pmatrix} \sigma_1 & 0 \\ 0 & -\sigma_1 \end{pmatrix}\partial_1 + \begin{pmatrix} \sigma_2 & 0 \\ 0 & -\sigma_2 \end{pmatrix}\partial_2 + \begin{pmatrix} \sigma_3 & 0 \\ 0 & -\sigma_3 \end{pmatrix}\partial_3 - \begin{pmatrix} 0 & m\mathbf{I}_2 \\ m\mathbf{I}_2 & 0 \end{pmatrix}\right]\begin{pmatrix} \psi_R(x) \\ \psi_L(x) \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}.$$

- Standard form of the Dirac equation (in the Weyl representation).

  First rewrite the above equation by inserting before $\begin{pmatrix} \psi_R(x) \\ \psi_L(x) \end{pmatrix}$, the matrix product $\mathbf{I}_4 = \begin{pmatrix} 0 & \mathbf{I}_2 \\ \mathbf{I}_2 & 0 \end{pmatrix}\begin{pmatrix} 0 & \mathbf{I}_2 \\ \mathbf{I}_2 & 0 \end{pmatrix}$. Then the left-most of those 2 matrices in the product should be multiplied (on the right) by all the terms that precede it, and the right-most of those 2 matrices should be multiplied by $\begin{pmatrix} \psi_R(x) \\ \psi_L(x) \end{pmatrix}$, resulting in $\begin{pmatrix} \psi_L(x) \\ \psi_R(x) \end{pmatrix}$. The end result of these manipulations is

$$i\left[\begin{pmatrix} 0 & \mathbf{I}_2 \\ \mathbf{I}_2 & 0 \end{pmatrix}\partial_0 + \begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix}\partial_1 + \begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix}\partial_2 + \begin{pmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{pmatrix}\partial_3\right]\begin{pmatrix} \psi_L(x) \\ \psi_R(x) \end{pmatrix} = \begin{pmatrix} m\mathbf{I}_2 & 0 \\ 0 & m\mathbf{I}_2 \end{pmatrix}\begin{pmatrix} \psi_L(x) \\ \psi_R(x) \end{pmatrix}. \tag{7}$$

  Compare Lancaster eq. (36.13).

- Gamma matrix notation.

  The equation can be simplified by defining some new matrices,

$$\gamma^0 = \begin{pmatrix} 0 & \mathbf{I}_2 \\ \mathbf{I}_2 & 0 \end{pmatrix}, \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \tag{8}$$

and also defining $\psi(x) = \begin{pmatrix} \psi_L(x) \\ \psi_R(x) \end{pmatrix}$. Compare the γ matrices with Lancaster eqs. (36.9), (36.10) and notice also that Lancaster writes out all 4x4 components in eq. (36.8).

Then eq. (7) becomes the *canonical* Dirac equation.

$$\boldsymbol{i\gamma^\mu \partial_\mu \psi(x) = m\psi(x)}.$$

Compare Lancaster eq. (36.12).

The γ matrices are technically called "Dirac matrices in the Weyl representation". An even more streamlined notation is attributed to Feynman, using the Feynman slash notation

$$\partial\!\!\!/ \equiv \gamma^\mu \partial_\mu.$$

In that notation, the Dirac equation becomes

$$\boldsymbol{\left(i\partial\!\!\!/ - m\right) \psi(x) = 0.}$$

- Dirac algebra.

  By direct computation, you can show that

  $$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}$$

  where $g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$.

  Compare Lancaster eq. (36.3). These relations are known as the Dirac algebra, which is a special case of a *Clifford algebra*.

- Other representations.

  If you make a similarity transform on the γ matrices, that will preserve the Dirac algebra but will change the form of the Dirac matrices. If you define $\gamma'^\mu = \mathbf{S}\gamma^\mu\mathbf{S}^{-1}$ and $\psi'(x) = \mathbf{S}\psi(x)$, then it's easy to show (go ahead, do it) that

  $$i(\partial\!\!\!/' - m)\psi'(x) = 0$$

  where $\partial\!\!\!/'$ is defined with γ'^μ instead of γ^μ.

4 Lorentz invariance

If we can find a linear transformation on ψ_L etc., which preserves the form of the Dirac equation under a change of reference frame, then we say the Dirac equation is Lorentz invariant. **To simplify things, only look at rotations and boosts around the z-axis.**

4.1 2-spinors

4.1.1 ψ_R equation

- Start with $i(\mathbf{I}_2\partial_0 + \boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\psi_R(\mathbf{x}) = 0$.
- Propose transformation laws:

  z-rotation by θ:

$$\psi_R(\mathbf{x}') = \begin{pmatrix} e^{i\frac{\theta}{2}}\psi_R'^1 \\ e^{-i\frac{\theta}{2}}\psi_R'^2 \end{pmatrix}(\mathbf{x}').$$

  where $\mathbf{x}' = (t', x', y, z') = (t, x\cos\theta - y\sin\theta, x\sin\theta + ycos\theta, z)$.

  z-boost by β:

$$\psi_R(\mathbf{x}'') = \begin{pmatrix} e^{-\frac{\beta}{2}}\psi_R''^1 \\ e^{\frac{\beta}{2}}\psi_R''^2 \end{pmatrix}(\mathbf{x}'').$$

  where $\mathbf{x}'' = (t'', x'', y'', z'') = (t\cosh\beta + z\sinh\beta, x, y, t\sinh\beta + z\cosh\beta)$.

- Apply chain rule as we've done numerous times before. For example:

$$\begin{aligned}\partial_1\psi_R(\mathbf{x}) = \frac{\partial}{\partial x}\psi_R(\mathbf{x}) &= \left(\frac{\partial x'}{\partial x}\frac{\partial}{\partial x'} + \frac{\partial y'}{\partial x}\frac{\partial}{\partial y'}\right)\psi_R(\mathbf{x}') \\ &= \left(\cos\theta\frac{\partial}{\partial x'} + \sin\theta\frac{\partial}{\partial y'}\right)\psi_R(\mathbf{x}')\end{aligned} \tag{9}$$

  and

$$\begin{aligned}\partial_0\psi_R(\mathbf{x}) = \frac{\partial}{\partial t}\psi_R(\mathbf{x}) &= \left(\frac{\partial t''}{\partial t}\frac{\partial}{\partial t''} + \frac{\partial z'}{\partial t}\frac{\partial}{\partial z''}\right)\psi_R(\mathbf{x}'') \\ &= \left(\cosh\beta\frac{\partial}{\partial t''} + \sinh\beta\frac{\partial}{\partial z''}\right)\psi_R(\mathbf{x}'')\end{aligned} \tag{10}$$

- Now check Lorentz invariance of the z-rotation of our Dirac (Weyl) equation $i(\mathbf{I}_2\partial_0 + \boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\psi_R(\mathbf{x}) = 0$. Start with the expansion of eq. 2

$$\begin{pmatrix}(i\partial_0 + i\partial_3)\psi_R^1(x) + (i\partial_1 + \partial_2)\psi_R^2(x) \\ (i\partial_0 - i\partial_3)\psi_R^2(x) + (i\partial_1 - \partial_2)\psi_R^1(x)\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \tag{11}$$

Then apply the field transformations and chain rule to this equation.

$$\begin{pmatrix} e^{i\frac{\theta}{2}}(i\partial'_0 + i\partial'_3)\psi'^1_R(\mathbf{x}') + e^{-i\frac{\theta}{2}}e^{i\theta}(i\partial'_1 + \partial'_2)\psi'^2_R(\mathbf{x}') \\ e^{-i\frac{\theta}{2}}(i\partial'_0 - i\partial'_3)\psi'^1_R(\mathbf{x}') + e^{i\frac{\theta}{2}}e^{-i\theta}(i\partial'_1 - \partial'_2)\psi'^2_R(\mathbf{x}') \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{12}$$

This simplifies to

$$\begin{pmatrix} e^{i\frac{\theta}{2}}\left((i\partial'_0 + i\partial'_3)\psi'^1_R(\mathbf{x}') + (i\partial'_1 + \partial'_2)\psi'^2_R(\mathbf{x}')\right) \\ e^{-i\frac{\theta}{2}}\left((i\partial'_0 - i\partial'_3)\psi'^1_R(\mathbf{x}') + (i\partial'_1 - \partial'_2)\psi'^2_R(\mathbf{x}')\right) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{13}$$

and even simpler

$$\begin{pmatrix} e^{i\frac{\theta}{2}} & 0 \\ 0 & e^{-i\frac{\theta}{2}} \end{pmatrix} \begin{pmatrix} (i\partial'_0 + i\partial'_3)\psi'^1_R(\mathbf{x}') + (i\partial'_1 + \partial'_2)\psi'^2_R(\mathbf{x}') \\ (i\partial'_0 - i\partial'_3)\psi'^1_R(\mathbf{x}') + (i\partial'_1 - \partial'_2)\psi'^2_R(\mathbf{x}') \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{14}$$

We see that you can factor out the first matrix (since the RHS is $\mathbf{0}$) and get a form identical to eq. (11). So rotational invariance (around the z-axis)has been demonstrated.

- Can we introduce a mass term in the Weyl equation? Let's try $i(\mathbf{I}_2\partial_0 + \boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\psi_R(\mathbf{x}) = m\psi_R(\mathbf{x})$. If you now see how that mass term transforms, you get

$$\begin{pmatrix} e^{i\frac{\theta}{2}} & 0 \\ 0 & e^{-i\frac{\theta}{2}} \end{pmatrix} \begin{pmatrix} (i\partial'_0 + i\partial'_3)\psi'^1_R(\mathbf{x}') + (i\partial'_1 + \partial'_2)\psi'^2_R(\mathbf{x}') \\ (i\partial'_0 - i\partial'_3)\psi'^1_R(\mathbf{x}') + (i\partial'_1 - \partial'_2)\psi'^2_R(\mathbf{x}') \end{pmatrix} = \begin{pmatrix} e^{i\frac{\theta}{2}} & 0 \\ 0 & e^{-i\frac{\theta}{2}} \end{pmatrix} m \begin{pmatrix} \psi'^1_R(\mathbf{x}') \\ \psi'^2_R(\mathbf{x}') \end{pmatrix}.$$

You can divide by $\begin{pmatrix} e^{i\frac{\theta}{2}} & 0 \\ 0 & e^{-i\frac{\theta}{2}} \end{pmatrix}$ on both sides, again ending up with an equation of the same form that we started with. So again, we've shown rotational invariance. You might wonder why we didn't start with a mass term.

- Check Lorentz invariance of the z-boost of our massless Weyl equation. Following the same procedures as before, we obtain

$$\begin{pmatrix} e^{\frac{\beta}{2}} & 0 \\ 0 & e^{-\frac{\beta}{2}} \end{pmatrix} \begin{pmatrix} (i\partial''_0 + i\partial''_3)\psi''^1_R(\mathbf{x}'') + (i\partial''_1 + \partial''_2)\psi''^2_R(\mathbf{x}'') \\ (i\partial''_0 - i\partial''_3)\psi''^1_R(\mathbf{x}'') + (i\partial''_1 - \partial''_2)\psi''^2_R(\mathbf{x}'') \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{15}$$

Dividing by the first matrix, we end up with the Weyl equation again, so we've now shown that the equation is boost-invariant (in the z-direction.)

- What about the massive Weyl equation? The mass term on the right-hand side will transform to $\begin{pmatrix} e^{-\frac{\beta}{2}} & 0 \\ 0 & e^{\frac{\beta}{2}} \end{pmatrix} m \begin{pmatrix} \psi_R''^1(\mathbf{x}'') \\ \psi_R''^2(\mathbf{x}'') \end{pmatrix}$. Unlike the situation with rotations, the first matrix on the RHS doesn't cancel the first matrix on the LHS. **So with the mass term, the Weyl equation is NOT boost-invariant! A mass term breaks Lorentz invariance in the Weyl equation.**

4.1.2 ψ_L equation

- Start with $i(\mathbf{I}_2\partial_0 - \boldsymbol{\sigma} \cdot \boldsymbol{\nabla})\psi_L(\mathbf{x}) = 0$.
- Propose transformation laws:

  z-rotation by θ:

$$\psi_L(\mathbf{x}') = \begin{pmatrix} e^{i\frac{\theta}{2}} \psi_L'^1 \\ e^{-i\frac{\theta}{2}} \psi_L'^2 \end{pmatrix} (\mathbf{x}').$$

  where $\mathbf{x}' = (t', x', y, z') = (t, x\cos\theta - y\sin\theta, x\sin\theta + ycos\theta, z)$. This is the same transformation proposed for ψ_R.

  z-boost by β:

$$\psi_L(\mathbf{x}'') = \begin{pmatrix} e^{\frac{\beta}{2}} \psi_L''^1 \\ e^{-\frac{\beta}{2}} \psi_L''^2 \end{pmatrix} (\mathbf{x}'').$$

  where $\mathbf{x}'' = (t'', x'', y'', z'') = (t\cosh\beta + z\sinh\beta, x, y, t\sinh\beta + z\cosh\beta)$. This transformation differs from ψ_R by switching the sign of β.

- The analysis of rotational and boost invariance proceeds just as before. Since rotational transformations are the same for ψ_L as for ψ_R, the Weyl equation is again rotationally invariant. And just as before, the massive Weyl equation for ψ_L breaks Lorentz invariance.

- A trick! Notice that our boost transformation rule for ψ_L is $\psi_L(\mathbf{x}'') = \begin{pmatrix} e^{\frac{\beta}{2}} & 0 \\ 0 & e^{-\frac{\beta}{2}} \end{pmatrix} \begin{pmatrix} \psi_L''^1 \\ \psi_L''^2 \end{pmatrix}$. The matrix prefactor is exactly the same as the one on the left of the Weyl equation for ψ_R (see eq. (15)). So, if on the RHS of the right-Weyl equation (15) , we put $m \begin{pmatrix} \psi_L''^1 \\ \psi_L''^2 \end{pmatrix}$, then the prefactors will cancel out and the equation will be boost-invariant. In summary,

$$i(\mathbf{I}_2\partial_0 + \boldsymbol{\sigma} \cdot \boldsymbol{\nabla})\psi_R(\mathbf{x}) = m\psi_L(\mathbf{x}) \tag{16}$$

IS a Lorentz-invariant equation. There are some details you might want to verify like the rotational invariance.

Similarly,

$$i(\mathbf{I}_2\partial_0 - \boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\psi_L(\mathbf{x}) = m\psi_R(\mathbf{x}) \tag{17}$$

is **also** Lorentz-invariant.

4.2 4-spinors

We've done all the hard work! Way back in eq. (6), or if you prefer, Lancaster eq. (36.14), we showed that the Dirac equation was equivalent to the two equations examined above:

$$\begin{aligned} i(\mathbf{I}_2\partial_0 + \boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\psi_R(\mathbf{x}) &= m\psi_L(\mathbf{x}) \\ i(\mathbf{I}_2\partial_0 - \boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\psi_L(\mathbf{x}) &= m\psi_R(\mathbf{x}) \end{aligned} \tag{18}$$

But we showed that these were Lorentz-invariant. **Therefore the massive (and massless) Dirac equation is Lorentz-invariant!** What we've also seen, is that you can't make a Lorentz-invariant theory out of only one 2-spinor. That's why electrons (which are massive) are represented by 4-spinors.

4.3 Parity

So far, we've only discussed rotations and boosts. But most of our experience with natural laws, tells us we can't distinguish left from right. That's called *parity invariance.* The transformation that accomplishes this, must take (x,y,z) to $-$(x,y,z), while leaving the t-component alone. There's a simple group property. Do a parity-transform twice, and you end up where you started. So we describe the transformation under parity as, for example,

$$\psi'_L(t', x', y', z') = \mathbf{P}\psi_L(t, -x, -y, -z) \tag{19}$$

where $\mathbf{P}^2 = \mathbf{I}$. Now consider the massless Weyl equation $i(\mathbf{I}_2\partial_0 - \boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\psi_L(\mathbf{x}) = 0$. Under the parity transformation, this becomes $i(\mathbf{I}_2\partial_0 + \boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\mathbf{P}\psi_L(\mathbf{x}) = 0$. For this to be a symmetry, we'd need to find a matrix $\mathbf{P}$ so that $\mathbf{P}\sigma_i\mathbf{P} = -\sigma_i$ for each i. It can't be done. Therefore the massless Weyl equation violates parity.

This is important, because neutrinos were once thought to be massless and could therefore be described by a 2-spinor with the Weyl equation. This description would therefore result in a parity-violating theory. In fact, it turns out that the complete theory of neutrinos and their interactions **are** parity-violating.

5 The Dirac Lagrangian

For now, this section is brief. The Dirac action is taken to be

$$\mathcal{S} = \int d^4x \mathcal{L} = \int d^4x \psi^\dagger(x)\gamma_0 \left(i\partial\!\!\!/ - m\right) \psi(x) \tag{20}$$

where $\mathcal{L}$ is the Lagrangian. The Euler-Lagrange equations (equations of motion) include the equation

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi^\dagger)} - \frac{\partial \mathcal{L}}{\partial \psi^\dagger} = 0. \tag{21}$$

Since the field $\psi^\dagger$ doesn't appear in the Lagrangian with any derivatives, the first term of the equation is 0. The remaining equation becomes

$$-\gamma_0(i\partial\!\!\!/ - m)\psi(x) = 0. \tag{22}$$

We can factor out the γ_0, leaving the Dirac equation.

The point of this section, is that when we want to construct Lorentz-invariant equations of motion, we often do that by first constructing a Lorentz-invariant Lagrangian. A Lagrangian is just a real number, so for Lorentz-invariance, the Lagrangian is a scalar – i.e. it is unchanged by a Lorentz transformation. I haven't demonstrated this, but that's where we're heading.

6 More sections to come!

UNDER CONSTRUCTION