The Dirac equation and its solutions

Bill Celmaster

April 6, 2021

1 Introduction

These notes are an extension of the notes from February previously entitled *Introducing the Dirac Equation.* At the time, these notes followed the outline below up to the point "Deriving the Dirac equation from a Larangian".

We then detoured into Lorentz-group representation theory, with the objective of deriving the Dirac Lagrangian, which had led to the Dirac equation which we'd explored in these notes.

Now, in section (7) we'll return to the Dirac Lagrangian and solve its equations of motion – namely the Dirac equation. We'll also find out how what particles and antiparticles are described by the theory.

2 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).

3 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

4 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.

4.1 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).

[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.

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

  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)).
  - 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

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

4.2 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.

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

  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.

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

  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

$$\not{\partial}\equiv\gamma^\mu\partial_\mu.$$

  In that notation, the Dirac equation becomes

$$\boldsymbol{\left(i\not{\partial}-m\right)\psi(x)=0.}$$

- 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}.

5 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.**

5.1 2-spinors

5.1.1 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.

$$\begin{pmatrix}e^{i\frac{\theta}{2}}(i\partial_0' + i\partial_3')\psi_R'^1(\mathbf{x}') + e^{-i\frac{\theta}{2}}e^{i\theta}(i\partial_1' + \partial_2')\psi_R'^2(\mathbf{x}') \\ e^{-i\frac{\theta}{2}}(i\partial_0' - i\partial_3')\psi_R'^1(\mathbf{x}') + e^{i\frac{\theta}{2}}e^{-i\theta}(i\partial_1' - \partial_2')\psi_R'^2(\mathbf{x}')\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}. \tag{12}$$

This simplifies to

$$\begin{pmatrix}e^{i\frac{\theta}{2}}\left((i\partial_0' + i\partial_3')\psi_R'^1(\mathbf{x}') + (i\partial_1' + \partial_2')\psi_R'^2(\mathbf{x}')\right) \\ e^{-i\frac{\theta}{2}}\left((i\partial_0' - i\partial_3')\psi_R'^1(\mathbf{x}') + (i\partial_1' - \partial_2')\psi_R'^2(\mathbf{x}')\right)\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}. \tag{13}$$

and even simpler

$$\begin{pmatrix}e^{i\frac{\theta}{2}} & 0\\ 0 & e^{-i\frac{\theta}{2}}\end{pmatrix}\begin{pmatrix}(i\partial_0' + i\partial_3')\psi_R'^1(\mathbf{x}') + (i\partial_1' + \partial_2')\psi_R'^2(\mathbf{x}') \\ (i\partial_0' - i\partial_3')\psi_R'^1(\mathbf{x}') + (i\partial_1' - \partial_2')\psi_R'^2(\mathbf{x}')\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}. \tag{14}$$

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

$$\begin{pmatrix}e^{i\frac{\theta}{2}} & 0\\ 0 & e^{-i\frac{\theta}{2}}\end{pmatrix}\begin{pmatrix}(i\partial_0' + i\partial_3')\psi_R'^1(\mathbf{x}') + (i\partial_1' + \partial_2')\psi_R'^2(\mathbf{x}') \\ (i\partial_0' - i\partial_3')\psi_R'^1(\mathbf{x}') + (i\partial_1' - \partial_2')\psi_R'^2(\mathbf{x}')\end{pmatrix} = \begin{pmatrix}e^{i\frac{\theta}{2}} & 0\\ 0 & e^{-i\frac{\theta}{2}}\end{pmatrix} m \begin{pmatrix}\psi_R'^1(\mathbf{x}') \\ \psi_R'^2(\mathbf{x}')\end{pmatrix}.$$

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

$$\begin{pmatrix} e^{\frac{\beta}{2}} & 0 \\ 0 & e^{-\frac{\beta}{2}} \end{pmatrix} \begin{pmatrix} (i\partial_0'' + i\partial_3'')\psi_R''^1(\mathbf{x}'') + (i\partial_1'' + \partial_2'')\psi_R''^2(\mathbf{x}'') \\ (i\partial_0'' - i\partial_3'')\psi_R''^1(\mathbf{x}'') + (i\partial_1'' - \partial_2'')\psi_R''^2(\mathbf{x}'') \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (15)$$

  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.**

5.1.2 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.

5.2 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.

5.3 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.

6 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.

7 Quick Review

7.1 The Program of Quantum Field Theory

- We start with a Lagrangian which is Lorentz invariant.
- An example: The free theory of the electron is described by the Dirac Lagrangian.
- Another example: The interacting theory of the electron with electromagnetism is described by a Lagrangian of free electrons and photons **plus** an interactive term which we treat as a perturbation.
- We learned, when we studied path integrals, that we could compute all Green's functions – which are used to obtain the scattering matrix – by moments of a probability measure specified by the integral of the action (integral of Lagrangian).
- Those moments can be computed exactly for free Lagrangians such as the free electron and free photon Lagrangian
- The interactive term can be expanded as a Taylor series of moments with respect to a free-Lagrangian measure.
- This series can be systematically examined by using schematics, called Feynman diagrams, that represent integrals.

7.2 An intermediate step: Particles

- The above program is built on the theorem that the S-matrix can be computed from the Green's functions (the LSZ theorem).
- That theorem requires a connection between free fields and particles, which we come to next.
- First we solve the Dirac equation as though the fields in the Lagrangian were complex-valued functions.
- Then we promote those fields to operators. The only thing we need to know about the field operators are their commutation relations (remember, they aren't commutative).
- From that, we can establish the Hilbert space of states on which the operators act, and we find these states represent particles and antiparticles.

8 Solving the Dirac equation and anticommutation relations

Start with the Dirac equation in the form of eq. (7) and simplify notation by writing $\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}$. This is a set of 4 linear differential equations (one for each component) and these can be solved. For now, we assume that ψ is a complex-valued function.

There's no good reason for us to describe how one arrives at a solution. Instead, let's write down the answer and make sure we understand what it is. First, we write the answer in a general form.

$$\psi(x) = \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}^m}} \left(a_p^s u^s(p) e^{-ipx} + b_p^{s*} v^s(p) e^{ipx} \right), \tag{23}$$

where $\omega_{\mathbf{p}}^m = \sqrt{\mathbf{p}^2 + m^2} = p_0$. In the above expression, for each index value s, $u^s(p)$ and $v^s(p)$ have 4 complex components. a_p^s and b_p^s each have 1 complex component. In the exponents, the term px means $\omega_{\mathbf{p}}^m x_0 - \mathbf{p} \cdot \mathbf{x}$. The index s over which we sum, is called the "spin".

You might wonder why we separately factor out a_p^s and b_p^s since these could have been absorbed in $u^s(p)$ and $v^s(p)$. In fact, we have a lot of freedom in how to choose those vectors. Nothing up to this point helps us decide which basis set to use. However, a little later, we will invoke anti-commutation relations and if we choose the vectors $u^s(p)$ and $v^s(p)$ in just "the right way", we'll be able to interpret a_p^s and b_p^s as particle and antiparticle annihilation operators.

So, armed with this foresight, we'll display an appropriate set of solutions for 4-momenta that look like $p^\mu = (E, 0, p_y, 0)$.

$$u^1(p) = \frac{1}{2} \begin{pmatrix} \sqrt{\omega_{\mathbf{p}}^m + p_y} + \sqrt{\omega_{\mathbf{p}}^m - p_y} \\ i\left(\sqrt{\omega_{\mathbf{p}}^m - p_y} - \sqrt{\omega_{\mathbf{p}}^m + p_y} \right) \\ \sqrt{\omega_{\mathbf{p}}^m + p_y} + \sqrt{\omega_{\mathbf{p}}^m - p_y} \\ i\left(\sqrt{\omega_{\mathbf{p}}^m + p_y} - \sqrt{\omega_{\mathbf{p}}^m - p_y} \right) \end{pmatrix}, u^2(p) = \frac{1}{2} \begin{pmatrix} i\left(\sqrt{\omega_{\mathbf{p}}^m + p_y} - \sqrt{\omega_{\mathbf{p}}^m - p_y} \right) \\ \sqrt{\omega_{\mathbf{p}}^m + p_y} + \sqrt{\omega_{\mathbf{p}}^m - p_y} \\ i\left(\sqrt{\omega_{\mathbf{p}}^m - p_y} - \sqrt{\omega_{\mathbf{p}}^m + p_y} \right) \\ \sqrt{\omega_{\mathbf{p}}^m + p_y} + \sqrt{\omega_{\mathbf{p}}^m - p_y} \end{pmatrix} \tag{24}$$

$$v^1(p) = \frac{1}{2} \begin{pmatrix} \sqrt{\omega_{\mathbf{p}}^m + p_y} + \sqrt{\omega_{\mathbf{p}}^m - p_y} \\ i\left(\sqrt{\omega_{\mathbf{p}}^m - p_y} - \sqrt{\omega_{\mathbf{p}}^m + p_y} \right) \\ -\sqrt{\omega_{\mathbf{p}}^m + p_y} - \sqrt{\omega_{\mathbf{p}}^m - p_y} \\ -i\left(\sqrt{\omega_{\mathbf{p}}^m + p_y} - \sqrt{\omega_{\mathbf{p}}^m - p_y} \right) \end{pmatrix}, v^2(p) = \frac{1}{2} \begin{pmatrix} i\left(\sqrt{\omega_{\mathbf{p}}^m + p_y} - \sqrt{\omega_{\mathbf{p}}^m - p_y} \right) \\ \sqrt{\omega_{\mathbf{p}}^m + p_y} + \sqrt{\omega_{\mathbf{p}}^m - p_y} \\ -i\left(\sqrt{\omega_{\mathbf{p}}^m - p_y} - \sqrt{\omega_{\mathbf{p}}^m + p_y} \right) \\ -\sqrt{\omega_{\mathbf{p}}^m + p_y} - \sqrt{\omega_{\mathbf{p}}^m - p_y} \end{pmatrix}$$

In calculations of Green's functions, the exact forms of the solutions matter but for our purposes, we display the solutions just to illustrate that solutions

can be found. Feel free to plug these into the Dirac equation to see that they indeed are solutions.

There remain arbitrary coefficients, a_p^s and b_p^s. We see from eq. (23) that by performing some kind of inverse Fourier transform, and by plugging in the appropriate values of $u^s(p)$ (given above for momenta in the z-direction), we can express a_p^s and b_p^s as functions of ψ. We'll write these functions as

$$\begin{aligned} a_p^s &= \mathcal{F}_a(\psi, s, p) \\ b_p^s &= \mathcal{F}_b(\psi, s, p) \end{aligned} \tag{25}$$

So far, a_p^s, b_p^s and $\psi(x)$ have all been complex numbers. But for quantum mechanics we must promote them to be operators. Furthermore, the quantization conditions are obtained from the Lagrangian by the equal-time ($x_0 = y_0$ in the equations below) anti-commutation rules

$$\begin{aligned} \{\psi^\alpha(x), \Pi^\beta(y)\} &= i\delta(\mathbf{x}-\mathbf{y})\delta^{\alpha\beta} \\ \{\psi^\alpha(x), \psi^\beta(y)\} &= 0 \\ \{\Pi^\alpha(x), \Pi^\beta(y)\} &= 0. \end{aligned} \tag{26}$$

where the anticommutator brace symbol has the meaning $\{A, B\} = AB + BA$. The first of these equations becomes

$$\begin{aligned} \{\psi^\alpha(x), \frac{\partial \mathcal{L}}{\partial \dot{\psi}_\beta}(y)\} &= \{\psi^\alpha(x), i\psi^{\dagger\beta}(y)\} \\ &= i\delta(\mathbf{x}-\mathbf{y})\delta^{\alpha\beta} \\ \{\psi^\alpha(x), i\psi^\beta(y)\} &= 0 \\ \{\psi^{\dagger\alpha}(x), i\psi^{\dagger\beta}(y)\} &= 0 \end{aligned} \tag{27}$$

The anti-commutator instead of the commutator is used for spin-1/2 fields. That turns out to be a deep consequence of the axioms of quantum field theory – axioms which I haven't provided but include things like causality etc. When the anti-commutator is used, we say the fields are fermion fields, and the resulting particles (which we're about to get to) are fermions.

Recall from eq. (25) that a_p^s and b_p^s are expressed as functions of ψ. Their adjoints are expressed as functions of $\psi^\dagger$. We can therefore employ the anti-commutation relations of eq. (27) to obtain

$$\begin{aligned} \{a_p^s, a_{p'}^{\dagger s'}\} &= (2\pi)^3\delta^3(\mathbf{p}-\mathbf{p}')\delta^{ss'} \\ \{b_p^s, b_{p'}^{\dagger s'}\} &= (2\pi)^3\delta^3(\mathbf{p}-\mathbf{p}')\delta^{ss'} \\ \{a_p^s, a_{p'}^{s'}\} = \{a_p^{\dagger s}, a_{p'}^{\dagger s'}\} &= \{b_p^s, b_{p'}^{s'}\} = \{b_p^{\dagger s}, b_{p'}^{\dagger s'}\} = 0 \end{aligned} \tag{28}$$

8.1 The Hilbert space

Thusfar, we've solved the Dirac equation and expressed it in terms of operators a_p^s and b_p^s which satisfy the anticommutation relations of eq. (28). But we haven't said what space these operators act upon. That's what we'll do next – to construct the underlying Hilbert space.

8.1.1 The Energy operator

First we need one more ingredient, also obtained from the Lagrangian. That ingredient is the energy operator, E. One way to define that operator, is as the conserved quantity associated by Noether's theorem, with time-translation symmetry. In any case, various texts derive E as

$$E = \int d^3x \left(i\psi^\dagger(x)\partial_t\psi(x) \right). \tag{29}$$

Then, with some effort, we can expand the fields ψ and $\psi^\dagger$ using eq. (23) to obtain finally

$$E = \sum_s \int \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}}^m \left(a_p^{s\dagger} a_p^s - b_p^s b_p^{s\dagger} \right). \tag{30}$$

8.1.2 The Vacuum

OK. How do we construct a Hilbert space? To begin with, suppose (as is required by the axioms of field theory) that the energy operator E is bounded from below. Then there is an energy eigenstate in the Hilbert space, which has the lowest energy eigenvalue. Let's call this state $|\Omega\rangle$. We call this state the vacuum.[2] Sometimes we write $|0\rangle$ although this is misleading because it looks like 'nothing'. The vacuum is a state like any other state, so it isn't more or less of anything.

8.1.3 A simplified theory

Now I'm going to greatly simplify things by considering a different example that doesn't involve integrals over p. I'll change some notations so that we don't confuse this simplified theory with the real thing. So, in this simplified theory, there is only one annihilation operator $\tilde{a}$ with its corresponding

[2] Actually, nothing we've said so far implies that there is a unique vacuum. There could, in principle, be multiple states with the same lowest eigenvalue. So, to avoid this, we stipulate that an axiom of QFT is 'the vacuum is unique'. In point of fact, there are situations where that axiom is too restrictive, but for our purposes, let's go with it.

creation operator that satisfies the anticommutation relations

$$\begin{aligned} \{\tilde{a}, \tilde{a}^\dagger\} &= 1 \\ \{\tilde{a}, \tilde{a}\} &= 0 \\ \{\tilde{a}^\dagger, \tilde{a}^\dagger\} &= 0 \end{aligned} \tag{31}$$

In this same theory, the energy operator is $\tilde{E} = \tilde{\omega}\tilde{a}^\dagger\tilde{a}$. The lowest-energy eigenstate will be denoted $|\tilde{\Omega}\rangle$ and has energy eigenvalue $\mathcal{E}_0$, so

$$\tilde{E}|\tilde{\Omega}\rangle = \mathcal{E}_0|\tilde{\Omega}\rangle.$$

8.1.4 What the annihilation operator does to the vacuum

We'll now see that when we apply the annihilation operator $\tilde{a}$ to the vacuum state, it will result in 0. That's one reason we call it the *annihilation operator*. Define

$$|\tilde{\Omega}^-\rangle = \tilde{a}|\tilde{\Omega}\rangle.$$

Apply the energy operator

$$\begin{aligned} \tilde{E}|\tilde{\Omega}^-\rangle &= \tilde{\omega}\tilde{a}^\dagger\tilde{a}\tilde{a}|\tilde{\Omega}\rangle \\ &= \tilde{\omega}\left(-\tilde{a}\tilde{a}^\dagger\tilde{a} + \tilde{a}\right)|\tilde{\Omega}\rangle \\ &= \left(-\tilde{a}\tilde{E} + \tilde{a}\tilde{\omega}\right)|\tilde{\Omega}\rangle \\ &= (-\mathcal{E}_0 + \tilde{\omega})|\tilde{\Omega}^-\rangle. \end{aligned} \tag{32}$$

But since $\tilde{a}\tilde{a} = 0$ from the anticommutation relations, the first equality above shows that $\tilde{E}|\tilde{\Omega}^-\rangle = 0$. Therefore $(-\mathcal{E}_0 + \tilde{\omega})|\tilde{\Omega}^-\rangle = 0$ from which we conclude that either $|\tilde{\Omega}^-\rangle = 0$ or $(-\mathcal{E}_0 + \tilde{\omega}) = 0$. Which is it? Suppose it's the latter, so $\mathcal{E}_0 = \tilde{\omega}$. In other words, the lowest energy eigenvalue is $\tilde{\omega}$. But by assumption $|\tilde{\Omega}^-\rangle \neq 0$ and we've just shown that $\tilde{E}|\tilde{\Omega}^-\rangle = 0$. Therefore $|\tilde{\Omega}^-\rangle$ is an eigenvector of the energy, with an eigenvalue of 0, which is less than $\mathcal{E}_0$. That's a contradiction and therefore we've proven that $|\tilde{\Omega}^-\rangle = 0$ or equivalently, that $\tilde{a}|\tilde{\Omega}\rangle = 0$.

8.1.5 Particles – what the creation operator does to the vacuum

Still in this simplified theory, let's see what the creation operator does to the vacuum. First, notice that

$$\begin{aligned} \tilde{E}|\tilde{\Omega}\rangle &= \tilde{\omega}\tilde{a}^\dagger\left(\tilde{a}|\tilde{\Omega}\rangle\right) \\ &= 0 \end{aligned} \tag{33}$$

which shows that $\mathcal{E}_0 = 0$. Next, define the state $|1\rangle = \tilde{a}^\dagger|\tilde{\Omega}\rangle$. We'll see that this is an eigenstate of the energy operator, with the eigenvalue $\tilde{\omega}$.

$$\begin{aligned}\tilde{E}|1\rangle &= \tilde{\omega}\tilde{a}^\dagger\tilde{a}\tilde{a}^\dagger|\tilde{\Omega}\rangle \\ &= \tilde{\omega}\left(-\tilde{a}^\dagger\tilde{a}^\dagger\tilde{a} + \tilde{a}^\dagger\right)|\tilde{\Omega}\rangle \\ &= \tilde{\omega}\tilde{a}^\dagger|\tilde{\Omega}\rangle \\ &= \tilde{\omega}|1\rangle.\end{aligned} \tag{34}$$

We've made use of the fact that $\tilde{a}^\dagger\tilde{a}^\dagger = 0$.

The interpretation (in this toy model) of the state $|1\rangle$ is that it is a *particle* of energy $\tilde{\omega}$. We obtained it by applying the creation operator to the vacuum state. Hence the name *creation operator*.

8.1.6 Particles are field vibrations

You'll sometimes hear that particles are just vibrations of a quantum field. I've never been especially comfortable with that statement, although I suppose it's a good picture to carry in one's head. My problem has to do with the fact that a quantum field is an operator. I know what it means for a real-valued function to vibrate. If the value of the function goes up and down as a function of time or space, we say it's vibrating. Or if we look at a complex function, we can separately look at oscillations of the real and imaginary parts. But it's less clear to me what should be meant by the vibration or oscillation of an operator.

Be that as it may, consider the field decomposition from before,

$$\psi^*(x) = \sum_{s=1}^{2}\int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\mathbf{p}}^m}}\left(a_p^{s\dagger}u^{*s}(p)e^{ipx} + b_p^s v^{*s}(p)e^{-ipx}\right). \tag{35}$$

We now recognize that the operator-coefficients $a_p^{s\dagger}$ are particle-creation operators. For each momentum, those particle-creation operators are multiplied by a vibratory term e^{ipx} with wave-style vibrations in time and space. We can describe this as 'a particle represented by a vibration of the field'. In some sense, you can think of the field Fourier component (the 'vibration') as having *potentiality* (don't confuse this with potential energy) to create a particle with that value of the momentum. If you like the terminology, then what you see is that the field has the potential to create particles of all possible momenta **but** all having the property that $p_0^2 - \mathbf{p}\cdot\mathbf{p} = m^2$ (the dispersion relation).

Notice that the minimum energy of any of these particles, is the mass m.

8.1.7 Fermi statistics

The notation $|1\rangle$ denotes a state with one particle. What happens when we apply the creation operator to this state? Does it become a state with two particles? We have

$$\begin{aligned} \tilde{a}^\dagger|1\rangle &= \tilde{a}^\dagger\tilde{a}^\dagger|\tilde{\Omega}\rangle \\ &= 0 \end{aligned} \tag{36}$$

because $\tilde{a}^\dagger\tilde{a}^\dagger = 0$ from the anticommutation relations. So the creation operator does **not** turn a 1-particle state into a 2-particle state. In fact, it annihilates the 1-particle state.

This doesn't necessarily prove that we can't create states that would be easily interpreted as 2-particle states. All we've shown is that we can't do this using creation operators. However, it turns out to be true that we can't do it any other way either.

Now let's turn to the full theory we started with where we integrated over momenta and summed over spins. We can apply the anti-commutation operators in exactly the same kind of way we did in the simplified model, but separately for each momentum and spin. The same argument as before tells us that the vacuum is annihilated by all of the annihilation operators. We could define states

$$|p, s\rangle = \frac{1}{\sqrt{2\omega_p^m}} a_p^{s\dagger}|\Omega\rangle$$

to describe a single particle state of momentum p and spin s and by applying the energy operator, we would see that it has energy $\omega_{\mathbf{p}}^m$. You should be slightly bothered by the fact that we had an integral over momenta, so that there is a continuum of states. But this kind of thing should be familiar from plain old quantum mechanics, where we cavalierly sum over a continuum of momentum states (for example).

But this still doesn't give us our entire Hilbert space. what about a state corresponding to the operation (on the vacuum) of a creation operator with spin s and momentum p, followed by a creation operator with spin s' and momentum p'. Unlike the simplified case, this does not result in 0 because the two creation operators are different. The resulting state describes one particle with momentum p and spin s, and another particle with momentum p' and spin s'. The most general state is a state with an *occupation number* of either 0 or 1 for each of the possible momenta and spins.

We say that such a Hilbert space describes fermions. Here you should recognize the Pauli exclusion principle that says you can't have more than one particle in a single quantum state.

8.1.8 Antiparticles

I've exclusively been talking about the a_p^s operators. What about the b_p^s's? Those also appear in the energy operator. Recall from before

$$E = \sum_s \int \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}}^m \left(a_p^{s\dagger} a_p^s - b_p^s b_p^{s\dagger} \right).$$

If we rewrite the terms involving the b operators as

$$-b_p^s b_p^{s\dagger} = b_p^{s\dagger} b_p^s - (2\pi)^3$$

and temporarily ignore the constant $(-(2\pi)^3)$, we see that the resulting expression is

$$E = \sum_s \int \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}}^m \left(a_p^{s\dagger} a_p^s + b_p^{s\dagger} b_p^s \right) - ....$$

The $b^\dagger$ operators behave in just the same way as the $a^\dagger$ operators, but create a completely different set of states which we call *the antiparticle states*. So now our Hilbert space has been extended so that the states include both particles and antiparticles. For each value of momentum and spin, there is either one or no particles that have that value and similarly there is either one or no antiparticles that have that value. The designations 'particle' and 'antiparticle', without interaction terms, don't help us distinguish the two kinds of particles. However, it's worth noting that in the expansion of the field operators, particle annihilation operators are the coefficients of e^{-ipx} and antiparticle creation operators are the coefficients of e^{ipx}. By convention, and primarily for historical reasons, the first exponential is said to describe positive frequencies, and the second exponential is said to describe negative frequencies. In the early days of quantum mechanics, those two situations were associated with particles of positive energy and negative energy respectively. However, that interpretation has long ago been replaced by the particle/antiparticle interpretation.

8.1.9 The zero-point energy

In the above discussion of antiparticles, there was a constant term generated from anticommuting the products of b operators. Such a term is present for each momentum and spin, and since those appear in an integral over momenta, the total is infinite. That total constant is known as *the zero-point energy* and has been the subject of much hair-pulling.

The modern understanding of this constant, is that it plays no role in physics – other than potentially physics that includes gravity (more in a

moment). A similar kind of thing occurs in classical physics. Energy is defined only up to a constant. That statement is common, but isn't especially meaningful unless we say how energy is defined. One definition is 'energy is a conserved quantity'. There are lots of conserved quantities, so that definition isn't all that illuminating, but it suffices to help us understand that an overall constant would play no role since it appears both before and after an interaction and thus makes no difference to the conservation of energy. Another definition has to do with the relation between energy and time-translation invariance (via Noether's theorem). Again, an overall constant energy difference has no significance in generating time translations.

The one exception to all of this is relativistic gravity (general relativity). The source of gravity is the energy-momentum tensor, so an overall constant would be expected to change the value of the source and thus change the amount of gravitational attraction. Frankly, I haven't examined the consequences of zero-point energy (or other constants) to the theory of gravity. I know it's a subject of interest but haven't properly examined how modern physicists deal with this.

For our purposes, where gravity isn't included in the theory, the constant has no effect so we can simply get rid of it. Or, more elegantly, we can subtract it by adding a term to the Lagrangian. One way this is done systematically is to assert that the terms in the Lagrangian are normal-ordered – that is, when expanded in annihilation and creation operators, all products of terms are re-ordered so that the annihilation operators appear on the left and the creation operators appear on the right.