Peeter Joot
peeterjoot@pm.me

# PHY2403H Quantum Field Theory. Lecture 20: Dirac Lagrangian, spinor solutions to the KG equation, Dirac matrices, plane wave solution, helicity. Taught by Prof. Erich Poppitz

*DISCLAIMER: Very rough notes from class, with some additional side notes.* These are notes for the UofT course PHY2403H, Quantum Field Theory, taught by Prof. Erich Poppitz, fall 2018.

## 1.1 Review.

Last time we

- introduced the Clifford algebra Dirac matrix (gamma matrices) elements satisfying

$$\left\{ \gamma^\mu, \gamma^\nu \right\} = 2 g^{\mu\nu}, \tag{1.1}$$

  where we use the Weyl representation

$$\begin{aligned} \gamma^0 &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ \gamma^k &= \begin{bmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{bmatrix}. \end{aligned} \tag{1.2}$$

  In particular $(\gamma^0)^2 = 1$, $(\gamma^0)^\dagger = \gamma^0$.

- and left off after showing that the Dirac Lagrangian

$$\mathcal{L} = \overline{\Psi} \left( i \gamma^\mu \partial_\mu - m \right) \Psi, \tag{1.3}$$

  where $\overline{\Psi} = \Psi^\dagger \gamma^0$, is Lorentz invariant. We argued that a single spinor can only describe a massless field, and that a two spinor construction can be used for a massive field. We skipped from there to the Dirac Lagrangian above.

## 1.2 Dirac equation.

Varying the Dirac action (exercise 1.1), we find the Dirac equation

$$\boxed{\left(i\gamma^\mu \partial_\mu - m\right)\Psi = 0.} \tag{1.4}$$

**Theorem 1.1: Dirac equations solutions satisfy the Klein-Gordon equation**

If $\Psi$ obeys eq. (1.4), the Dirac equation, then $\Psi$ obeys the Klein-Gordon equation.

Theorem 1.1 follows by pre-multiplying by a sort of "conjugate" operator[1] $i\gamma^\mu + m$ to find

$$\begin{aligned} 0 &= \left(i\gamma^\mu \partial_\mu + m\right)\left(i\gamma^\mu \partial_\mu - m\right)\Psi \\ &= \left(-\gamma^\mu \gamma_\nu \partial_\mu \partial_\nu - m^2\right)\Psi \\ &= \left(-\frac{1}{2}\left(\gamma^\mu \gamma_\nu + \gamma^\nu \gamma_\mu\right)\partial_\mu \partial_\nu - m^2\right)\Psi \\ &= \left(-g^{\mu\nu}\partial_\mu \partial_\nu - m^2\right)\Psi \\ &= -\left(\partial_\mu \partial^\mu + m^2\right)\Psi \\ &= -\left(\partial_{00} - \boldsymbol{\nabla}^2 + m^2\right)\Psi, \end{aligned} \tag{1.5}$$

which is a Klein-Gordon equation for $\Psi$.

*Goal:* Expand $\Psi(\mathbf{x}, t)$ in a basis of solutions of the Dirac equation. Call the coefficients $a, b, \cdots$. This will be like

$$\phi \sim \int \frac{d^3 p}{(2\pi)^3 2\omega_{\mathbf{p}}}\left(e^{ip\cdot x} a_{\mathbf{p}}^\dagger + e^{-ip\cdot x} a_{\mathbf{p}}\right) \tag{1.6}$$

As with the scalar field, let's look for plane wave solutions. We'll first look for solutions of the form

$$\Psi(x) = u(p) e^{-ip\cdot x}, \tag{1.7}$$

where $p^2 = m^2$, $p^0 > 0 \,\forall p$. Plugging into eq. (1.4) we find

$$\left(\gamma^\mu p_\mu - m\right) u(p) = 0. \tag{1.8}$$

[1] Q: Is there a name for such a conjugation operation?

Let's write this out explicitly for exposition, first noting that

$$\begin{aligned}\gamma^\mu p_\mu &= \gamma^0 p_0 + \gamma^k p_k \\ &= p_0 \begin{bmatrix} 0 & \sigma^0 \\ \sigma^0 & 0 \end{bmatrix} + p_k \begin{bmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & p^0\sigma^0 - \boldsymbol{\sigma}\cdot\mathbf{p} \\ p^0\sigma^0 + \boldsymbol{\sigma}\cdot\mathbf{p} & 0 \end{bmatrix},\end{aligned} \tag{1.9}$$

so eq. (1.8) becomes

$$\begin{aligned}0 &= \begin{bmatrix} -m & p^0\sigma^0 - \boldsymbol{\sigma}\cdot\mathbf{p} \\ p^0\sigma^0 + \boldsymbol{\sigma}\cdot\mathbf{p} & -m \end{bmatrix} \begin{bmatrix} u_1(p) \\ u_2(p) \end{bmatrix} \\ &= \begin{bmatrix} -m & p\cdot\sigma \\ p\cdot\bar{\sigma} & -m \end{bmatrix} \begin{bmatrix} u_1(p) \\ u_2(p) \end{bmatrix},\end{aligned} \tag{1.10}$$

where the following handy shorthand[2] has been used to group the momentum related block matrices

$$\begin{aligned}p\cdot\sigma &= p^0\sigma^0 - \mathbf{p}\cdot\boldsymbol{\sigma} \\ p\cdot\bar{\sigma} &= p^0\sigma^0 + \mathbf{p}\cdot\boldsymbol{\sigma}.\end{aligned} \tag{1.11}$$

Note that these $p\cdot\sigma, p\cdot\bar{\sigma}$'s are both block matrices. In particular

$$\begin{aligned}p\cdot\sigma &= p_\mu\sigma^\mu \\ &= \begin{bmatrix} p_0 + p_3 & p_1 - ip_2 \\ p_1 + ip_2 & p_0 - p_3 \end{bmatrix}.\end{aligned} \tag{1.12}$$

The question is what $u's$ obey such an equation.

We can gain some insight by first considering the rest frame, where $\mathbf{p} = 0, p^0 = m$. Going back to eq. (1.8), the rest frame Dirac equation becomes

$$\begin{aligned}0 &= \left(\gamma^0 p_0 - m\right) u \\ &= m\left(\gamma^0 - 1\right) u.\end{aligned} \tag{1.13}$$

Our block matrix equation is now reduced to a set of $2\times 2$ identity matrices

$$0 = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} u(\mathbf{p} = 0). \tag{1.14}$$

[2] Assuming I wrote this down correctly, this follows the usual convention $x\cdot p = x^\mu p_\mu = x^0 p^0 - \mathbf{x}\cdot\mathbf{p}$. I had some doubt that I got the signs right in my notes from class, since a peek at [1] seemingly showed the opposite sign convention where $\sigma\cdot x$ was first defined, namely eq. 3.41/3.43. There they write $\sigma\cdot\partial = \partial_0 + \boldsymbol{\sigma}\cdot\boldsymbol{\nabla}$, not $\sigma\cdot\partial = \partial_0 - \boldsymbol{\sigma}\cdot\boldsymbol{\nabla}$. What explains this is the fact that the four-gradient in coordinate form should really considered a lower index quantity ($\partial_\mu$), so in the scalar+vector tuple form, we should write $\partial^\mu = (\partial^0, -\boldsymbol{\nabla})$, or $\partial_\mu = (\partial_0, \boldsymbol{\nabla})$. This means that $\sigma\cdot\partial = \sigma^0\partial^0 - \boldsymbol{\sigma}\cdot(-\boldsymbol{\nabla}) = \partial_0 + \boldsymbol{\sigma}\cdot\boldsymbol{\nabla}$. Having a tuple notation that can be used to represent either lower or upper index quantities is very confusing, and probably justifies avoiding that notation for any lower index quantity whenever possible for clarity!

The solution space is given by

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \zeta \\ \zeta \end{bmatrix} = 0, \tag{1.15}$$

where $\zeta$ is itself a $2 \times 1$ column matrix, say

$$\zeta = \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix}. \tag{1.16}$$

so our solutions are all proportional to column

$$u(\mathbf{p} = 0) \sim \sqrt{m} \begin{bmatrix} \zeta \\ \zeta \end{bmatrix}. \tag{1.17}$$

We'll figure out the desired normalization later[3], and have added a $\sqrt{m}$ factor into the mix for later convenience. Equation (1.17) is a solution of the Dirac equation in the rest frame where $\mathbf{p} = 0$. A solution in a frame where $\mathbf{p} \neq 0$ can be found using a boost. We won't work that out explicitly here, but instead show the answer and argue that it must be valid, but the interested student can find that boost calculated explicitly in [1].

*Claim:* in a boosted frame where $\mathbf{p} \neq 0$ solution is

$$u(p) = \begin{bmatrix} \sqrt{p \cdot \sigma} \zeta \\ \sqrt{p \cdot \bar{\sigma}} \zeta \end{bmatrix} \tag{1.18}$$

What do we mean by these square roots? Since $p \cdot \sigma, p \cdot \bar{\sigma}$ are both Hermitian $2 \times 2$ matrices, we can define the square root as the matrix of the square roots of the eigenvalues.

*Check:* In the rest frame

$$\begin{aligned} \sqrt{p \cdot \sigma}\Big|_{\mathbf{p}=0} &= \sqrt{p \cdot \bar{\sigma}}\Big|_{\mathbf{p}=0} \\ &= \begin{bmatrix} \sqrt{p_0} & 0 \\ 0 & \sqrt{p_0} \end{bmatrix} \\ &= \begin{bmatrix} \sqrt{m} & 0 \\ 0 & \sqrt{m} \end{bmatrix}, \end{aligned} \tag{1.19}$$

so

$$\begin{aligned} u(p) &= \sqrt{m} \begin{bmatrix} \sigma^0 \zeta \\ \sigma^0 \zeta \end{bmatrix} \\ &= \sqrt{m} \begin{bmatrix} \zeta \\ \zeta \end{bmatrix}, \end{aligned} \tag{1.20}$$

[3] [1] says of this that we pick a normalization with $\zeta^\dagger \zeta = 1$.

as we already found.

We claim that the structure of the boost is

$$u(\mathbf{p}) = \begin{bmatrix} \frac{\sqrt{p\cdot\sigma}}{m} & 0 \\ 0 & \frac{\sqrt{p\cdot\bar{\sigma}}}{m} \end{bmatrix} \underbrace{\sqrt{m}\begin{bmatrix} \zeta \\ \zeta \end{bmatrix}}_{u(\mathbf{p}=0)} \tag{1.21}$$

We'd like to check that this is an element of $SL(2,\mathbb{C})$. We'll also see in the end that we don't have to calculate these square roots, since we always end up with two spinors and when all is said we end up with products of these roots.

**Lemma 1.1: Determinant of square root.**

If matrix $A$ is diagonalizable, then $\det\sqrt{A} = \sqrt{\det A}$.

Proof: Suppose that

$$A = U\,\mathrm{diag}(\lambda_1, \cdots \lambda_n)U^\dagger, \tag{1.22}$$

then

$$\begin{aligned} \det\sqrt{A} &= \det\left(U\,\mathrm{diag}(\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_n})U^\dagger\right) \\ &= \prod_j \sqrt{\lambda_j} \\ &= \sqrt{\prod_j \lambda_j} \\ &= \sqrt{\det A}. \end{aligned} \tag{1.23}$$

**Lemma 1.2: Determinant of $p\cdot\sigma$.**

$$\det\frac{\sqrt{p\cdot\sigma}}{\sqrt{m}} = 1$$

Proof:

$$
\begin{aligned}
\det \frac{\sqrt{p\cdot\sigma}}{\sqrt{m}} &= \sqrt{\det \frac{(p\cdot\sigma)}{m}}, \\
&= \sqrt{\det \frac{1}{m} \begin{bmatrix} p^0 + p^3 & -p_1 + ip_2 \\ -p_1 - ip_2 & p^0 + p^3 \end{bmatrix}} \\
&= \sqrt{\frac{1}{m^2}\left((p^0)^2 - \mathbf{p}^2\right)} \\
&= \sqrt{\frac{m^2}{m^2}} \\
&= 1.
\end{aligned} \tag{1.24}
$$

**Lemma 1.3:** $(p\cdot\sigma)(p\cdot\bar{\sigma})$

$$(p\cdot\sigma)(p\cdot\bar{\sigma}) = m^2.$$

Proof:

$$
\begin{aligned}
(p\cdot\sigma)(p\cdot\bar{\sigma}) &= \left(p^0 - \mathbf{p}\cdot\boldsymbol{\sigma}\right)\left(p^0 + \mathbf{p}\cdot\boldsymbol{\sigma}\right) \\
&= (p^0)^2 - (\mathbf{p}\cdot\boldsymbol{\sigma})^2 \\
&= (p^0)^2 - \mathbf{p}^2 \\
&= m^2.
\end{aligned} \tag{1.25}
$$

**Theorem 1.2:** $u(p)$ **is a solution to the Dirac equation.**

Equation (1.18) is a solution of eq. (1.4), the Dirac equation.

Proof:

$$
\begin{aligned}
\left(\gamma^\mu p_\mu - m\right) u(p) &= \begin{bmatrix} -m & p\cdot\sigma \\ p\cdot\bar{\sigma} & -m \end{bmatrix} \begin{bmatrix} \sqrt{p\cdot\sigma}\zeta \\ \sqrt{p\cdot\bar{\sigma}}\zeta \end{bmatrix} \\
&= \begin{bmatrix} \left(-m\sqrt{p\cdot\sigma} + p\cdot\sigma\sqrt{p\cdot\bar{\sigma}}\right)\zeta \\ \left(p\cdot\bar{\sigma}\sqrt{p\cdot\sigma} - m\sqrt{p\cdot\bar{\sigma}}\right)\zeta \end{bmatrix} \\
&= \begin{bmatrix} \sqrt{p\cdot\sigma}\left(-m + \sqrt{p\cdot\sigma}\sqrt{p\cdot\bar{\sigma}}\right)\zeta \\ \sqrt{p\cdot\bar{\sigma}}\left(p\sqrt{\cdot\bar{\sigma}}\sqrt{p\cdot\sigma} - m\right)\zeta \end{bmatrix} \\
&= \begin{bmatrix} \sqrt{p\cdot\sigma}\left(-m + \sqrt{m^2}\right)\zeta \\ \sqrt{p\cdot\bar{\sigma}}\left(p\sqrt{m^2} - m\right)\zeta \end{bmatrix} \\
&= 0.
\end{aligned} \tag{1.26}
$$

*Summary:* For $p^0 > 0, p^2 = m^2$

$$\Psi(x) = e^{-ip\cdot x}u(p) \tag{1.27}$$

$$u(p) = \begin{bmatrix} \sqrt{p\cdot\sigma}\zeta \\ \sqrt{p\cdot\bar{\sigma}}\zeta \end{bmatrix} \tag{1.28}$$

*Example:*

$$p = \left(E, 0, 0, p^3\right) \tag{1.29}$$

We have

$$\begin{aligned} \sqrt{\sigma\cdot p} &= \sqrt{E - p^3\sigma^3} \\ &= \sqrt{\left(\begin{bmatrix} E - p^3 & 0 \\ 0 & E + p^3 \end{bmatrix}\right)} \\ &= \begin{bmatrix} \sqrt{E - p^3} & 0 \\ 0 & \sqrt{E + p^3} \end{bmatrix}. \end{aligned} \tag{1.30}$$

Similarly

$$\sqrt{\bar{\sigma}\cdot p} = \begin{bmatrix} \sqrt{E + p^3} & 0 \\ 0 & \sqrt{E - p^3} \end{bmatrix}, \tag{1.31}$$

so

$$u(p) = \begin{bmatrix} \begin{bmatrix} \sqrt{E - p^3} & 0 \\ 0 & \sqrt{E + p^3} \end{bmatrix}\zeta \\ \begin{bmatrix} \sqrt{E + p^3} & 0 \\ 0 & \sqrt{E - p^3} \end{bmatrix}\zeta \end{bmatrix} \tag{1.32}$$

Suppose we let

$$\zeta = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{1.33}$$

we are left with

$$u(p) = \begin{bmatrix} \sqrt{E - p^3}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ \sqrt{E + p^3}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{bmatrix} \tag{1.34}$$

Alternatively for $\zeta = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$$u(p) = \begin{bmatrix} \sqrt{E + p^3}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ \sqrt{E - p^3}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{bmatrix} \tag{1.35}$$

If we pick $p_3 = E$[4], then we find two solutions

$$u(p)\big|_{\zeta=(1,0)^{\mathrm{T}},p_3=E} = \begin{bmatrix} 0\begin{bmatrix}1\\0\end{bmatrix} \\ \sqrt{2E}\begin{bmatrix}1\\0\end{bmatrix}\end{bmatrix} \tag{1.36}$$

and

$$u(p)\big|_{\zeta=(0,1)^{\mathrm{T}},p_3=E} = \begin{bmatrix} \sqrt{2E}\begin{bmatrix}0\\1\end{bmatrix} \\ 0\begin{bmatrix}0\\1\end{bmatrix}\end{bmatrix} \tag{1.37}$$

## 1.3 Helicity

.

Let $h$ (the helicity) be

$$\begin{aligned} h &= \frac{1}{2}\begin{bmatrix}\hat{\mathbf{p}}\cdot\sigma & 0 \\ 0 & \hat{\mathbf{p}}\cdot\sigma\end{bmatrix} \\ &= \hat{\mathbf{p}}\cdot\mathbf{S}, \end{aligned} \tag{1.38}$$

where

$$\mathbf{S} = \begin{bmatrix}\frac{\sigma}{2} & 0 \\ 0 & \frac{\sigma}{2}\end{bmatrix}. \tag{1.39}$$

$h$ has eigenvalues $\pm 1/2$.

It turns out that eq. (1.36), and eq. (1.37) are both eigenstates of the helicity operator.

$$hu^{(1)} = \frac{1}{2}u^{(1)} \tag{1.40}$$

$$hu^{(2)} = -\frac{1}{2}u^{(2)}, \tag{1.41}$$

corresponding to momentum aligned with and opposing the spin directions as sketched in fig. 1.1.

## 1.4 Next time.

We found $\Psi = ue^{-ip\cdot x}$. Next time we will seek another solution $\Psi = ve^{+ip\cdot x}$, and we will also figure out how to normalize things.

---

[4] we can boost a massive particle to be arbitrarily close to $p_3 = E$.

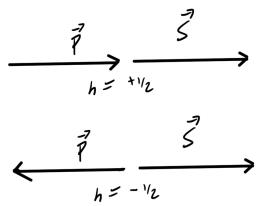


**Figure 1.1:** Helicity orientation.

## 1.5 Problems:

Exercise 1.1 Vary the Dirac action.

Answer for Exercise 1.1

From the action

$$S = \int d^4x \overline{\Psi} \left( i\gamma^\mu \partial_\mu - m \right) \Psi, \tag{1.42}$$

we find

$$\delta S = \int d^4x \delta\overline{\Psi} \left( i\gamma^\mu \partial_\mu - m \right) \Psi + \int d^4x \overline{\Psi} \left( i\gamma^\mu \partial_\mu - m \right) \delta\Psi. \tag{1.43}$$

There are two ways to deal with this. One (somewhat unsatisfactory seeming to me) is to treat both $\delta\overline{\Psi}$ and $\delta\Psi$ as independent variations, requiring that $\delta S = 0$ for any such variations. In that case we find that $\left( i\gamma^\mu \partial_\mu - m \right) \Psi = 0$ if the total variation of the action is zero. That leaves the somewhat awkward question of what to do with the $0 = \int d^4x \overline{\Psi} \left( i\gamma^\mu \partial_\mu - m \right) \delta\Psi$ constraint. However, that question can be resolved by observing that these two contributions to the variation are not independent. In particular

$$
\begin{aligned}
\left( \int d^4x \overline{\Psi} \left( i\gamma^\mu \partial_\mu - m \right) \delta\Psi \right)^\dagger &= \int d^4x \delta\Psi^\dagger \left( -i(\gamma^\mu)^\dagger \partial_\mu - m \right) \gamma^0 \Psi \\
&= \int d^4x \delta\Psi^\dagger \gamma^0 \left( -i\gamma^0 (\gamma^\mu)^\dagger \gamma^0 \overleftarrow{\partial}_\mu - m \right) \Psi \\
&= \int d^4x \delta\overline{\Psi} \left( -i\gamma^\mu \overleftarrow{\partial}_\mu - m \right) \Psi \\
&= \int d^4x \left( \partial_\mu \left( -i\delta\overline{\Psi}\gamma^\mu\Psi \right) - \left( -i\delta\overline{\Psi}\gamma^\mu\partial_\mu\Psi \right) - m\delta\overline{\Psi}\Psi \right) \\
&= \int d^4x \delta\overline{\Psi} \left( i\gamma^\mu \partial_\mu - m \right) \Psi,
\end{aligned}
\tag{1.44}
$$

where the boundary integral has been assumed to be zero. This shows that the total variation is

$$
\delta S = \int d^4x \left( \delta\overline{\Psi} D\Psi + \left( \delta\overline{\Psi} D\Psi \right)^\dagger \right), \tag{1.45}
$$

where

$$
D = i\gamma^\mu \partial_\mu - m, \tag{1.46}
$$

represents the Dirac operator. Requiring that the action variation $\delta S = 0$ is zero for all $\delta\overline{\Psi}$, means that $D\Psi = 0$, which proves eq. (1.4).

Given that the action itself is real, it makes sense for it's variation to be real, as demonstrated above. A nice side effect of demonstrating this is the removal of the redundant variation variable.

# Bibliography

[1] Michael E Peskin and Daniel V Schroeder. *An introduction to Quantum Field Theory*. Westview, 1995. 2, 1.2, 3