Quantum Field Theory: Example Sheet 3

Michaelmas 2018, Mathematical Tripos Part III, University of Cambridge

Model solutions by L. K. Wong*. Last updated November 17, 2018.

Question 1

Proposition 1.1: *The Clifford algebra*

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}\mathbb{1}_4, \tag{1.1}$$

with $\mathbb{1}_n$ denoting the $n \times n$ identity matrix,[1] *is satisfied by Dirac matrices in the chiral representation,*

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1}_2 \\ \mathbb{1}_2 & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}. \tag{1.2}$$

Proof.—This follows from direct evaluation, while using the identity

$$\sigma^i\sigma^j = \delta^{ij}\mathbb{1}_2 + i\epsilon^{ijk}\sigma^k \tag{1.3}$$

for the Pauli matrices. Note that the index k is being implicitly summed over. ■

Proposition 1.2: *Different representations of the Clifford algebra can be obtained through a unitary transformation $\gamma^\mu \to (\gamma')^\mu = U\gamma^\mu U^\dagger$. The unitary matrix*

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix} \mathbb{1}_2 & \mathbb{1}_2 \\ -\mathbb{1}_2 & \mathbb{1}_2 \end{pmatrix} \tag{1.4}$$

results in the Dirac representation

$$(\gamma')^0 = \begin{pmatrix} \mathbb{1}_2 & 0 \\ 0 & -\mathbb{1}_2 \end{pmatrix}, \quad (\gamma')^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}. \tag{1.5}$$

Proof.—Given the representations in Eqs. (1.2) and (1.5), we wish to determine the corresponding unitary transformation U. As both γ^μ and $(\gamma')^\mu$ can be written as blocks of 2×2 matrices, it is natural to expect that the same is true for U, namely let

$$U = \begin{pmatrix} A & B \\ C & D \end{pmatrix}. \tag{1.6}$$

Unitarity ($U^\dagger U = \mathbb{1}_4$) enforces three independent constraint equations on these 2×2 complex matrices,

$$A^\dagger A + C^\dagger C = \mathbb{1}_2, \quad A^\dagger B + C^\dagger D = 0, \quad B^\dagger B + D^\dagger D = \mathbb{1}_2. \tag{1.7}$$

We then use the equation $(\gamma')^0 U = U\gamma^0$ to learn that $A = B$ and $C = -D$. Plugging these back into the equations above, we get

$$A^\dagger A + D^\dagger D = \mathbb{1}_2, \quad A^\dagger A = D^\dagger D; \tag{1.8}$$

telling us that $A/\sqrt{2}$ and $D/\sqrt{2}$ are themselves unitary matrices. Finally, we use $(\gamma')^i U = U\gamma^i$ to find

$$\sigma^i A = D\sigma^i, \quad \sigma^i D = A\sigma^i. \tag{1.9}$$

*Please send any comments or corrections to L.K.Wong@damtp.cam.ac.uk
[1] When the context is clear, we will often omit writing $\mathbb{1}_n$ explicitly.

It is easy to correctly guess that $A = D = z\mathbb{1}_2$ is a solution, where $z \in \mathbb{C}$ is an appropriate normalization factor. Substitute this back into Eq. (1.8) to learn that $2|z|^2 = 1$, thus a valid choice is $z = 1/\sqrt{2}$. Putting everything together, we get $A = B = D = \mathbb{1}_2/\sqrt{2}$, and $C = -D = -\mathbb{1}_2/\sqrt{2}$. This returns the desired result. ■

Question 2

To improve readability, in Questions 2 and 3 we label spacetime indices using the Roman alphabet—$a, b, \ldots \in \{0, 1, 2, 3\}$—rather than the Greek.

Lemma 2.1: *The commutator of two Dirac matrices satisfies* $[\gamma^a, \gamma^b] = 2(\gamma^a\gamma^b - \eta^{ab})$.

Proof.—By definition, we have that

$$\gamma^a\gamma^b = \frac{1}{2}[\gamma^a, \gamma^b] + \frac{1}{2}\{\gamma^a, \gamma^b\}. \tag{2.1}$$

Rearranging this equation while using Eq. (1.1) returns the desired result. ■

Proposition 2.1: *The Dirac matrices satisfy*

$$[\gamma^a\gamma^b, \gamma^c\gamma^d] = 2\eta^{bc}\gamma^a\gamma^d - 2\eta^{ac}\gamma^b\gamma^d + 2\eta^{bd}\gamma^c\gamma^a - 2\eta^{ad}\gamma^c\gamma^b. \tag{2.2}$$

Proof.—Expand out the commutator to find

$$\begin{aligned}[\gamma^a\gamma^b, \gamma^c\gamma^d] &= [\gamma^a\gamma^b, \gamma^c]\gamma^d + \gamma^c[\gamma^a\gamma^b, \gamma^d] \\ &= \gamma^a[\gamma^b, \gamma^c]\gamma^d + [\gamma^a, \gamma^c]\gamma^b\gamma^d + \gamma^c\gamma^a[\gamma^b, \gamma^d] + \gamma^c[\gamma^a, \gamma^d]\gamma^b.\end{aligned} \tag{2.3}$$

The next step is to use Lemma 2.1. Schematically, each term will yield $\gamma[\gamma, \gamma]\gamma \sim 2(\gamma\gamma\gamma\gamma - \eta\gamma\gamma)$. We want to engineer it in such a way that all terms with four factors of γ cancel each other. We achieve this by exploiting the antisymmetry $[A, B] = -[B, A]$ of the commutator to write

$$[\gamma^a\gamma^b, \gamma^c\gamma^d] = -\gamma^a[\gamma^c, \gamma^b]\gamma^d + [\gamma^a, \gamma^c]\gamma^b\gamma^d - \gamma^c\gamma^a[\gamma^d, \gamma^b] + \gamma^c[\gamma^a, \gamma^d]\gamma^b. \tag{2.4}$$

Using Lemma 2.1 now neatly recovers the desired result. ■

We define the matrices $S^{ab} = \frac{1}{4}[\gamma^a, \gamma^b]$. Using Lemma 2.1, it is easy to see that $S^{ab} = \frac{1}{2}(\gamma^a\gamma^b - \eta^{ab})$.

Lemma 2.2: *The matrices* S^{ab} *satisfy* $[S^{ab}, \gamma^c] = \gamma^a\eta^{bc} - \gamma^b\eta^{ac}$.

Proof.—Expand the LHS to find

$$[S^{ab}, \gamma^c] = \frac{1}{2}[\gamma^a\gamma^b, \gamma^c] = \frac{1}{2}\left(\gamma^a[\gamma^b, \gamma^c] + [\gamma^a, \gamma^c]\gamma^b\right) = \frac{1}{2}\left(-\gamma^a[\gamma^c, \gamma^b] + [\gamma^a, \gamma^c]\gamma^b\right). \tag{2.5}$$

In the last step, we make use of the antisymmetry of the commutator such that, when we use Lemma 2.1, the terms with three factors of γ cancel each other and leave us with the desired result. ■

Proposition 2.2: *The six matrices* S^{ab} *form a representation of the Lie algebra of the Lorentz group,*

$$[S^{ab}, S^{cd}] = S^{ad}\eta^{bc} - S^{bd}\eta^{ac} + S^{ca}\eta^{bd} - S^{cb}\eta^{ad}. \tag{2.6}$$

Proof.—Use the definition of S^{cd} on the LHS to find

$$[S^{ab}, S^{cd}] = \frac{1}{2}[S^{ab}, \gamma^c\gamma^d] = \frac{1}{2}[S^{ab}, \gamma^c]\gamma^d + \frac{1}{2}\gamma^c[S^{ab}, \gamma^d]. \tag{2.7}$$

Using Lemma 2.2, this further simplifies to

$$\begin{aligned}[S^{ab}, S^{cd}] &= \frac{1}{2}\left(\gamma^a\eta^{bc} - \gamma^b\eta^{ac}\right)\gamma^d + \frac{1}{2}\gamma^c\left(\gamma^a\eta^{bd} - \gamma^b\eta^{ad}\right) \\ &= \frac{1}{2}\gamma^a\gamma^d\eta^{bc} - \frac{1}{2}\gamma^b\gamma^d\eta^{ac} + \frac{1}{2}\gamma^c\gamma^a\eta^{bd} - \frac{1}{2}\gamma^c\gamma^b\eta^{ad}.\end{aligned} \tag{2.8}$$

Finally, use the definition of S^{ab} to write $\frac{1}{2}\gamma^a\gamma^d = S^{ad} + \frac{1}{2}\eta^{ad}$. All terms with two factors of η cancel each other, leaving us with the desired result. ■

Question 3

In the interest of efficiency, it will be convenient to prove the desired identities slightly out of order.

Proposition (d): *The Dirac matrix $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ satisfies $(\gamma^5)^2 = 1$.*

Proof.—The Clifford algebra in Eq. (1.1) tells us that γ^μ and γ^ν anticommute with each other when $\mu \neq \nu$. This fact can be exploited to show that

$$(\gamma^5)^2 = -\gamma^0\gamma^1\gamma^2\gamma^3\gamma^0\gamma^1\gamma^2\gamma^3 = -(-1)^{3+2+1}(\gamma^0)^2(\gamma^1)^2(\gamma^2)^2(\gamma^3)^2. \tag{3.1}$$

We pick up three factors of -1 when we anticommute the second γ^0 further to the right until it meets the first γ^0, another two factors of -1 to anticommute γ^1, and one additional factor of -1 to anticommute γ^2. The Clifford algebra also tells us that $(\gamma^0)^2 = 1$ and $(\gamma^i)^2 = -1$. These can be used to deduce the desired result. ■

Lemma 3.1: *The Dirac matrix γ^5 anticommutes with all the other Dirac matrices, $\{\gamma^\mu, \gamma^5\} = 0$.*

Proof.—Consider the quantity $\gamma^5\gamma^a$. Since γ^a will anticommute with three of the four Dirac matrices contained within γ^5, and will of course commute with itself, it follows that

$$\gamma^5\gamma^a = (-1)^3 i\gamma^a\gamma^0\gamma^1\gamma^2\gamma^3 = -\gamma^a\gamma^5, \tag{3.2}$$

where each factor of -1 comes from each anticommutation. ■

Proposition (a): $\operatorname{tr}\gamma^a = 0$

Proof.—We insert the identity $1 = (\gamma^5)^2$ into the trace and use its cyclicity property to write

$$\operatorname{tr}\gamma^a = \operatorname{tr}(\gamma^5\gamma^5\gamma^a) = \operatorname{tr}(\gamma^5\gamma^a\gamma^5). \tag{3.3}$$

We now use the fact that γ^5 anticommutes with γ^a to find $\operatorname{tr}(\gamma^5\gamma^a\gamma^5) = -\operatorname{tr}(\gamma^5\gamma^5\gamma^a) = -\operatorname{tr}\gamma^a$, which implies that $\operatorname{tr}\gamma^a = 0$. In fact, it was not necessary to use γ^5 to prove this; any matrix M that anticommutes with γ^a and satisfies $M^2 = 1$ will do the job. For instance, one could have used $M = \gamma^0$ to prove that $\operatorname{tr}\gamma^i = 0$, and then used a different choice $M = i\gamma^i$ to prove that $\operatorname{tr}\gamma^0 = 0$. ■

Proposition (c): $\operatorname{tr}(\gamma^a\gamma^b\gamma^c) = 0$ *and, more generally,* $\operatorname{tr}(\gamma^{a_1}\cdots\gamma^{a_n}) = 0$ *for all odd integers n.*

Proof.—The proof proceeds identically as for Proposition (a). Suppose there exists a matrix M that anticommutes with γ^a for all $a \in \{0,1,2,3\}$ and that satisfies $M^2 = 1$. Clearly, a valid choice is $M = \gamma^5$. Insert the identity $1 = M^2$, use the cyclicity of the trace, and anticommute M with all the other Dirac matrices to find $\operatorname{tr}(\gamma^{a_1}\cdots\gamma^{a_n}) = (-1)^n \operatorname{tr}(\gamma^{a_1}\cdots\gamma^{a_n})$, from which one deduces the desired result. ■

Proposition (e): $\operatorname{tr}\gamma^5 = 0$

Proof.—The proof is again identical, except we use $1 = (\gamma^0)^2$, or $1 = -(\gamma^i)^2$, for the identity. ■

Proposition (b): $\operatorname{tr}(\gamma^a\gamma^b) = 4\eta^{ab}$

Proof.—Using the Clifford algebra, we write $\operatorname{tr}(\gamma^a\gamma^b) = \operatorname{tr}(2\eta^{ab}\mathbb{1}_4) - \operatorname{tr}(\gamma^b\gamma^a)$. The second term on the RHS is equal to $\operatorname{tr}(\gamma^a\gamma^b)$ due to the cyclicity of the trace, hence $\operatorname{tr}(\gamma^a\gamma^b) = \operatorname{tr}(\eta^{ab}\mathbb{1}_4)$. Then using the fact that $\operatorname{tr}(\mathbb{1}_4) = 4$ yields the desired result. ■

Proposition (f): $\not{p}\not{q} = 2p\cdot q - \not{q}\not{p} = p\cdot q + 2S^{ab}p_a q_b$

Proof.—The first equality follows from using the Clifford algebra, $\gamma^a\gamma^b = 2\eta^{ab} - \gamma^b\gamma^a$. Contract this with $p_a q_b$ to get $\not{p}\not{q} = 2p\cdot q - \not{q}\not{p}$. The second equality comes from using the identity $S^{ab} = \frac{1}{2}(\gamma^a\gamma^b - \eta^{ab})$ established in Question 2. Contract it with $p_a q_b$ and rearrange to get $\not{p}\not{q} = p\cdot q + 2S^{ab}p_a q_b$. ■

Proposition (g): $\operatorname{tr}(\not{p}\not{q}) = 4p\cdot q$

Proof.—Contract the result of Proposition (b) with $p_a q_b$ to obtain the desired result. Alternatively, one can also take the trace of Proposition (f) while using the fact that $\operatorname{tr}([\gamma^a,\gamma^b]) = 0$ due to the cyclicity of the trace. ■

Proposition (h): $\operatorname{tr}(\not{p}_1\cdots\not{p}_n) = 0$ *for all odd integers* n

Proof.—Contract the result of Proposition (c) with $(p_1)_{a_1}\cdots(p_n)_{a_n}$ to get the desired result. ■

Proposition (i): $\operatorname{tr}(\not{p}_1\not{p}_2\not{p}_3\not{p}_4) = 4\left[(p_1\cdot p_2)(p_3\cdot p_4) + (p_1\cdot p_4)(p_2\cdot p_3) - (p_1\cdot p_3)(p_2\cdot p_4)\right]$

Proof.—We begin by considering the object

$$\begin{aligned}\operatorname{tr}(\gamma^a\gamma^b\gamma^c\gamma^d) &= 2\eta^{ab}\operatorname{tr}(\gamma^c\gamma^d) - \operatorname{tr}(\gamma^b\gamma^a\gamma^c\gamma^d)\\ &= 2\eta^{ab}\operatorname{tr}(\gamma^c\gamma^d) - 2\eta^{ac}\operatorname{tr}(\gamma^b\gamma^d) + \operatorname{tr}(\gamma^b\gamma^c\gamma^a\gamma^d)\\ &= 2\eta^{ab}\operatorname{tr}(\gamma^c\gamma^d) - 2\eta^{ac}\operatorname{tr}(\gamma^b\gamma^d) + 2\eta^{ad}\operatorname{tr}(\gamma^b\gamma^c) - \operatorname{tr}(\gamma^b\gamma^c\gamma^d\gamma^a),\end{aligned} \tag{3.4}$$

having made judicious application of the Clifford algebra three times. The final term on the last line is equal to $\operatorname{tr}(\gamma^a\gamma^b\gamma^c\gamma^d)$ due to cyclicity of the trace, thus

$$\begin{aligned}\operatorname{tr}(\gamma^a\gamma^b\gamma^c\gamma^d) &= \eta^{ab}\operatorname{tr}(\gamma^c\gamma^d) - \eta^{ac}\operatorname{tr}(\gamma^b\gamma^d) + \eta^{ad}\operatorname{tr}(\gamma^b\gamma^c)\\ &= 4(\eta^{ab}\eta^{cd} - \eta^{ac}\eta^{bd} + \eta^{ad}\eta^{bc}),\end{aligned} \tag{3.5}$$

having used Proposition (b) to obtain the second line. Contracting this with $(p_1)_a(p_2)_b(p_3)_c(p_4)_d$ yields the desired result. ■

Proposition (j): $\operatorname{tr}(\gamma^5\not{p}_1\not{p}_2) = 0$

Proof.—We start by considering the object $\operatorname{tr}(\gamma^5\gamma^a\gamma^b)$. Since there are only three Dirac matrices within the trace, we can always find a matrix M that anticommutes with all three matrices $\{\gamma^5,\gamma^a,\gamma^b\}$ and satisfies $M^2 = 1$. Specifically, we can choose $M = \sqrt{\eta_{cc}}\gamma^c$ for any $c \notin \{a,b,5\}$. Then using the same arguments as in the proof of Proposition (c), we learn that $\operatorname{tr}(\gamma^5\gamma^a\gamma^b) = 0$. Contracting this with $(p_1)_a(p_2)_b$ yields the desired result. ■

Proposition (k): $\gamma_a \not{p} \gamma^a = -2\not{p}$

Proof.—Using the Clifford algebra, we can write $\gamma^a\gamma^b\gamma^c = -\gamma^a\gamma^c\gamma^b + \gamma^a 2\eta^{bc}$. Contract this with $\eta_{ac}p_b$ to find $\gamma_a \not{p} \gamma^a = -(\gamma^a\gamma^c\eta_{ac})\not{p} + 2\not{p} = -2\not{p}$, having used the identity $\gamma^a\gamma^c\eta_{ac} = \frac{1}{2}\{\gamma^a, \gamma^c\}\eta_{ac} = 4$ in the last step. ■

Proposition (l): $\gamma_a \not{p}_1 \not{p}_2 \gamma^a = 4p_1 \cdot p_2$

Proof.—In a similar vein to Proposition (k), what we would like to do is take the γ^a matrix on the right and push it through until it meets γ_a on the left. We find

$$\gamma_a \not{p}_1 \not{p}_2 \gamma^a = \gamma_a \not{p}_1 \underbrace{(-\gamma^a\gamma^b + 2\eta^{ab})}_{\gamma^b\gamma^a}(p_2)_b = \gamma_a \not{p}_1(-\gamma^a \not{p}_2 + 2p_2^a) = -\gamma_a \not{p}_1 \gamma^a \not{p}_2 + 2\not{p}_2 \not{p}_1. \tag{3.6}$$

Now use Proposition (k) to simplify the last expression, such that

$$\gamma_a \not{p}_1 \not{p}_2 \gamma^a = 2\not{p}_1 \not{p}_2 + 2\not{p}_2 \not{p}_1 = 2\{\not{p}_1, \not{p}_2\} = 2\{\gamma^a, \gamma^b\}(p_1)_a(p_2)_b = 4p_1 \cdot p_2, \tag{3.7}$$

as desired. ■

Proposition (m): $\gamma_a \not{p}_1 \not{p}_2 \not{p}_3 \gamma^a = -2\not{p}_3 \not{p}_2 \not{p}_1$

Proof.—The proof proceeds in a similar fashion as for Proposition (l). We use the Clifford algebra to write

$$\gamma_a \not{p}_1 \not{p}_2 \not{p}_3 \gamma^a = \gamma_a \not{p}_1 \not{p}_2(-\gamma^a \not{p}_3 + 2p_3^a) = -\gamma_a \not{p}_1 \not{p}_2 \gamma^a \not{p}_3 + 2\not{p}_3 \not{p}_1 \not{p}_2. \tag{3.8}$$

Now use Proposition (l) to simplify this, yielding

$$\gamma_a \not{p}_1 \not{p}_2 \not{p}_3 \gamma^a = -(4p_1 \cdot p_2)\not{p}_3 + 2\not{p}_3 \not{p}_1 \not{p}_2 = -2\not{p}_3 \underbrace{(2\eta^{ab} - \gamma^a\gamma^b)}_{\gamma^b\gamma^a}(p_1)_a(p_2)_b = -2\not{p}_3 \not{p}_2 \not{p}_1, \tag{3.9}$$

as desired. ■

Proposition (n): $\mathrm{tr}(\gamma^5 \not{p}_1 \not{p}_2 \not{p}_3 \not{p}_4) = 4i\epsilon_{abcd}p_1^a p_2^b p_3^c p_4^d$

Proof.—First consider the properties of the object $S^{abcd} = \mathrm{tr}(\gamma^5\gamma^a\gamma^b\gamma^c\gamma^d)$. Using the Clifford algebra on $\gamma^a\gamma^b$, we learn that

$$\mathrm{tr}(\gamma^5\gamma^a\gamma^b\gamma^c\gamma^d) = -\,\mathrm{tr}(\gamma^5\gamma^b\gamma^a\gamma^c\gamma^d) + 2\eta^{ab}\,\mathrm{tr}(\gamma^5\gamma^c\gamma^d). \tag{3.10}$$

The second term on the RHS vanishes according to Proposition (j), thus we learn that S^{abcd} is antisymmetric in its first two indices. We can repeat the same procedure for any adjacent pair of Dirac matrices. This is sufficient to deduce that S^{abcd} is totally antisymmetric, so must be proportional to the Levi–Civita symbol,[2] $S^{abcd} \propto \epsilon^{abcd}$. Regardless of our metric signature convention, a four-dimensional spacetime is taken to have $\epsilon_{0123} = +1$. Raising indices with the Minkowski metric tells us that $\epsilon^{0123} = -1$, hence

$$S^{abcd} = -\epsilon^{abcd}\,\mathrm{tr}(\gamma^5\gamma^0\gamma^1\gamma^2\gamma^3) = i\epsilon^{abcd}\,\mathrm{tr}\big((\gamma^5)^2\big) = 4i\epsilon^{abcd}, \tag{3.11}$$

having used Proposition (d) in the last step. Contracting this with $(p_1)_a(p_2)_b(p_3)_c(p_4)_d$ yields the desired result. ■

[2] Also often called the totally antisymmetric tensor or the alternating symbol.

Question 4

The Dirac equation $(i\partial\!\!\!/ - m)\psi = 0$ admits plane-wave solutions $\psi = u(\mathbf{p})e^{-ip\cdot x}$ and $\psi = v(\mathbf{p})e^{ip\cdot x}$, with

$$u^s(\mathbf{p}) = \begin{pmatrix} \sqrt{p\cdot\sigma}\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\xi^s \end{pmatrix}, \quad v^s(\mathbf{p}) = \begin{pmatrix} \sqrt{p\cdot\sigma}\xi^s \\ -\sqrt{p\cdot\bar{\sigma}}\xi^s \end{pmatrix}; \tag{4.1}$$

where $\sigma^\mu = (1, \boldsymbol{\sigma})$, $\bar{\sigma}^\mu = (1, -\boldsymbol{\sigma})$, and ξ^s forms an orthonormal basis for 2-spinors satisfying

$$(\xi^r)^\dagger \cdot \xi^s = \delta^{rs}. \tag{4.2}$$

Before we begin proving the desired identities, it is worthwhile thinking about what it really means to take the root $\sqrt{p\cdot\sigma}$. Since the 2×2 matrices σ^μ are hermitian, it follows that $p\cdot\sigma$ is hermitian. It is a fact that any hermitian matrix H is diagonalizable, meaning there exists a matrix U such that $H = UDU^{-1}$ and D is a diagonal matrix. It is easy to see that the columns of U are the eigenvectors of H, while the diagonal elements of D are its eigenvalues.

If $D = \mathrm{diag}(d_1, \ldots, d_n)$ is an $n\times n$ diagonal matrix, its root is $\sqrt{D} = \mathrm{diag}(\sqrt{d_1}, \ldots, \sqrt{d_n})$. Now notice that

$$(U\sqrt{D}U^{-1})^2 = U\sqrt{D}U^{-1}U\sqrt{D}U^{-1} = U\sqrt{D}\sqrt{D}U^{-1} = UDU^{-1} = H.$$

Hence, we learn that the root of a diagonalizable matrix H is $\sqrt{H} = U\sqrt{D}U^{-1}$. The takeaway message here is that since $p\cdot\sigma$ is hermitian and therefore diagonalizable, we always know how to determine its root $\sqrt{p\cdot\sigma}$. In other words, it is a well-defined quantity.

Lemma 4.1: $(p\cdot\sigma)(p\cdot\bar{\sigma}) = m^2$

Proof.—We prove this by direct evaluation. Starting with the LHS, we find

$$(p\cdot\sigma)(p\cdot\bar{\sigma}) = p_\mu p_\nu \sigma^\mu\bar{\sigma}^\nu = (p^0)^2 + p_0 p_i \underbrace{(\bar{\sigma}^i + \sigma^i)}_{0} + p_i p_j \underbrace{\sigma^i\bar{\sigma}^j}_{-\sigma^i\sigma^j} = m^2, \tag{4.3}$$

where the last step follows from using Eq. (1.3) and the fact that $p^2 = m^2$. ■

Lemma 4.2: $[(p\cdot\sigma), (p\cdot\bar{\sigma})] = 0$

Proof.—Again, we can prove this by direct evaluation. One finds

$$[(p\cdot\sigma), (p\cdot\bar{\sigma})] = p_\mu p_\nu[\sigma^\mu, \bar{\sigma}^\nu] = -p_i p_j[\sigma^i, \sigma^j]. \tag{4.4}$$

The last step follows since $\sigma^0 = \bar{\sigma}^0 = \mathbb{1}_2$ commutes with everything. Using Eq. (1.3), we see that $[\sigma^i, \sigma^j] \propto \epsilon^{ijk}\sigma^k$. Since this is being contracted with the symmetric object $p_i p_j$ in the equation above, we deduce that the commutator must vanish. ■

One more fact is needed about the matrices $p\cdot\sigma$ and $p\cdot\bar{\sigma}$ before we can satisfactorily prove the desired identities. We first need a theorem from undergraduate mathematics, which we will state without proof.

Theorem 4.1: *Two matrices A and B are simultaneously diagonalizable if and only if they commute.*

Lemma 4.3: $\sqrt{p\cdot\sigma}\sqrt{p\cdot\bar{\sigma}} = \sqrt{(p\cdot\sigma)(p\cdot\bar{\sigma})}$

Proof.—Consider two matrices H_1 and H_2. In general, $\sqrt{H_1}\sqrt{H_2} \neq \sqrt{H_1 H_2}$. To see this, take the square of the LHS to find $(\sqrt{H_1}\sqrt{H_2})^2 = \sqrt{H_1}\sqrt{H_2}\sqrt{H_1}\sqrt{H_2}$. This is equal to $(\sqrt{H_1 H_2})^2 = H_1 H_2$ only if $\left[\sqrt{H_1}, \sqrt{H_2}\right] = 0$. Let us now consider the specific case $H_1 = p\cdot\sigma$ and $H_2 = p\cdot\bar{\sigma}$. From Lemma 4.2, we know that $[H_1, H_2] = 0$. Combined with Theorem 4.1, this implies that H_1 and H_2 are simultaneously diagonalizable, meaning we can write

$$H_i = UD_iU^{-1}, \quad \sqrt{H_i} = U\sqrt{D_i}U^{-1} \tag{4.5}$$

for $i \in \{1, 2\}$. This immediately tells us that

$$[\sqrt{H_1}, \sqrt{H_2}] = [U\sqrt{D_1}U^{-1}, U\sqrt{D_2}U^{-1}] = U[\sqrt{D_1}, \sqrt{D_2}]U^{-1} = 0, \tag{4.6}$$

since diagonal matrices commute. This completes the proof. ■

These three lemmas tell us that the intuition we have for how square roots behave for real numbers will also work for the matrices $p \cdot \sigma$ and $p \cdot \bar{\sigma}$ because they are hermitian and commute with each other. We can now turn to addressing the question proper.

Proposition 4.1: *The inner products between the plane-wave solutions and their hermitian conjugates are*

$$u^r(\mathbf{p})^\dagger \cdot u^s(\mathbf{p}) = 2p^0\delta^{rs}, \tag{4.7a}$$
$$v^r(\mathbf{p})^\dagger \cdot v^s(\mathbf{p}) = 2p^0\delta^{rs}, \tag{4.7b}$$
$$u^r(\mathbf{p})^\dagger \cdot v^s(-\mathbf{p}) = 0. \tag{4.7c}$$

Proof.—Using its definition in Eq. (4.1), we find

$$u^r(\mathbf{p})^\dagger \cdot u^s(\mathbf{p}) = \left(\xi^{r\dagger}\sqrt{p\cdot\sigma}, \xi^{r\dagger}\sqrt{p\cdot\bar{\sigma}}\right)\begin{pmatrix}\sqrt{p\cdot\sigma}\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\xi^s\end{pmatrix} = \xi^{r\dagger}\left[p\cdot\sigma + p\cdot\bar{\sigma}\right]\xi^s = 2p_0\,\xi^{r\dagger}\cdot\xi^s. \tag{4.8}$$

Note that $(\sqrt{p\cdot\sigma})^\dagger = \sqrt{p\cdot\sigma}$ since σ^μ are hermitian. Making use of Eq. (4.2) and noting that $p_0 \equiv p^0$ in our conventions, we get the desired result in Eq. (4.7a). Exactly the same procedure can be used to prove Eq. (4.7b). To prove the last identity, let us define $\bar{p}^\mu = (p^0, -\mathbf{p})$ for a given 4-momentum $p^\mu = (p^0, \mathbf{p})$. Then

$$\begin{aligned}
u^r(\mathbf{p})^\dagger \cdot v^s(-\mathbf{p}) &= \left(\xi^{r\dagger}\sqrt{p\cdot\sigma}, \xi^{r\dagger}\sqrt{p\cdot\bar{\sigma}}\right)\begin{pmatrix}\sqrt{\bar{p}\cdot\sigma}\xi^s \\ -\sqrt{\bar{p}\cdot\bar{\sigma}}\xi^s\end{pmatrix} \\
&= \xi^{r\dagger}\left[\sqrt{p\cdot\sigma}\sqrt{\bar{p}\cdot\sigma} - \sqrt{p\cdot\bar{\sigma}}\sqrt{\bar{p}\cdot\bar{\sigma}}\right]\xi^s \\
&= \xi^{r\dagger}\left[\sqrt{p\cdot\sigma}\sqrt{p\cdot\bar{\sigma}} - \sqrt{p\cdot\bar{\sigma}}\sqrt{p\cdot\sigma}\right]\xi^s.
\end{aligned} \tag{4.9}$$

The third line follows since $p\cdot\bar{\sigma} = \bar{p}\cdot\sigma$ and $\bar{p}\cdot\bar{\sigma} = p\cdot\sigma$. Lemmas 4.2 and 4.3 can now be used to show that this vanishes. ■

Definition—*The Dirac adjoint $\bar{\psi}$ is defined as $\bar{\psi} = \psi^\dagger\gamma^0$.*

Proposition 4.2: *The inner products between these plane-wave solutions and their Dirac adjoints are*

$$\bar{u}^r(\mathbf{p}) \cdot u^s(\mathbf{p}) = 2m\delta^{rs}, \tag{4.10a}$$
$$\bar{v}^r(\mathbf{p}) \cdot v^s(\mathbf{p}) = -2m\delta^{rs}, \tag{4.10b}$$
$$\bar{u}^r(\mathbf{p}) \cdot v^s(\mathbf{p}) = 0. \tag{4.10c}$$

Proof.—It is instructive to prove this in two ways. One option is to proceed just like we did for Proposition 4.1. The only difference is that there is an extra γ^0 matrix in between the two Dirac spinors. The solutions in Eq. (4.1) are written in the chiral representation, so γ^0 is given by Eq. (1.2).

One therefore finds

$$\bar{u}^r(\mathbf{p}) \cdot u^s(\mathbf{p}) = \xi^{r\dagger} \left[\sqrt{p \cdot \sigma} \sqrt{p \cdot \bar{\sigma}} + \sqrt{p \cdot \bar{\sigma}} \sqrt{p \cdot \sigma} \right] \xi^s = \xi^{r\dagger} \left[2\sqrt{(p \cdot \sigma)(p \cdot \bar{\sigma})} \right] \xi^s, \tag{4.11}$$

having used Lemmas 4.2 and 4.3 in the last step. Lemma 4.1 and Eq. (4.2) can then be used to show that this is equal to $2m\delta^{rs}$. Similar steps can be used to prove the other identities.

While this is all fine, there is actually a much simpler way. Since $\bar{\psi}\psi$ is a Lorentz scalar for any Dirac spinor ψ,[3] we can evaluate the inner product in any frame. In the particle's rest frame, $p \cdot \sigma = p \cdot \bar{\sigma} = m$, hence

$$u^s(\mathbf{0}) = \sqrt{m} \begin{pmatrix} \xi^s \\ \xi^s \end{pmatrix}, \quad v^s(\mathbf{0}) = \sqrt{m} \begin{pmatrix} \xi^s \\ -\xi^s \end{pmatrix}. \tag{4.12}$$

It follows that

$$\bar{v}^r(\mathbf{0}) \cdot v^s(\mathbf{0}) = m \left(-\xi^{r\dagger}, \xi^{r\dagger} \right) \begin{pmatrix} \xi^s \\ -\xi^s \end{pmatrix} = -2m(\xi^r)^\dagger \cdot \xi^s = -2m\delta^{rs}. \tag{4.13}$$

As this is true in all frames, we get Eq. (4.10b) as desired. Similar steps can be used to prove the other identities. ■

Question 5

Proposition 5.1: *The plane-wave solutions to the Dirac equation satisfy the "spin sum relations"*

$$\sum_s u^s(\mathbf{p})\bar{u}^s(\mathbf{p}) = \not{p} + m, \quad \sum_s v^s(\mathbf{p})\bar{v}^s(\mathbf{p}) = \not{p} - m. \tag{5.1}$$

Proof.—Using the definitions in Eq. (4.1), we get

$$\begin{aligned} \sum_s u^s(\mathbf{p})\bar{u}^s(\mathbf{p}) &= \sum_s \begin{pmatrix} \sqrt{p \cdot \sigma}\xi^s \\ \sqrt{p \cdot \bar{\sigma}}\xi^s \end{pmatrix} \left(\sqrt{p \cdot \sigma}\xi^{s\dagger}, \sqrt{p \cdot \bar{\sigma}}\xi^{s\dagger} \right) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} p \cdot \sigma & \sqrt{(p \cdot \sigma)(p \cdot \bar{\sigma})} \\ \sqrt{(p \cdot \bar{\sigma})(p \cdot \sigma)} & p \cdot \bar{\sigma} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} m & p \cdot \sigma \\ p \cdot \bar{\sigma} & m \end{pmatrix}, \end{aligned} \tag{5.2}$$

where judicious use has been made of Lemmas 4.1–4.3. In obtaining the second line, we have also used the identity $\sum_s \xi^s(\xi^s)^\dagger = \mathbb{1}_2$. This is obviously true if we choose the basis $\xi^1 = (1, 0)$ and $\xi^2 = (0, 1)$. A new basis that remains orthonormal can then be constructed via a unitary transformation $\xi^s \to U\xi^s$. One can then show that $\sum_s \xi^s(\xi^s)^\dagger = \mathbb{1}_2$ is true in any basis. The last line of Eq. (5.2) is indeed equal to $\not{p} + m$, hence we obtain the desired result. The same can be done to prove the identity for v. ■

Question 6

The Dirac field and its hermitian conjugate admit the Fourier decomposition

$$\psi(\mathbf{x}) = \sum_s \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left[b_p^s u^s(\mathbf{p}) e^{i\mathbf{p}\cdot\mathbf{x}} + c_p^{s\dagger} v^s(\mathbf{p}) e^{-i\mathbf{p}\cdot\mathbf{x}} \right], \tag{6.1a}$$

$$\psi^\dagger(\mathbf{x}) = \sum_s \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left[b_p^{s\dagger} u^s(\mathbf{p})^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}} + c_p^s v^s(\mathbf{p})^\dagger e^{i\mathbf{p}\cdot\mathbf{x}} \right] \tag{6.1b}$$

[3] See, e.g., Claim 4.3 of David Tong's notes [1] for a proof.

in terms of the creation and annihilation operators (b, c), which satisfy

$$\{b_p^r, b_q^{s\dagger}\} = (2\pi)^3 \delta^{rs} \delta^{(3)}(\mathbf{p}-\mathbf{q}), \quad \{c_p^r, c_q^{s\dagger}\} = (2\pi)^3 \delta^{rs} \delta^{(3)}(\mathbf{p}-\mathbf{q}), \tag{6.2}$$

with all other anticommutations vanishing.

Proposition 6.1: *The anticommutation relations in Eq. (6.2) imply that*

$$\{\psi_\alpha(\mathbf{x}), \psi_\beta(\mathbf{y})\} = \{\psi_\alpha^\dagger(\mathbf{x}), \psi_\beta^\dagger(\mathbf{y})\} = 0, \tag{6.3a}$$

$$\{\psi_\alpha(\mathbf{x}), \psi_\beta^\dagger(\mathbf{y})\} = \delta_{\alpha\beta} \delta^{(3)}(\mathbf{x}-\mathbf{y}). \tag{6.3b}$$

Proof.—This just follows by direct evaluation. Proving Eq. (6.3a) is trivial, since schematically we have $\{\psi, \psi\} \sim \{b + c^\dagger, b + c^\dagger\} = 0$, using Eq. (6.2) in the last step. The same is true for $\{\psi^\dagger, \psi^\dagger\} = 0$. Thus, all we have to do is prove Eq. (6.3b). The LHS evaluates to

$$\begin{aligned}
\{\psi_\alpha(\mathbf{x}), \psi_\beta^\dagger(\mathbf{y})\} &= \sum_{r,s} \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \int \frac{\mathrm{d}^3 q}{(2\pi)^3} \frac{1}{\sqrt{2E_q}} \\
&\quad \times \left\{ b_p^r u_\alpha^r(\mathbf{p}) e^{i\mathbf{p}\cdot\mathbf{x}} + c_p^{r\dagger} v_\alpha^r(\mathbf{p}) e^{-i\mathbf{p}\cdot\mathbf{x}},\ b_q^{s\dagger} u_\beta^s(\mathbf{q})^\dagger e^{-i\mathbf{q}\cdot\mathbf{y}} + c_q^s v_\beta^s(\mathbf{q})^\dagger e^{i\mathbf{q}\cdot\mathbf{y}} \right\} \\
&= \sum_{r,s} \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \int \frac{\mathrm{d}^3 q}{(2\pi)^3} \frac{1}{\sqrt{2E_q}} \\
&\quad \times \left[\{b_p^r, b_q^{s\dagger}\}\, u_\alpha^r(\mathbf{p}) u_\beta^s(\mathbf{q})^\dagger\, e^{i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{y})} + \{c_p^{r\dagger}, c_q^s\}\, v_\alpha^r(\mathbf{p}) v_\beta^s(\mathbf{q})^\dagger\, e^{-i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{y})} \right].
\end{aligned} \tag{6.4}$$

Now use the anticommutation relations in Eq. (6.2) and integrate over $\mathbf{q}$ to get

$$\begin{aligned}
\{\psi_\alpha(\mathbf{x}), \psi_\beta^\dagger(\mathbf{y})\} &= \sum_s \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{2E_p} \left[u_\alpha^s(\mathbf{p}) \bar{u}_{\dot\beta}^s(\mathbf{p}) e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} + v_\alpha^s(\mathbf{p}) \bar{v}_{\dot\beta}^s(\mathbf{p}) e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} \right] (\gamma^0)^{\dot\beta}{}_\beta \\
&= \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{2E_p} \left[(\not{p} + m)_{\alpha\dot\beta}\, e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} + (\not{p} - m)_{\alpha\dot\beta}\, e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} \right] (\gamma^0)^{\dot\beta}{}_\beta,
\end{aligned} \tag{6.5}$$

having written $u^\dagger = \bar{u}\gamma^0$, and likewise for v, in the first line. This facilitates use of Proposition 5.1 to obtain the second line. Note that $\dot\beta$ is just another index in addition to α and β. We now expand $\not{p} = \gamma^0 p_0 + \gamma^i p_i$ and relabel $\mathbf{p} \to -\mathbf{p}$ in the second term to get

$$\begin{aligned}
\{\psi_\alpha(\mathbf{x}), \psi_\beta^\dagger(\mathbf{y})\} &= \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{2E_p} \left[(\gamma^0 p_0 + \gamma^i p_i + m) + (\gamma^0 p_0 - \gamma^i p_i - m) \right]_{\alpha\dot\beta} (\gamma^0)^{\dot\beta}{}_\beta e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} \\
&= \delta_{\alpha\beta} \delta^{(3)}(\mathbf{x}-\mathbf{y}).
\end{aligned} \tag{6.6}$$

To obtain the last line, we have integrated over $\mathbf{p}$ and used the fact that $(\gamma^0)^2 = \mathbb{1}_4$. ■

Question 7

Proposition 7.1: *After normal ordering, the Hamiltonian for the free Dirac field is*

$$:H: = \int \mathrm{d}^3 x\ :\bar\psi(-i\gamma^i \partial_i + m)\psi: = \sum_s \int \frac{\mathrm{d}^3 p}{(2\pi)^3} E_p \left(b_p^{s\dagger} b_p^s + c_p^{s\dagger} c_p^s \right). \tag{7.1}$$

Proof.—Let us begin by considering the object

$$(-i\gamma^i \partial_i + m)\psi = (-i\gamma^i \partial_i + m) \sum_s \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left[b_p^s u^s(\mathbf{p}) e^{i\mathbf{p}\cdot\mathbf{x}} + c_p^{s\dagger} v^s(\mathbf{p}) e^{-i\mathbf{p}\cdot\mathbf{x}} \right]$$

$$
\begin{aligned}
&= \sum_s \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left[b_p^s (-\gamma^i p_i + m) u^s(\mathbf{p}) e^{i\mathbf{p}\cdot\mathbf{x}} + c_p^{s\dagger} (\gamma^i p_i + m) v^s(\mathbf{p}) e^{-i\mathbf{p}\cdot\mathbf{x}} \right] \\
&= \sum_s \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \gamma^0 p_0 \left[b_p^s u^s(\mathbf{p}) e^{i\mathbf{p}\cdot\mathbf{x}} - c_p^{s\dagger} v^s(\mathbf{p}) e^{-i\mathbf{p}\cdot\mathbf{x}} \right]. \qquad (7.2)
\end{aligned}
$$

In obtaining the second line, there is a subtle minus sign associated with the dot product $\mathbf{p}\cdot\mathbf{x} = \delta_{ij} p^i x^i = -\eta_{ij} p^i x^j = -p_i x^i$, which we pick up when differentiating the exponential.[4] In the third line, we have used the fact that u and v are plane-wave solutions to the Dirac equation, satisfying

$$
\begin{aligned}
(\not{p} - m) u = 0 \quad &\Rightarrow \quad (-\gamma^i p_i + m) u = \gamma^0 p_0 u, \\
(\not{p} + m) v = 0 \quad &\Rightarrow \quad (\gamma^i p_i + m) v = -\gamma^0 p_0 v. \qquad (7.3)
\end{aligned}
$$

Multiply this on the left with $\bar{\psi}(\mathbf{x})$ and integrate over $\mathbf{x}$ to obtain the Hamiltonian,

$$
\begin{aligned}
H &= \int \mathrm{d}^3 x \; \bar{\psi}(-i\gamma^i p_i + m)\psi \\
&= \sum_{r,s} \int \frac{\mathrm{d}^3 q}{(2\pi)^3} \frac{1}{\sqrt{2E_q}} \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \sqrt{\frac{E_p}{2}} \int \mathrm{d}^3 x \\
&\quad \times \left[b_q^{r\dagger} \bar{u}^r(\mathbf{q}) e^{-i\mathbf{q}\cdot\mathbf{x}} + c_q^r \bar{v}^r(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{x}} \right] \gamma^0 \left[b_p^s u^s(\mathbf{p}) e^{i\mathbf{p}\cdot\mathbf{x}} - c_p^{s\dagger} v^s(\mathbf{p}) e^{-i\mathbf{p}\cdot\mathbf{x}} \right] \\
&= \sum_{r,s} \int \frac{\mathrm{d}^3 q}{(2\pi)^3} \frac{1}{\sqrt{2E_q}} \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \sqrt{\frac{E_p}{2}} \int \mathrm{d}^3 x \\
&\quad \times \Big\{ \left[b_q^{r\dagger} b_p^s \, u^r(\mathbf{q})^\dagger u^s(\mathbf{p}) - c_q^r c_p^{s\dagger} \, v^r(\mathbf{q})^\dagger v^s(\mathbf{p}) \right] e^{i(\mathbf{p}-\mathbf{q})\cdot\mathbf{x}} \\
&\qquad + \left[c_q^r b_p^s \, v^r(\mathbf{q})^\dagger u^s(\mathbf{p}) - b_q^{r\dagger} c_p^{s\dagger} \, u^r(\mathbf{q})^\dagger v^s(\mathbf{p}) \right] e^{i(\mathbf{p}+\mathbf{q})\cdot\mathbf{x}} \Big\}. \qquad (7.4)
\end{aligned}
$$

To obtain the last line, we have used the freedom to relabel $\mathbf{x} \to -\mathbf{x}$ in half of the terms. Performing the integral over $\mathbf{x}$ will now generate delta functions that impose the conditions $\mathbf{q} = \mathbf{p}$ or $\mathbf{q} = -\mathbf{p}$. Integrating over $\mathbf{q}$ then yields

$$
\begin{aligned}
H = \frac{1}{2} \sum_{r,s} \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \Big[& b_p^{r\dagger} b_p^s \, u^r(\mathbf{p})^\dagger u^s(\mathbf{p}) - c_p^r c_p^{s\dagger} \, v^r(\mathbf{p})^\dagger v^s(\mathbf{p}) \\
& + c_{-p}^r b_p^s \, v^r(-\mathbf{p})^\dagger u^s(\mathbf{p}) - b_{-p}^{r\dagger} c_p^{s\dagger} \, u^r(-\mathbf{p})^\dagger v^s(\mathbf{p}) \Big]. \qquad (7.5)
\end{aligned}
$$

The results of Proposition 4.1 enable us to simplify the terms in square brackets. In particular, Eq. (4.7c) and its hermitian conjugate tells us that the last two terms vanish. Also using Eqs. (4.7a) and (4.7b), we get

$$
H = \sum_s \int \frac{\mathrm{d}^3 p}{(2\pi)^3} E_p \left(b_p^{s\dagger} b_p^s - c_p^s c_p^{s\dagger} \right). \qquad (7.6)
$$

This returns the desired result after normal ordering. Note that fermionic operators pick up a minus sign when normal ordered (i.e., $:cc^\dagger: = -c^\dagger c$) due to the anticommutation relations in Eq. (6.2). ■

[4] We have met this minus sign before in Question 1 of Sheet 2.

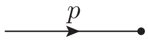
Question 8

Consider nucleon-nucleon scattering, $\psi\psi \to \psi\psi$, in Yukawa theory with Lagrangian

$$\mathcal{L} = \bar{\psi}(i\partial\!\!\!/ - m)\psi + \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}\mu^2\phi^2 - \lambda\bar{\psi}\psi\phi. \tag{8.1}$$

The Feynman rules relevant for this process are as follows:

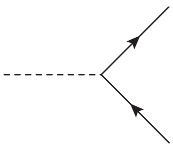
- Each ingoing fermion gets a factor of $u^s(p)$, and each outgoing fermion gets a factor of $\bar{u}^s(p)$:

p, s  $= u^s(p)$,  $p, s = \bar{u}^s(p)$.

- Each internal scalar, denoted by a dotted line, gets a relevant factor of its propagator:

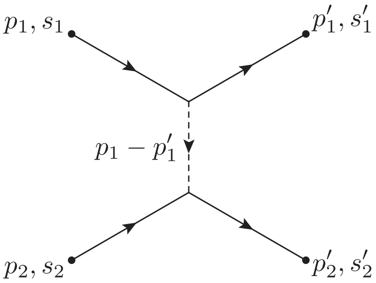
 $= \dfrac{i}{p^2 - \mu^2}$.

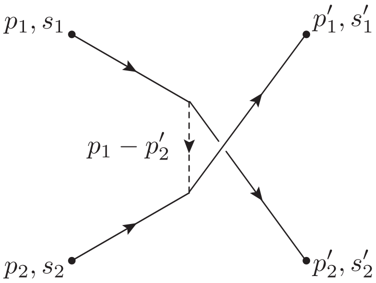
- Each interaction vertex picks up a factor of $-i\lambda$:

 $= -i\lambda$.

- Add an extra minus sign for Fermi-Dirac statistics.

These rules tell us that nucleon-nucleon scattering has the tree-level amplitude

$i\mathcal{M} =$ 

$+$ 

$$= i(-i\lambda)^2 \left(\frac{[\bar{u}^{s_1'}(p_1') \cdot u^{s_1}(p_1)][\bar{u}^{s_2'}(p_2') \cdot u^{s_2}(p_2)]}{(p_1 - p_1')^2 - \mu^2} - \frac{[\bar{u}^{s_1'}(p_1') \cdot u^{s_2}(p_2)][\bar{u}^{s_2'}(p_2') \cdot u^{s_1}(p_1)]}{(p_1 - p_2')^2 - \mu^2} \right). \tag{8.2}$$

To contract the 4-component spinors in the right way, we start from the head (the direction in which the arrow is pointing) of each solid line and work our way backwards to the tail. For instance, in the t-channel diagram on the left, the head of the top solid line gives us a factor of $\bar{u}^{s_1'}(p_1')$. We contract this with the tail of this line, which gives a factor of $u^{s_1}(p_1)$. Similarly, the bottom line gives $\bar{u}^{s_2'}(p_2') \cdot u^{s_2}(p_2)$. Note that the u-channel diagram has an overall minus sign relative to the t-channel diagram, which comes from Fermi-Dirac statistics associated with swapping the labels $(p_1', s_1') \leftrightarrow (p_2', s_2')$ on the outgoing particles.

Lemma 8.1: *If all four external particles have the same mass m, the inner products between the external momenta satisfy the relations*

$$p_1 \cdot p_2 = p_1' \cdot p_2' = \frac{1}{2}(s - 2m^2), \tag{8.3a}$$

$$p_1 \cdot p_1' = p_2 \cdot p_2' = -\frac{1}{2}(t - 2m^2), \tag{8.3b}$$

$$p_1 \cdot p_2' = p_1' \cdot p_2 = -\frac{1}{2}(u - 2m^2). \tag{8.3c}$$

Proof.—By definition, the Mandelstam variable $s = (p_1 + p_2)^2$. Expanding the dot product and using the fact that $(p_1)^2 = (p_2)^2 = m^2$, we get $s = 2m^2 + 2p_1 \cdot p_2$. By momentum conservation, $s = (p'_1 + p'_2)^2$ also, which implies $s = 2m^2 + 2p'_1 \cdot p'_2$. These can be rearranged to yield Eq. (8.3a). The same procedure can be followed to obtain the other results. ■

Proposition 8.1: *The spin-averaged probability for nucleon-nucleon scattering is proportional to*

$$\langle |\mathcal{M}|^2 \rangle = \lambda^4 \left(\frac{(u-4m^2)^2}{(u-\mu^2)^2} + \frac{(t-4m^2)^2}{(t-\mu^2)^2} + \frac{(s-4m^2)^2 - (u-4m^2)^2 - (t-4m^2)^2}{2(u-\mu^2)(t-\mu^2)} \right). \tag{8.4}$$

Proof.—Let us write $u_i \equiv u^{s_i}(p_i)$ and $\bar{u}_i \equiv \bar{u}^{s_i}(p_i)$ for $i \in \{1, 2\}$. These enable us to write the matrix element in Eq. (8.2) more compactly as

$$\mathcal{M} = (-i\lambda)^2 \left(\frac{(\bar{u}'_1 u_1)(\bar{u}'_2 u_2)}{t-\mu^2} - \frac{(\bar{u}'_1 u_2)(\bar{u}'_2 u_1)}{u-\mu^2} \right). \tag{8.5}$$

In general, the inner products $\bar{u}'_i u_j$ are complex numbers, so we cannot just take the square of this object like we would for scattering processes involving only scalar fields. Instead, we first compute the conjugate

$$\mathcal{M}^* = (-i\lambda)^2 \left(\frac{(\bar{u}_1 u'_1)(\bar{u}_2 u'_2)}{t-\mu^2} - \frac{(\bar{u}_2 u'_1)(\bar{u}_1 u'_2)}{u-\mu^2} \right), \tag{8.6}$$

which we then multiply by $\mathcal{M}$ to get

$$\begin{aligned} |\mathcal{M}|^2 = \lambda^4 \Bigg(& \frac{(\bar{u}'_1 u_2)(\bar{u}'_2 u_1)(\bar{u}_2 u'_1)(\bar{u}_1 u'_2)}{(u-\mu^2)^2} + \frac{(\bar{u}'_1 u_1)(\bar{u}'_2 u_2)(\bar{u}_1 u'_1)(\bar{u}_2 u'_2)}{(t-\mu^2)^2} \\ & - \frac{(\bar{u}'_1 u_1)(\bar{u}'_2 u_2)(\bar{u}_2 u'_1)(\bar{u}_1 u'_2) + (\bar{u}'_1 u_2)(\bar{u}'_2 u_1)(\bar{u}_1 u'_1)(\bar{u}_2 u'_2)}{(u-\mu^2)(t-\mu^2)} \Bigg), \end{aligned} \tag{8.7}$$

where $t = (p_1 - p'_1)^2$ and $u = (p_1 - p'_2)^2$ are the usual Mandelstam variables. For reasons that will become clear later, let us write this as

$$|\mathcal{M}|^2 = \lambda^4 \left(\frac{\Phi_{uu}}{(u-\mu^2)^2} + \frac{\Phi_{tt}}{(t-\mu^2)^2} - \frac{\Phi_{tu} + \Phi_{ut}}{(u-\mu^2)(t-\mu^2)} \right). \tag{8.8}$$

In an experiment, we usually cannot control whether an ingoing fermion is produced with spin up or spin down. Both options are generally produced with equal probability. When we monitor the outgoing particles, we may similarly lack the ability to measure their spins, or we may wish to be agnostic and only ask questions that are blind to the spins of the final state. In such cases, we consider the spin-averaged quantity

$$\langle |\mathcal{M}|^2 \rangle = \left(\frac{1}{2} \sum_{s_1} \right) \left(\frac{1}{2} \sum_{s_2} \right) \sum_{s'_1} \sum_{s'_2} |\mathcal{M}|^2 = \frac{1}{4} \sum_{\text{spins}} |\mathcal{M}|^2. \tag{8.9}$$

While we use the conventional term "spin average," what we really mean is that we average over the ingoing spins and sum over all possible outgoing spins.

We can now get a simplified expression for $\langle |\mathcal{M}|^2 \rangle$ by exploiting the identities in Proposition 5.1. To see how this goes, first consider one of the strings of inner products that reads

$$\begin{aligned} \sum \Phi_{uu} = \sum (\bar{u}'_1 u_2)(\bar{u}'_2 u_1)(\bar{u}_2 u'_1)(\bar{u}_1 u'_2) &= \sum (\bar{u}'_1 u_2)(\bar{u}_2 u'_1)(\bar{u}'_2 u_1)(\bar{u}_1 u'_2) \\ &= \sum \left[\bar{u}'_1 (\not{p}_2 + m) u'_1 \right] \left[\bar{u}'_2 (\not{p}_1 + m) u'_2 \right]. \end{aligned} \tag{8.10}$$

The first line follows since each inner product $(\bar{u} \cdot u)$ is just a complex number that commutes with all other complex numbers. We have chosen an especially convenient arrangement such that $\bar{u}$ appears directly to the right of u. This facilitates the use of Proposition 5.1, which gives us the second line after summing over s_1 and s_2. We can simplify this further by noting that for any inner product $\bar{\psi}A\psi$, where A is some 4×4 matrix,

$$\bar{\psi}A\psi = \bar{\psi}_\alpha A^\alpha{}_\beta \psi^\beta = \psi^\beta \bar{\psi}_\alpha A^\alpha{}_\beta = \mathrm{tr}(\psi\bar{\psi}A). \tag{8.11}$$

This identity can be used to show that

$$\sum \left[\bar{u}_1'(\not{p}_2 + m)u_1'\right] = \sum \mathrm{tr}[u_1'\bar{u}_1'(\not{p}_2 + m)] = \mathrm{tr}[(\not{p}_1' + m)(\not{p}_2 + m)]. \tag{8.12}$$

Now we use the results from Question 3 to simplify things even further. We know that the trace over an odd number of Dirac matrices vanishes, hence

$$\mathrm{tr}[(\not{p}_1' + m)(\not{p}_2 + m)] = \mathrm{tr}(\not{p}_1'\not{p}_2) + \mathrm{tr}(m^2) = 4p_1' \cdot p_2 + 4m^2 = -2(u - 4m^2), \tag{8.13}$$

where the last steps follow from using Proposition (g) and Lemma 8.1. Doing the same for $[\bar{u}_2'(\not{p}_1 + m)u_2']$, we end up with

$$\langle \Phi_{uu} \rangle = \frac{1}{4} \sum \Phi_{uu} = (u - 4m^2)^2. \tag{8.14}$$

Exactly the same steps tell us that

$$\langle \Phi_{tt} \rangle = (t - 4m^2)^2. \tag{8.15}$$

Finally, we are left with the task of simplifying

$$\Phi_{tu} = (\bar{u}_1'u_1)(\bar{u}_1u_2')(\bar{u}_2'u_2)(\bar{u}_2u_1'), \quad \Phi_{ut} = (\bar{u}_1'u_2)(\bar{u}_2u_2')(\bar{u}_2'u_1)(\bar{u}_1u_1'). \tag{8.16}$$

The two are related by the interchange $(p_1, s_1) \leftrightarrow (p_2, s_2)$, hence it suffices to evaluate the spin average of only one of them. We find

$$\begin{aligned}
\langle \Phi_{tu} \rangle &= \frac{1}{4} \sum \bar{u}_1(\not{p}_1 + m)(\not{p}_2' + m)(\not{p}_2 + m)u_1' \\
&= \frac{1}{4}\, \mathrm{tr}\left[(\not{p}_1' + m)(\not{p}_1 + m)(\not{p}_2' + m)(\not{p}_2 + m)\right] \\
&= \frac{1}{4}\Big\{ \mathrm{tr}(\not{p}_1'\not{p}_1\not{p}_2'\not{p}_2) + m^2 \Big[\mathrm{tr}(\not{p}_1'\not{p}_1) + \mathrm{tr}(\not{p}_1'\not{p}_2') + \mathrm{tr}(\not{p}_1'\not{p}_2) \\
&\qquad + \mathrm{tr}(\not{p}_1\not{p}_2') + \mathrm{tr}(\not{p}_1\not{p}_2) + \mathrm{tr}(\not{p}_2'\not{p}_2)\Big] + 4m^4 \Big\}.
\end{aligned} \tag{8.17}$$

As before, any term that traces over an odd number of Dirac matrices vanishes. Propositions (g) and (i) can now be used to simplify this further, yielding

$$\begin{aligned}
\langle \Phi_{tu} \rangle &= (p_1' \cdot p_1)(p_2' \cdot p_2) + (p_1' \cdot p_2)(p_1 \cdot p_2') - (p_1' \cdot p_2')(p_1 \cdot p_2) \\
&\quad + m^2\left(p_1' \cdot p_1 + p_1' \cdot p_2' + p_1' \cdot p_2 + p_1 \cdot p_2' + p_1 \cdot p_2 + p_2' \cdot p_2\right) + m^4 \\
&= \frac{1}{4}(t - 2m^2)^2 + \frac{1}{4}(u - 2m^2)^2 - \frac{1}{4}(s - 2m^2)^2 + m^2(s - t - u + 2m^2) + m^4 \\
&= \frac{1}{4}\left[t^2 + u^2 - s^2 + 8m^2(s - t - u) + 16m^4\right] \\
&= \frac{1}{4}\left[(t - 4m^2)^2 + (u - 4m^2)^2 - (s - 4m^2)^2\right],
\end{aligned} \tag{8.18}$$

where the second line follows from using Lemma 8.1, and the fourth from using $s + t + u = 4m^2$. The interchange $p_1 \leftrightarrow p_2$ is equivalent to interchanging $t \leftrightarrow u$, hence it follows that $\langle \Phi_{ut} \rangle = \langle \Phi_{tu} \rangle$. Putting everything together, the spin-averaged probability for nucleon-nucleon scattering is proportional to

$$\left\langle |\mathcal{M}|^2 \right\rangle = \lambda^4 \left(\frac{\langle \Phi_{uu} \rangle}{(u - \mu^2)^2} + \frac{\langle \Phi_{tt} \rangle}{(t - \mu^2)^2} - \frac{2\,\langle \Phi_{tu} \rangle}{(u - \mu^2)(t - \mu^2)} \right), \tag{8.19}$$

which yields the desired result upon substitution of Eqs. (8.14), (8.15), and (8.18).[5] ■

Corollary: *In the centre-of-mass frame, nucleon-nucleon scattering has the spin-averaged differential cross section*

$$\left\langle \frac{\mathrm{d}\sigma}{\mathrm{d}t} \right\rangle = \frac{\lambda^4}{16\pi s(s-4m^2)} \left(\frac{(u-4m^2)^2}{(u-\mu^2)^2} + \frac{(t-4m^2)^2}{(t-\mu^2)^2} + \frac{(s-4m^2)^2-(u-4m^2)^2-(t-4m^2)^2}{2(u-\mu^2)(t-\mu^2)} \right). \tag{8.21}$$

Proof.—From Eq. (16) of Ben Allanach's notes [2], any $2 \to 2$ process has the differential cross section

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = \frac{|\mathcal{M}|^2}{16\pi\lambda(s, m_1^2, m_2^2)} \tag{8.22}$$

in its centre-of-mass frame, with $\lambda(x,y,z) = x^2+y^2+z^2-2xy-2xz-2yz$.[6] Setting the masses of the ingoing particles to $m_1 = m_2 = m$, we get

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = \frac{|\mathcal{M}|^2}{16\pi s(s-4m^2)}. \tag{8.23}$$

The denominator is spin-independent, hence taking the spin average of this quantity is equivalent to replacing $|\mathcal{M}|^2$ with $\langle |\mathcal{M}|^2 \rangle$ in the numerator. Substituting in the result from Proposition 8.1 then returns Eq. (8.21). ■

The total spin-averaged cross section $\langle\sigma\rangle$ can then be obtained by integrating over t and dividing by a symmetry factor of 2! (because the two outgoing particles are indistinguishable),[7]

$$\langle\sigma\rangle = \frac{1}{2}\int_{t_{\min}}^{t_{\max}} \mathrm{d}t \left\langle \frac{\mathrm{d}\sigma}{\mathrm{d}t} \right\rangle. \tag{8.24}$$

Since all external particles are identical, the ingoing and outgoing momenta for the first nucleon can be written as $p_1^\mu = (E, \mathbf{p})$ and $p_1'^\mu = (E, \mathbf{p}')$ in the centre-of-mass frame, where $|\mathbf{p}| = |\mathbf{p}'|$. The Mandelstam variable

$$t = (p_1 - p_1')^2 = -|\mathbf{p} - \mathbf{p}'|^2 = -2|\mathbf{p}|^2(1-\cos\theta), \tag{8.25}$$

where $\theta \in [0, \pi]$ is the angle between $\mathbf{p}$ and $\mathbf{p}'$. It follows that we should integrate over the range $t \in [-4|\mathbf{p}|^2, 0]$. Since $s = (2E)^2 = 4|\mathbf{p}|^2 + 4m^2$, the integration range can also be written as $t \in [-(s-4m^2), 0]$.

References

[1] D. Tong, *Quantum Field Theory*, (2007) http://www.damtp.cam.ac.uk/user/tong/qft.html.

[2] B. C. Allanach, *QFT: Cross Sections and Decay Rates*, (2018) http://www.damtp.cam.ac.uk/user/examples/3P1l.pdf.

[3] M. Srednicki, *Quantum Field Theory* (Cambridge University Press, Cambridge, England, 2007).

---

[5] In fact, using the identity $s+t+u = 4m^2$ judiciously gives us a simpler (but equivalent) expression,

$$\left\langle |\mathcal{M}|^2 \right\rangle = \lambda^4 \left(\frac{(u-4m^2)^2}{(u-\mu^2)^2} + \frac{(t-4m^2)^2}{(t-\mu^2)^2} + \frac{ut - 4m^2 s}{(u-\mu^2)(t-\mu^2)} \right). \tag{8.20}$$

[6] An equivalent expression is $\lambda(s, m_1^2, m_2^2) = [s-(m_1+m_2)^2][s-(m_1-m_2)^2]$.

[7] We need the symmetry factor because merely integrating over all the outgoing momenta treats the final state as being labeled by an ordered list of these momenta. But if some outgoing particles are identical, this is not correct; the momenta of the identical particles should be specified by an unordered list. The symmetry factor provides the appropriate correction. See, e.g., Chapter 11 of Srednicki [3].