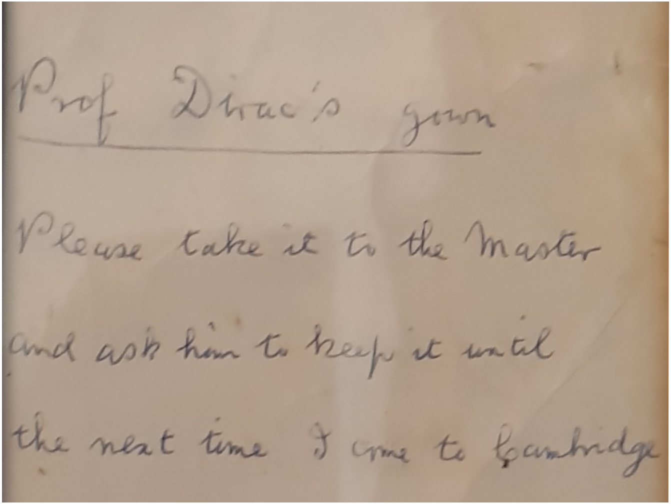
Dr C.G. Lester, 2023



Prof Dirac's gown

Please take it to the master
and ask him to keep it until
the next time I come to Cambridge

H2: The Dirac Equation

Non-Relativistic QM (Revision)

- For particle physics need a relativistic formulation of quantum mechanics
- Take as the starting point non-relativistic energy:

$$E = T + V = \frac{\vec{p}^2}{2m} + V$$

- In QM we identify the energy and momentum operators:

$$\vec{p} \to -i\vec{\nabla}, \quad E \to i\frac{\partial}{\partial t}$$

- which gives the time dependent Schrödinger equation (take V=0 for simplicity)

$$-\frac{1}{2m}\vec{\nabla}^2\psi = i\frac{\partial\psi}{\partial t} \tag{12}$$

with plane wave solutions: $\psi = Ne^{i(\vec{p}.\vec{r}-Et)}$, where $i\frac{\partial\psi}{\partial t} = E\psi$, $-i\nabla\psi = \vec{p}\psi$

- The SE is first order in the time derivatives and second order in spatial derivatives – and is manifestly not Lorentz invariant.
- In what follows we will use probability density/current extensively. For the non-relativistic case these are derived as follows. Firstly, (12) implies that

$$-\frac{1}{2m}\vec{\nabla}^2\psi^* = -i\frac{\partial\psi^*}{\partial t} \tag{13}$$

- $\psi^* \times (12) - \psi \times (13)$:

$$
\begin{aligned}
-\frac{1}{2m}\left(\psi^*\nabla^2\psi - \psi\nabla^2\psi^*\right) &= i\left(\psi^*\frac{\partial\psi}{\partial t} + \psi\frac{\partial\psi^*}{\partial t}\right) \\
-\frac{1}{2m}\vec{\nabla}.\left(\psi^*\vec{\nabla}\psi - \psi\vec{\nabla}\psi^*\right) &= i\frac{\partial}{\partial t}(\psi^*\psi)
\end{aligned}
$$

- Which by comparison with the continuity equation

$$\vec{\nabla}.\vec{j} + \frac{\partial\rho}{\partial t} = 0$$

leads to the following expressions for probability density and current: $\rho = \psi^*\psi = |\psi|^2$,

$$\vec{j} = \frac{1}{2mi}\left(\psi^*\vec{\nabla}\psi - \psi\vec{\nabla}\psi^*\right)$$

- For a plane wave $\psi = Ne^{i(\vec{p}.\vec{r}-Et)}$

$$\rho = |N|^2$$

and

$$\vec{j} = |N|^2, \qquad \text{and} \qquad \frac{\vec{p}}{m} = |N|^2\vec{v}$$

- The number of particles per unit volume is $|N|^2$
- For $|N|^2$ particles per unit volume moving at velocity $\vec{v}$, have $|N|^2|\vec{v}|$ passing through a unit area per unit time (particle flux). Therefore $\vec{j}$ is a vector in the particle's direction with magnitude equal to the flux.

The Klein-Gordon Equation

- Applying $\vec{p} \to -i\vec{\nabla}, \quad E \to i\partial/\partial t$ to the relativistic equation for energy:

$$E^2 = |\vec{p}|^2 + m^2 \tag{14}$$

gives the Klein-Gordon equation:

$$\frac{\partial^2 \psi}{\partial t^2} = \vec{\nabla}^2 \psi - m^2 \psi \tag{15}$$

- Using $\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$ $\partial^\mu \partial_\mu \equiv \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}$
- KG can be expressed compactly as

$$(\partial^\mu \partial_\mu + m^2)\psi = 0 \tag{16}$$

- For plane wave solutions, $\psi = Ne^{i(\vec{p}.\vec{r} - Et)}$, the KG equation gives: $-E^2\psi = -|\vec{p}|^2\psi - m^2\psi \to E = \pm\sqrt{|\vec{p}|^2 + m^2}$
- Not surprisingly, the KG equation has negative energy solutions – they are allowed by our starting equation (14).
- Historically the –ve energy solutions were viewed as problematic. But for the KG there is also a problem with the probability density ...

- Proceeding as before to calculate the probability and current densities, the complex conjugate of (15) is:

$$\frac{\partial^2 \psi^*}{\partial t^2} = \vec{\nabla}^2 \psi^* - m^2 \psi^* \tag{17}$$

$\psi^* \times (15) - \psi \times (17)$:

$$\begin{aligned} \psi^* \frac{\partial^2 \psi}{\partial t^2} - \psi \frac{\partial^2 \psi^*}{\partial t^2} &= \psi^* (\nabla^2 \psi - m^2 \psi) - \psi (\nabla^2 \psi^* - m^2 \psi^*) \\ \frac{\partial}{\partial t} \left(\psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t} \right) &= \vec{\nabla}.(\psi^* \vec{\nabla} \psi - \psi \vec{\nabla} \psi^*) \end{aligned}$$

- Which, again, by comparison with the continuity equation allows us to identify

$$\rho = i \left(\psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t} \right) \quad \text{and} \quad \vec{j} = i(\psi^* \vec{\nabla} \psi - \psi \vec{\nabla} \psi^*)$$

- For a plane wave $\psi = N e^{i(\vec{p}.\vec{r} - Et)}$ so that

$$\rho = 2E|N|^2 \text{ and } \vec{j} = 2\vec{p}|N|^2$$

- Particle densities are proportional to E and thus to γ. We might have anticipated this from the previous discussion of Lorentz invariant phase space (i.e. density of 1 in the particle's rest frame will appear as γ in a frame where the particle has energy E due to length contraction).

The Dirac Equation

Historically, it was thought that there were two main problems with the Klein-Gordon equation:

- Negative energy solutions
- The negative particle densities associated with these solutions $\rho = 2E|N|^2$

We now know that **in Quantum Field Theory these problems do not arise** and the KG equation is used to describe spin-0 particles (inherently single particle description $\rightarrow$ multi-particle quantum excitations of a scalar field).

Nevertheless:

- These problems motivated Dirac (1928) to search for a different formulation of relativistic quantum mechanics in which all particle densities are positive.
- As we will see, the solutions of the resulting wave equation solved not only this problem but also explained intrinsic spin and why antiparticles exist.
- [Not examinable: The magnetic moment of the electron is also explained by the Dirac Equation. See Appendix IV.]

The Dirac Equation :

- Schrödinger eqn: $-\frac{1}{2m}\vec{\nabla}^2\psi = i\frac{\partial\psi}{\partial t}$ is 1st order in $\partial/\partial t$
- Klein-Gordon $(\partial^\mu\partial_\mu + m^2)\psi = 0$ is 2nd order in $\partial/\partial x, \partial/\partial y, \partial/\partial z$
- Dirac looked for an alternative which was 1st order throughout:

$$\hat{H}\psi = (\vec{\alpha}.\vec{p} + \beta m)\psi = i\frac{\partial\psi}{\partial t} \tag{18}$$

where $\hat{H}$ is the Hamiltonian operator and, as usual $\vec{p} = -i\vec{\nabla}$

- Writing (18) in full:

$$\left(-i\alpha_x\frac{\partial}{\partial x} - i\alpha_y\frac{\partial}{\partial y} - i\alpha_z\frac{\partial}{\partial z} + \beta m\right)\psi = \left(i\frac{\partial}{\partial t}\right)\psi$$

"squaring" this equation gives

$$\left(-i\alpha_x\frac{\partial}{\partial x} - i\alpha_y\frac{\partial}{\partial y} - i\alpha_z\frac{\partial}{\partial z} + \beta m\right)\left(-i\alpha_x\frac{\partial}{\partial x} - i\alpha_y\frac{\partial}{\partial y} - i\alpha_z\frac{\partial}{\partial z} + \beta m\right)\psi = -\frac{\partial^2\psi}{\partial t^2}$$

- Which can be expanded in gory details as...

$$
\begin{aligned}
-\frac{\partial^2\psi}{\partial t^2} \quad = \quad & -\alpha_x^2\frac{\partial^2\psi}{\partial x^2} - \alpha_y^2\frac{\partial^2\psi}{\partial y^2} - \alpha_z^2\frac{\partial^2\psi}{\partial z^2} + \beta^2 m^2\psi \\
& -(\alpha_x\alpha_y + \alpha_y\alpha_x)\frac{\partial^2\psi}{\partial x\partial y} - (\alpha_y\alpha_z + \alpha_z\alpha_y)\frac{\partial^2\psi}{\partial y\partial z} - (\alpha_z\alpha_x + \alpha_x\alpha_z)\frac{\partial^2\psi}{\partial z\partial x} \\
& -(\alpha_x\beta + \beta\alpha_x)m\frac{\partial\psi}{\partial x} - (\alpha_y\beta + \beta\alpha_y)m\frac{\partial\psi}{\partial y} - (\alpha_z\beta + \beta\alpha_z)m\frac{\partial\psi}{\partial z}
\end{aligned}
$$

- For this to be a reasonable formulation of relativistic QM, a free particle must also obey $E^2 = \vec{p}^2 + m^2$, i.e. it must satisfy the Klein-Gordon equation:

$$
-\frac{\partial^2\psi}{\partial t^2} = -\frac{\partial^2\psi}{\partial x^2} - \frac{\partial^2\psi}{\partial y^2} - \frac{\partial^2\psi}{\partial z^2} + m^2\psi
$$

- Hence for the Dirac Equation to be consistent with the KG equation require:

$$
\begin{aligned}
\alpha_x^2 = \alpha_y^2 = \alpha_z^2 = \beta^2 = 1 && (19) \\
\alpha_j\beta + \beta\alpha_j = 0 && (20) \\
\alpha_j\alpha_k + \alpha_k\alpha_j = 0 \qquad (j \neq k) && (21)
\end{aligned}
$$

- Immediately we see that the α_j and β cannot be numbers. Require four mutually anti-commuting matrices.
- Must be (at least) 4x4 matrices (see Appendix III)

Consequence: these Dirac Spinors have at least four components!

Somewhat surprising conclusion:

A consequence of introducing a Lorentz Covariant wave equation that is first-order in time/space derivatives is that the wave-function has to have extra degrees of freedom we didn't previously know we would need!

Schematically:

$$\text{(first-order Lorentz-covariant wave equation)} \quad \Longrightarrow \quad \psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}$$

A representation of Dirac's $\vec{\alpha}$ and β matrices

We also want the Hamiltonian $\hat{H}\psi = (\vec{\alpha}.\vec{p} + \beta m)\psi$ to be Hermitian.

This extra constraint forces:

$$\alpha_x = \alpha_x^\dagger; \qquad \alpha_y = \alpha_y^\dagger; \qquad \alpha_z = \alpha_z^\dagger; \qquad \beta = \beta^\dagger; \tag{22}$$

i.e. we need to find four anti-commuting Hermitian 4x4 matrices.

At this point it is convenient to introduce an explicit representation for $\vec{\alpha}, \beta$. It should be noted that physical results do not depend on the particular representation – everything important could be derived from the anti-commutation relations (19)-(21).

A convenient choice is

$$\beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \qquad \alpha_j = \begin{pmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{pmatrix}$$

based on the Pauli spin matrices:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Dirac Equation: Probability Density and Current

- Now consider probability density/current – this is where the perceived problems with the Klein-Gordon equation arose.
- Start with the Dirac equation

$$-i\alpha_x \frac{\partial \psi}{\partial x} - i\alpha_y \frac{\partial \psi}{\partial y} - i\alpha_z \frac{\partial \psi}{\partial z} + m\beta\psi = i\frac{\partial \psi}{\partial t} \quad (23)$$

and its Hermitian conjugate

$$+i\frac{\partial \psi^\dagger}{\partial x}\alpha_x^\dagger + i\frac{\partial \psi^\dagger}{\partial y}\alpha_y^\dagger + i\frac{\partial \psi^\dagger}{\partial z}\alpha_z^\dagger + m\psi^\dagger\beta^\dagger = -i\frac{\partial \psi^\dagger}{\partial t} \quad (24)$$

- Consider $\psi^\dagger \times (23) - (24) \times \psi$ remembering α, β are Hermitian

$$\psi^\dagger\left(-i\alpha_x \frac{\partial \psi}{\partial x} - i\alpha_y \frac{\partial \psi}{\partial y} - i\alpha_z \frac{\partial \psi}{\partial z} + \beta m\psi\right) - \left(i\frac{\partial \psi^\dagger}{\partial x}\alpha_x + i\frac{\partial \psi^\dagger}{\partial y}\alpha_y + i\frac{\partial \psi^\dagger}{\partial z}\alpha_z + m\psi^\dagger\beta\right)\psi$$

$$= i\psi^\dagger\frac{\partial \psi}{\partial t} + i\frac{\partial \psi^\dagger}{\partial t}\psi$$

$\rightarrow$

$$\psi^\dagger\left(\alpha_x \frac{\partial \psi}{\partial x} + \alpha_y \frac{\partial \psi}{\partial y} + \alpha_z \frac{\partial \psi}{\partial z}\right) + \left(\frac{\partial \psi^\dagger}{\partial x}\alpha_x + \frac{\partial \psi^\dagger}{\partial y}\alpha_y + \frac{\partial \psi^\dagger}{\partial z}\alpha_z\right)\psi + \frac{\partial(\psi^\dagger\psi)}{\partial t} = 0$$

- Now using the identity:

$$\psi^\dagger\alpha_x\frac{\partial \psi}{\partial x} + \frac{\partial \psi^\dagger}{\partial x}\alpha_x\psi \equiv \frac{\partial(\psi^\dagger\alpha_x\psi)}{\partial x}$$

- gives the continuity equation

$$\vec{\nabla}.(\psi^\dagger\vec{\alpha}\psi) + \frac{\partial(\psi^\dagger\psi)}{\partial t} = 0 \tag{25}$$

where $\psi^\dagger = (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*)$

- The probability density and current can be identified as:

$$\rho = \psi^\dagger\psi$$

and

$$\vec{j} = \psi^\dagger\vec{\alpha}\psi$$

where $\rho = \psi^\dagger\psi = |\psi_1|^2 + |\psi_2|^2 + |\psi_3|^2 + |\psi_4|^2 > 0$

- Unlike the KG equation, the Dirac equation has probability densities which are always positive.
- In addition, the solutions to the Dirac equation are the four component Dirac Spinors. A great success of the Dirac equation is that these extra components naturally give rise to the property of intrinsic spin and antiparticles. (See (43) on page 98 for discussion of why Dirac spinors represent spin-half particles.)
- Such particles have an intrinsic magnetic moment of

$$\vec{\mu} = \frac{q}{m}\vec{S} \qquad \text{(see Appendix IV).}$$

Covariant Notation for Dirac Equation using the Dirac Gamma Matrices

- The Dirac equation can be written more elegantly by introducing the four Dirac gamma matrices: $\gamma^0 \equiv \beta; \quad \gamma^1 \equiv \beta\alpha_x; \quad \gamma^2 \equiv \beta\alpha_y; \quad \gamma^3 \equiv \beta\alpha_z$
- Premultiply the Dirac equation (23) by β

$$i\beta\alpha_x\frac{\partial\psi}{\partial x} + i\beta\alpha_y\frac{\partial\psi}{\partial y} + i\beta\alpha_z\frac{\partial\psi}{\partial z} - \beta^2 m\psi = -i\beta\frac{\partial\psi}{\partial t}$$

$$\rightarrow \gamma^1\frac{\partial\psi}{\partial x} + i\gamma^2\frac{\partial\psi}{\partial y} + i\gamma^3\frac{\partial\psi}{\partial z} - m\psi = -i\gamma^0\frac{\partial\psi}{\partial t}$$

- using $\partial_\mu = (\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$ this can be written compactly as:

$$\boxed{(i\gamma^\mu\partial_\mu - m)\psi = 0} \qquad (26)$$

- NOTE: it is important to realise that the Dirac gamma matrices are not four-vectors - they are constant matrices which remain invariant under a Lorentz transformation. However it can be shown that the Dirac equation is itself Lorentz covariant (see page 143 of Appendix V).

Properties of the gamma matrices

From the properties of the α and β matrices (19)-(21) one immediately obtains:

$$(\gamma^0)^2 = \beta^2 = 1 \qquad \text{and} \qquad (\gamma^1)^2 = \beta\alpha_x\,\beta\alpha_x = -\alpha_x\beta\beta\alpha_x = -\alpha_x^2 = -1.$$

The full set of relations is

$$\begin{aligned} (\gamma^0)^2 &= 1 \\ (\gamma^1)^2 = (\gamma^2)^2 = (\gamma^3)^2 &= -1 \\ \gamma^0\gamma^j + \gamma^j\gamma^0 &= 0 \\ \gamma^j\gamma^k + \gamma^k\gamma^j &= 0 \qquad (j \neq k) \end{aligned}$$

which can be expressed as:

$$\boxed{\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}} \qquad (27)$$

which defines an algebra.

- β is Hermitian so γ^0 is Hermitian.
- The α matrices are also Hermitian, giving

$$\gamma^{1\dagger} = (\beta\alpha_x)^\dagger = \alpha_x^\dagger\beta^\dagger = \alpha_x\beta = -\beta\alpha_x = -\gamma^1.$$

  hence γ^1, γ^2, γ^3 are anti-Hermitian:

$$\boxed{\gamma^{0\dagger} = \gamma^0, \quad \gamma^{1\dagger} = -\gamma^1, \quad \gamma^{2\dagger} = -\gamma^2, \quad \gamma^{3\dagger} = -\gamma^3}.$$

Pauli-Dirac Representation

- From now on we will use the Pauli-Dirac representation of the gamma matrices:

$$\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}; \quad \gamma^k = \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix}$$

which when written in full are

$$\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}; \quad \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix};$$

$$\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}; \quad \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

- Using the gamma matrices $\rho = \psi^\dagger \psi$ and $\vec{j} = \psi^\dagger \vec{\alpha} \psi$ can be written as $j^\mu = (\rho, \vec{j}) = \psi^\dagger \gamma^0 \gamma^\mu \psi$ where j^μ is the four-vector current. (The proof that j^μ is indeed a four vector concludes on page 151)
- In terms of the four-vector current the continuity equation becomes $\partial_\mu j^\mu = 0$
- Finally the expression for the four-vector current $j^\mu = \psi^\dagger \gamma^0 \gamma^\mu \psi$

can be simplified by introducing the adjoint spinor.

The Adjoint Spinor

The adjoint spinor is defined as follows:

$$\overline{\psi} = \psi^\dagger \gamma^0.$$

i.e. $\overline{\psi} = \psi^\dagger\gamma^0 = (\psi^*)^T\gamma^0 = (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*)\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}$. so

$$\overline{\psi} = (\psi_1^*, \psi_2^*, -\psi_3^*, -\psi_4^*)$$

In terms the adjoint spinor the four vector current can be written:

$$j^\mu = \overline{\psi}\gamma^\mu\psi$$

We will use this expression in deriving the Feynman rules for the Lorentz invariant matrix element for the fundamental interactions.

- That's enough notation, start to investigate the free particle solutions of the Dirac equation...

Dirac Equation: Plane Wave Solutions I

We are interested in plane wave solutions to the Dirac Equation (26) of the form:

$$\boxed{\psi = e^{\pm i(\vec{p}\cdot\vec{r} - Et)} u_{\pm}(E, \vec{p})} \tag{28}$$

where $u_{\pm}$ are appropriately chosen four-component 'spinors' and $E^2 = m^2 + \vec{p}^2$. [Aside: we will also name these spinors $(u_+, u_-) \leftrightarrow (u, v)$ in equations that feature just one but not the other.] Since the Dirac Equation (26) is:

$$(i\gamma^{\mu}\partial_{\mu} - m)\psi = 0$$

such solutions would need to satisfy: $(i\gamma^{\mu}(\mp ip_{\mu}) - m)e^{\pm i(\vec{p}\cdot\vec{r} - Et)} u_{\pm}(E, \vec{p}) = 0$ or

$$\boxed{(\gamma^{\mu}p_{\mu} \mp m)u_{\pm} = 0} \tag{29}$$

which is written by most sources as two separate expressions:

$$(\gamma^{\mu}p_{\mu} - m)u = 0 \tag{30}$$

$$(\gamma^{\mu}p_{\mu} + m)v = 0 \tag{31}$$

with (30) and (31) referred to as 'the momentum space representation of the Dirac Equation for particle spinors' and 'anti-particle spinors', respectively. [Aside: proper justification for those names will be provided later!]

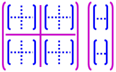
Note that (30) and (31) are algebraic rather than differential equations.

Dirac Equation: Plane Wave Solutions II

We will solve (29) to gain insight into the properties that u and v (i.e. that $u_\pm$) must have. Re-writing (29), i.e. $(\gamma^\mu p_\mu \mp m)u_\pm = 0$, in our representation of the Dirac Algebra gives us:

$$\left(E\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} - \vec{p}\cdot\begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix} \mp m\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}\right) u_\pm = 0 \tag{32}$$

or more succinctly:



$$\begin{pmatrix} E \mp m & -\vec{\sigma}\cdot\vec{p} \\ +\vec{\sigma}\cdot\vec{p} & -E \mp m \end{pmatrix} u_\pm = \begin{pmatrix} E \mp m & -\vec{\sigma}\cdot\vec{p} \\ +\vec{\sigma}\cdot\vec{p} & -E \mp m \end{pmatrix}\begin{pmatrix} A_\pm \\ B_\pm \end{pmatrix} = 0 \tag{33}$$

where we have broken the four-spinors $u_\pm$ into the their top two parts $A_\pm$ and their bottom two parts $B_\pm$. The expression (33) may be written as two simultaneous equations:

$$(E \mp m)A_\pm = (\vec{\sigma}\cdot\vec{p})B_\pm \tag{34}$$

$$(E \pm m)B_\pm = (\vec{\sigma}\cdot\vec{p})A_\pm. \tag{35}$$

These two equations turn out to be equivalent to each other (at least when $E^2 \neq m^2$) and so they amount to half as many constraints on $A_\pm$ and $B_\pm$ as one might naively imagine. To prove this statement, note first that $(\vec{\sigma}\cdot\vec{p})^2 = (E-m)(E+m)I$ since

$$(\vec{\sigma}\cdot\vec{p})^2 = \begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix}^2 = (p_x^2 + p_y^2 + p_z^2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = |\vec{p}|^2 I = (E^2 - m^2)I.$$

Dirac Equation: Plane Wave Solutions III

Then observe that when $E^2 \neq m^2$:

$$
\begin{aligned}
(34) \iff & [(E \mp m)A_\pm = (\vec{\sigma}\cdot\vec{p})B_\pm] \\
\iff & \left[(E \mp m)(\vec{\sigma}\cdot\vec{p})A_\pm = (\vec{\sigma}\cdot\vec{p})^2 B_\pm\right] \\
\iff & [(E \mp m)(\vec{\sigma}\cdot\vec{p})A_\pm = (E-m)(E+m)B_\pm] \\
\iff & [(\vec{\sigma}\cdot\vec{p})A_\pm = (E \pm m)B_\pm] \\
\iff & (35).
\end{aligned}
$$

[Aside: above we have required $E^2 \neq m^2$ (equivalently $\vec{p} \neq 0$) since it allows us to divide by $E \mp m$ which makes the maths easier to present. While equations (34) and (35) are not equivalent when $E^2 = m^2$, it may still be shown that the $\vec{p} \to 0$ limit of all our future result(s) is indeed the same as the result(s) one *would* have found by considering $E = \pm m$ as a special case at the outset. On other words: there is nothing special or magical about $\vec{p} = 0$; it the same as the limit $\vec{p} \to 0$. Thus we cheerfully assume that $E^2 \neq m^2$ wherever required hereafter, even though our results are also valid for $E^2 = m^2$. (That we can do this should follow from Einstein's equivalence principle!)]

Dirac Equation: Plane Wave Solutions IV

As $\boxed{(E \mp m)A_\pm = (\vec{\sigma} \cdot \vec{p})B_\pm \quad (34)}$ and $\boxed{(E \pm m)B_\pm = (\vec{\sigma} \cdot \vec{p})A_\pm \quad (35)}$

are equivalent (at least when $E^2 \neq m^2$) we may keep whichever one we like and discard the other, and then we may regard one of $A_\pm$ and $B_\pm$ free while the other is fixed by the retained equation. For reasons which may become clear later, we choose the following:

- when considering u_+ we use (35) to fix B_+ in terms of an unconstrained A_+; and
- when considering u_- we use (34) to fix A_- in terms of an unconstrained B_-, so that

$$u_+ = u \in \left\{ \left(\begin{matrix} A_+ \\ \frac{\vec{\sigma}\cdot\vec{p}}{E+m} A_+ \end{matrix} \right) \middle| \forall A_+ \right\} \text{ and } u_- = v \in \left\{ \left(\begin{matrix} \frac{\vec{\sigma}\cdot\vec{p}}{E+m} B_- \\ B_- \end{matrix} \right) \middle| \forall B_- \right\}$$

or, equivalently, we could say $u_+ = u \in \text{span}\{u_1, u_2\}$ and $u_- = v \in \text{span}\{v_1, v_2\}$ where:

$$u_1 = N \left(\begin{matrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \frac{\vec{\sigma}\cdot\vec{p}}{E+m} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{matrix} \right), \quad u_2 = N \left(\begin{matrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \frac{\vec{\sigma}\cdot\vec{p}}{E+m} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{matrix} \right), \quad v_1 = N \left(\begin{matrix} \frac{\vec{\sigma}\cdot\vec{p}}{E+m} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{matrix} \right), \quad v_2 = N \left(\begin{matrix} \frac{\vec{\sigma}\cdot\vec{p}}{E+m} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{matrix} \right)$$

where N may be freely chosen (see next slide).

Comments on the Dirac spinors u_1, u_2, v_1 and v_2.

1. It may be instructive to observe that the last definitions unpack to:

$$u_1 = N\begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{E+m} \\ \frac{p_x+ip_y}{E+m} \end{pmatrix}, \quad u_2 = N\begin{pmatrix} 0 \\ 1 \\ \frac{p_x-ip_y}{E+m} \\ \frac{-p_z}{E+m} \end{pmatrix}, \quad v_1 = N\begin{pmatrix} \frac{p_x-ip_y}{E+m} \\ \frac{-p_z}{E+m} \\ 0 \\ 1 \end{pmatrix}, \quad v_2 = N\begin{pmatrix} \frac{p_z}{E+m} \\ \frac{p_x+ip_y}{E+m} \\ 1 \\ 0 \end{pmatrix}$$

2. If one wishes to choose N so that $\psi^\dagger(u_1)\psi(u_1)$, $\psi^\dagger(u_2)\psi(u_2)$, $\psi^\dagger(v_1)\psi(v_1)$, and $\psi^\dagger(v_2)\psi(v_2)$ are all to equal $2E$, then a suitable choices is

$$N = \sqrt{E+m}.$$

This gives the '$2E$-particles-per-unit-volume' normalisation which we determined was needed in the previous handout. [Check this value of N is correct! You may find it easier to perform this check on the non-unpacked forms of the u and v spinors given on the previous slide.]

Does the Dirac Equation solve the negative energy states problem?

The answer to this question depends on the context.

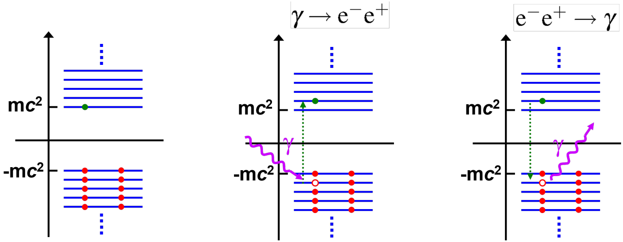
- Our definition of u and v via $u_\pm$ in (28) placed no explicit requirements on the sign of E. It demanded only that $E^2 = m^2 + \vec{p}^2$.
- Actual observed Dirac Particles clearly have positive energy (see e.g. evidence on page 88). This strongly motivates us taking $E = +\sqrt{m^2 + \vec{p}^2}$.
- Nonetheless, (28) shows that:

$$\hat{E}\ \psi(u_\pm) = \left(i\frac{\partial}{\partial t}\right)\left(e^{\pm i(\vec{p}\cdot\vec{r}-Et)}u_\pm(E,\vec{p})\right) = \pm E e^{\pm i(\vec{p}\cdot\vec{r}-Et)}u_\pm(E,\vec{p}) = \pm E\ \psi(u_\pm)$$

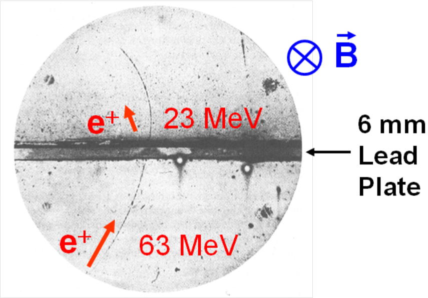
  implying that taking $E > 0$ would lead the 'usual' energy operator having positive eigenvalues on the $u = u_+$ spinors, but negative eigenvalues on the $v = u_-$ spinors. Dirac initially thought that half of his particles had negative energies. He developed various (now discredited) theories as workarounds (Dirac hole model, Dirac sea, *etc*).
- With the benefit of hindsight, all these were non-problems: the transition from Quantum Mechanics (1st quantisation) to Quantum Field Theories (2nd quantisation) gave a new interpretation to the $i\frac{\partial}{\partial t}$ operator and the energy operator in the contexts above. In Quantum Field Theory, all energies are positive, but the u-spinors have 'positive frequencies' and the v-spinors have 'negative frequencies' … see QFT course.
- As this course is QM rather than QFT based, we need to apply workaround to some of our operators. See page 93.

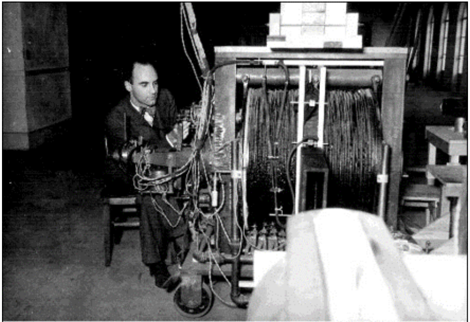
Discredited Dirac sea / hole model

- An attempt to explain why particles able to take negative energies would not fall down to ever lower energies radiating lots of energy in the process.
- Dirac Interpretation: the vacuum corresponds to all –ve energy states being full with the Pauli exclusion principle preventing electrons falling into -ve energy states. Holes in the –ve energy states correspond to +ve energy anti-particles with opposite charge. Provides a picture for pair-production and annihilation.



Discovery of Positron





- e+ enters at bottom, slows down in the lead plate – know direction
- Curvature in B-field shows that it is a positive particle
- Can't be a proton as would have stopped in the lead $\Longrightarrow$ Provided Verification of Predictions of Dirac Equation
- Anti-particle solutions exist ! But the picture of the vacuum corresponding to the state where all –ve energy states are occupied is rather unsatisfactory, what about bosons (no exclusion principle),. . . .

Chronology relating to Negative Energy Solutions

Not examinable

- 1928, Dirac invents his Equation. Probability density is positive, but negative energies are permitted (Proc. Roy. Soc. A117, 610-628) [1].
- 1930, Dirac tries to solve negative energies via the "hole" theory. He relates anti-particles to negative energy eigenstates. (Proc. Cam. Phil. Soc. 26, 376-381) [2].
- 1934, Paulu and Weisskopf present a new interpretation of Klein-Gordon equation: as field equation for a charged spin-0 field. ρ represents the charge density. The energy is given via

$$\frac{1}{2}\int d^3r\left[|\nabla\psi|^2 + m^2|\psi|^2\right]$$

  and thus positive by definition (Helv. Phys. Acta 7, 709-734) [3].
- 1934, The Dirac equation aquired a field-theoretic interpretation. It no longer represented a probability amplitude. Instead it became the field operator of a spin-$\frac{1}{2}$ field in a QFT. See the QFT and AQFT courses.

Not examinable

Charge Conjugation I

- In the part II Relativity and Electrodynamics course it was shown that the motion of a charged particle in an electromagnetic field $A^\mu = (\phi, \vec{A})$ can be obtained by making the minimal substitution $\vec{p} \to \vec{p} - e\vec{A}; \quad E \to E - e\phi$. With $\vec{p} = -i\vec{\nabla}$, $E = i\partial/\partial t$ and $A^\mu = (\phi, \vec{A})$ this can be written

$$\partial_\mu \to \partial_\mu + ieA_\mu$$

  and so under the above substitution the Dirac equation becomes:

$$\gamma^\mu(\partial_\mu + ieA_\mu)\psi + im\psi = 0 \tag{36}$$

- Now (for fun, and just because we can!) take the complex conjugate of the above and pre-multiplying by $-i\gamma^2$ to get this:

$$-i\gamma^2\gamma^{\mu*}(\partial_\mu - ieA_\mu)\psi^* - m\gamma^2\psi^* = 0. \tag{37}$$

  To simplify (37), note that in our representation of the gamma matrices we have:

$$\gamma^{0*} = \gamma^0;\ \gamma^{1*} = \gamma^1;\ \gamma^{2*} = -\gamma^2;\ \gamma^{3*} = \gamma^3$$

Charge Conjugation II

from which we can show that commuting a γ^2 past any gamma-matrix complex-conjugates-and-negates that gamma-matrix:

$$\gamma^2\gamma^{\mu *} = -\gamma^\mu\gamma^2.$$

We may therfore push the γ^2 past the $\gamma^{\mu\star}$ in (37) to get:

$$\gamma^\mu(\partial_\mu - ieA_\mu)i\gamma^2\psi^* + imi\gamma^2\psi^* = 0. \tag{38}$$

Since the expression $i\gamma^2\psi^\star$ features twice in the above, we may simplify things further by defining an operator $\hat{C}$ as follows

$$\boxed{\psi' = \hat{C}\psi = i\gamma^2\psi^*}$$

so that (38) becomes:

$$\gamma^\mu(\partial_\mu - ieA_\mu)\psi' + im\psi' = 0. \tag{39}$$

- Comparing (39) to the original equation (36):

$$\gamma^\mu(\partial_\mu + ieA_\mu)\psi + im\psi = 0$$

we see that the spinor ψ' describes a particle of the same mass but with opposite charge, i.e. an anti-particle ! $\hat{C}$: particle spinor $\Longleftrightarrow$ anti-particle spinor

Charge Conjugation III

- It cannot be stressed how revolutionary the result of the last slide was, when discpvered. That discovery shows that half of all spinor degrees of freedom relate to particles that 'go the other way' when experiencing an external field. In short, Dirac discovered that every fermion has an anti-fermionic partner.
- Because of the change in the sign of e in (36) compared to (39), we can henceforth name $\hat{C}$ the **Charge Conjugation Operator** for our representation of the gamma-matrices.
- For fun, consider the action of $\hat{C}$ on the free particle wave-function: $\psi = u_1 e^{i(\vec{p}.\vec{r}-Et)}$

  $\psi' = \hat{C}\psi = i\gamma^2\psi^* = i\gamma^2 u_1^* e^{-i(\vec{p}.\vec{r}-Et)}$

  $$i\gamma^2 u_1^* = i\begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}\sqrt{E+m}\begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{E+m} \\ \frac{p_x+ip_y}{E+m} \end{pmatrix}^* = \sqrt{E+m}\begin{pmatrix} \frac{p_x-ip_y}{E+m} \\ \frac{-p_z}{E+m} \\ 0 \\ 1 \end{pmatrix} = v_1 \text{ hence}$$

  $\psi = u_1 e^{i(\vec{p}.\vec{r}-Et)} \xrightarrow{\hat{C}} \psi' = v_1 e^{-i(\vec{p}.\vec{r}-Et)}$.

  Similarly $\psi = u_2 e^{i(\vec{p}.\vec{r}-Et)} \xrightarrow{\hat{C}} \psi' = v_2 e^{-i(\vec{p}.\vec{r}-Et)}$.
- Thus, **henceforth we may call v_1 the anti-partner of u_1 and we may call v_2 the anti-partner of u_2.**
- [Note to lecturer: tell class about general link between complex conjugation and anti-particles.]

Special operators for (some) anti-particle solutions

We noted on page 86 that if we were to apply the usual QM operators for energy and momentum

$$\hat{H} = i\partial/\partial t \qquad \text{and} \qquad \hat{p} = -i\vec{\nabla}$$

to particle or anti-particle solutions of the Dirac Equation of the form

$$\psi = u(E, \vec{p})e^{+i(\vec{p}.\vec{r}-Et)} \qquad \text{and} \qquad \psi = v(E, \vec{p})e^{-i(\vec{p}.\vec{r}-Et)}$$

then we would get eigenvalues for the anti-particles whose signs would be reversed from the physically/desired expected values. The signs for the anti-particles would be 'broken' by the $-$ in their expotent. This is not an issue if working in QFT from the start (different operators!)

- As the couse is in limbo, between experiment and theory: our workaround is use different operators (just for anti-particle states) to extract physical energies. I.e. define:

$$\hat{H}^{(v)} = -i\partial/\partial t \qquad \text{and} \qquad \hat{p}^{(v)} = i\vec{\nabla}.$$

- Need to do same for some other operators too. E.g.: under the transformation $(E, \vec{p}) \to (-E, -\vec{p})$ one would necessarily get $\vec{L} = \vec{r} \wedge \vec{p} \to -\vec{L}$. Conservation of total angular momentum is $[H, \vec{L} + \vec{S}] = 0$ so one needs to use $\hat{S}^{(\nu)} \to -\hat{S}$ to measure physical spin of anti-particles.

Summary of Solutions to the Dirac Equation

- The $2E$-per-unit-vol-normalised free PARTICLE solutions to the Dirac equation $\psi = u(E,\vec{p})e^{+i(\vec{p}.\vec{r}-Et)}$ satisfy $\boxed{(\gamma^\mu p_\mu - m)u = 0}$

$$\text{with } u_1 = \sqrt{E+m}\begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{E+m} \\ \frac{p_x+ip_y}{E+m} \end{pmatrix}; \qquad u_2 = \sqrt{E+m}\begin{pmatrix} 0 \\ 1 \\ \frac{p_x-ip_y}{E+m} \\ \frac{-p_z}{E+m} \end{pmatrix}$$

- The ANTI-PARTICLE solutions in terms of the physical energy and momentum: $\psi = v(E,\vec{p})e^{-i(\vec{p}.\vec{r}-Et)}$ satisfy $\boxed{(\gamma^\mu p_\mu + m)v = 0}$

$$\text{with } v_1 = \sqrt{E+m}\begin{pmatrix} \frac{p_x-ip_y}{E+m} \\ \frac{-p_z}{E+m} \\ 0 \\ 1 \end{pmatrix}; \qquad v_2 = \sqrt{E+m}\begin{pmatrix} \frac{p_z}{E+m} \\ \frac{p_x+ip_y}{E_m} \\ 1 \\ 0 \end{pmatrix}$$

- For the anti-particle states, operators whose eigenvalues are time-odd require reversed forms, e.g. $\hat{S}^{(v)} = -\hat{S}$.
- For both particle and anti-particle solutions $E = \sqrt{|\vec{p}|^2 + m^2}$.

(Now try question 7 – mainly about four-vector current.)

Why do Dirac particles have intrinsic angular momentum? (slide 1 of 3)

The Ehrenfest Theorem: $\frac{d}{dt}\langle A\rangle = \frac{i}{\hbar}\langle [H,A]\rangle + \left\langle \frac{\partial A}{\partial t}\right\rangle$.

- For a Dirac spinor is orbital angular momentum a good quantum number? i.e. does $L = \vec{r}\wedge\vec{p}$ commute with the Hamiltonian?

$$[H,\vec{L}] = [\vec{\alpha}.\vec{p}+\beta m, \vec{r}\wedge\vec{p}]$$
$$= [\vec{\alpha}.\vec{p}, \vec{r}\wedge\vec{p}]$$

Consider the x component of L:

$$[H,L_x] = [\vec{\alpha}.\vec{p}, (\vec{r}\wedge\vec{p})_x]$$
$$= [\alpha_x p_x + \alpha_y p_y + \alpha_z p_z, yp_z - zp_y]$$

The only non-zero contributions come from:

$$[H,L_x] = \alpha_y p_z[p_y, y] - \alpha_z p_y[p_z, z]$$
$$= -i(\alpha_y p_z - \alpha_z p_y)$$
$$= -i(\vec{\alpha}\wedge\vec{p})_x.$$

Consideration of other components shows that

$$[H,\vec{L}] = -i\vec{\alpha}\wedge\vec{p} \qquad (40)$$

- As this is not zero, the orbital angular momentum operator does not commute with the Hamiltonian, and so **orbital angular momentum is not a constant of motion** by Ehrenfest. **This should make you unhappy!**

Why do Dirac particles have intrinsic angular momentum? (slide 2 of 3)

Avoid depression by introducing a new mystery (perhaps even useless??) 4x4 operator:

$$\vec{S} = \frac{1}{2}\vec{\Sigma} = \frac{1}{2}\begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}$$

where $\vec{\sigma}$ are the Pauli spin matrices: i.e.

$$\Sigma_x = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}; \quad \Sigma_y = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}; \quad \Sigma_x = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Now consider the commutator

$$[H, \vec{\Sigma}] = [\vec{\alpha}.\vec{p}, \vec{\Sigma}]$$

here

$$[\beta, \vec{\Sigma}] = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} - \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} = 0$$

and hence

$$[H, \vec{\Sigma}] = [\vec{\alpha}.\vec{p}, \vec{\Sigma}]$$

Consider the x comp:

$$\begin{aligned} [H, \Sigma_x] &= [\alpha_x p_x + \alpha_y p_y + \alpha_z p_z, \Sigma_x] \\ &= p_x[\alpha_x, \Sigma_x] + p_y[\alpha_y, \Sigma_x] + p_z[\alpha_z, \Sigma_x] \end{aligned}$$

- Taking each of the commutators in turn:

$$
\begin{aligned}
[\alpha_x, \Sigma_x] &= \begin{pmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{pmatrix}\begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix} - \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix}\begin{pmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{pmatrix} = 0, \\
[\alpha_y, \Sigma_x] &= \begin{pmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{pmatrix}\begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix} - \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix}\begin{pmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & \sigma_y\sigma_x - \sigma_x\sigma_y \\ \sigma_y\sigma_x - \sigma_x\sigma_y & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & -2i\sigma_z \\ -2i\sigma_z & 0 \end{pmatrix} \\
&= -2i\alpha_z
\end{aligned}
$$

and similarly $[\alpha_z, \Sigma_x] = 2i\alpha_y$. Hence:

-

$$
\begin{aligned}
[H, S_x] &= \frac{1}{2}\left(p_x[\alpha_x, \Sigma_x] + p_y[\alpha_y, \Sigma_x] + p_z[\alpha_z, \Sigma_x]\right) \\
&= \frac{1}{2}\left(0 - 2ip_y\alpha_x + 2ip_z\alpha_y\right) = i(\vec{\alpha} \wedge \vec{p})_x
\end{aligned}
$$

and in general

$$
\left[H, \vec{S}\right] = i\vec{\alpha} \wedge \vec{p}. \tag{41}
$$

Why do Dirac particles have intrinsic angular momentum? (slide 3 of 3)

- Hence the observable corresponding to the operator $\vec{S}$ is also not a constant of motion. However, comparing (40) and (41) we see something quite amazing!

$$[H, \vec{L}] + [H, \vec{S}] = -i\vec{\alpha} \wedge \vec{p} + i\vec{\alpha} \wedge \vec{p} = 0 \tag{42}$$

Therefore the most amazing result in the whole course is that $L + S$ is conserved:

$$\boxed{\frac{d}{dt}(L + S) \propto [H, \vec{L} + \vec{S}] = 0}. \tag{43}$$

- In passing, but less excitingly, one might also note that because

$$\vec{S} = \frac{1}{2}\begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}$$

the commutation relationships for $\vec{S}$ are the same as for the $\vec{\sigma}$, e.g. $[S_x, S_y] = iS_z$. Furthermore both S^2 and S_z are diagonal

$$S^2 = \tfrac{1}{4}(\Sigma_x^2 + \Sigma_y^2 + \Sigma_z^2) = \tfrac{3}{4}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}; \quad S_z = \tfrac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

- Consequently $S^2\psi = S(S+1)\psi = \frac{3}{4}$ and for a particle travelling along the z-direction $S_z\psi = \pm\frac{1}{2}\psi$ or $\frac{\hbar}{2}$ in non-natural units and so dirac particles are spin-$\frac{1}{2}$.

Spin States

- In general the spinors u_1, u_2, v_1, v_2 are not Eigenstates of $\hat{S}_z$

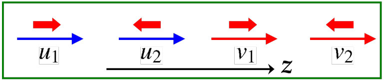
$$\hat{S}_z = \frac{1}{2}\Sigma_z = \frac{1}{2}\begin{pmatrix}\sigma_z & 0\\ 0 & \sigma_z\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}$$

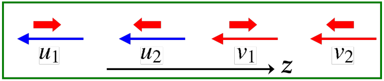
- However particles/anti-particles travelling in the z-direction: $p_z = \pm|\vec{p}|$
- $u_1 = N\begin{pmatrix}1\\ 0\\ \frac{\pm|\vec{p}|}{E+m}\\ 0\end{pmatrix}$; $u_2 = N\begin{pmatrix}0\\ 1\\ 0\\ \frac{\mp|\vec{p}|}{E+m}\end{pmatrix}$; $v_1 = N\begin{pmatrix}0\\ \frac{\mp|\vec{p}|}{E+m}\\ 0\\ 1\end{pmatrix}$; $v_2 = N\begin{pmatrix}\frac{\pm|\vec{p}|}{E+m}\\ 0\\ 1\\ 0\end{pmatrix}$

are Eigenstates of $\hat{S}_z$

$$\hat{S}_z u_2 = -\frac{1}{2}u_2, \hat{S}_z u_1 = +\frac{1}{2}u_1$$

$$\hat{S}_z^{(v)} v_1 = -\hat{S}_z v_1 = +\frac{1}{2}v_1, \hat{S}_z^{(v)} v_2 = -\hat{S}_z v_2 = -\frac{1}{2}v_2$$





Note the change of sign of $\hat{S}$ when dealing with antiparticle spinors

Spinors u_1 u_2 v_1 $v_2$are only eigenstates of $\hat{S}_z$ for $p_z = \pm|\vec{p}|$

Pause for Breath. . .

- We have found solutions to the Dirac equation which are also eigenstates $\hat{S}_z$ but only for particles travelling along the z axis. This is not a particularly useful basis!
- More generally, we want to label our states in terms of "good quantum numbers", i.e. a set of commuting observables.
- We can't use z component of spin since $[\hat{H}, \hat{S}_z] \neq 0$ as see in (42)
- We will introduce a new concept: "HELICITY". Helicity plays an important role in much that follows in the course...

Helicity

- The component of a particles spin along its direction of flight is a good quantum number (unless you overtake the particle) because $[H, \vec{S}.\vec{p}] = 0$. The same would be true even if $\vec{S}$ or $\vec{p}$ were scaled by (possibly different) constants.
- This motivates defining the helicity operator h by:

$$h \equiv \vec{\Sigma} \cdot \hat{\vec{p}} \equiv \vec{\Sigma} \cdot \left(\frac{\vec{p}}{|\vec{p}|} \right)$$

because with this definition:

1. the helicity of a particle will be independent of boosts along that particle's momentum direction (unless you overtake the particle!), and
2. the helicity values will be +1 or -1 for Dirac fermions. [We have already seen that Dirac fermions are spin-$\frac{1}{2}$ and that the S_z operator measures this if the particle is moving in the z-direction. Since $\Sigma = 2S$ the allowed values of helicity will be ± 1.]

- Conventionally, $h = +1$ and $h = -1$ are often referred to as follows:

- These are right and left handed HELICITY eigenstates. In Handout 4 we will discuss right and left handed CHIRAL eigenstates. Only in the limit $v \approx c$ are the HELICITY eigenstates the same as the CHIRAL eigenstates. **Do not confuse them!**

Helicity Eigenstates: $\{u_\uparrow, u_\downarrow, v_\uparrow, v_\downarrow\}$. I

- The basis spinors we already found $\{u_1, u_2, v_1, v_2\}$ were nice in that $\hat{C}(u_1, u_2) = (v_1, v_2)$. However, the difference between u_1 and u_2 was physically meaningless. It was set only by an arbitrary choice of two linearly independent vectors $A_+ = (1,0)^T$ or $(0,1)^T$ on page 84.
- We wish to remove that arbitrariness by finding a better spinor basis $\{u_\uparrow, u_\downarrow, v_\uparrow, v_\downarrow\}$ whose elements are all eigenstates of the helicity operator, e.g.:

$$(\vec{\Sigma}.\hat{p})u_\uparrow = +u_\uparrow \qquad \text{and}$$
$$(\vec{\Sigma}.\hat{p})u_\downarrow = -u_\downarrow.$$

- Since $\vec{\Sigma} \cdot \hat{p} = \begin{pmatrix} \vec{\sigma} \cdot \hat{\vec{p}} & 0 \\ 0 & \vec{\sigma} \cdot \hat{\vec{p}} \end{pmatrix}$ is proportional to $\begin{pmatrix} \vec{\sigma} \cdot \vec{p} & 0 \\ 0 & \vec{\sigma} \cdot \vec{p} \end{pmatrix}$ and since all of the 'chunked' spinor-halves on page 84 are already proportional to some power of $(\vec{\sigma} \cdot \vec{p})$ we can find the new basis we seek ($\{u_\uparrow, u_\downarrow, v_\uparrow, v_\downarrow\}$) by replacing the arbitrary choices for A_+ or B_- (namely $(1,0)^T$ or $(0,1)^T$) with the eigenvectors of the (2×2)-matrix $(\vec{\sigma} \cdot \vec{p})$.

Helicity Eigenstates: $\{u_\uparrow, u_\downarrow, v_\uparrow, v_\downarrow\}$. II

- If the momentum $\vec{p}$ points in the (θ,ϕ) direction in the usual spherical polar coordinates:

$$\vec{p} = |\vec{p}|\ (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$$

then

$$\begin{aligned}\vec{\sigma}.\hat{\vec{p}} &= \frac{1}{|\vec{p}|}\begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z\end{pmatrix} \\ &= \begin{pmatrix} \cos\theta & \sin\theta\cos\phi - i\sin\theta\sin\phi \\ \sin\theta\cos\phi + i\sin\theta\sin\phi & -\cos\theta\end{pmatrix} \\ &= \begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta\end{pmatrix}.\end{aligned}$$

- The eigenvectors of this last matrix (normalised such that $(\vec{e}_+)^\dagger\vec{e}_+ = (\vec{e}_-)^\dagger e_- = 1$) are:

$$\vec{e}_+ = \begin{pmatrix} \cos\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2}\end{pmatrix} \quad \text{and} \quad \vec{e}_- = \begin{pmatrix} -\sin\frac{\theta}{2} \\ e^{i\phi}\cos\frac{\theta}{2}\end{pmatrix}$$

which may be verified (test your memory of trig identities!) by checking that:

$$(\vec{\sigma}.\hat{\vec{p}})\ \vec{e}_+ = +\vec{e}_+ \quad \text{and} \quad (\vec{\sigma}.\hat{\vec{p}})\ \vec{e}_- = -\vec{e}_-.$$

Helicity Eigenstates: $\{u_\uparrow, u_\downarrow, v_\uparrow, v_\downarrow\}$. III

Accordingly, we augment/replace our previous definition from page 81 ...

$$u_1 = N\begin{pmatrix} \begin{pmatrix}1\\0\end{pmatrix} \\ \frac{\vec\sigma\cdot\vec p}{E+m}\begin{pmatrix}1\\0\end{pmatrix}\end{pmatrix}, \quad u_2 = N\begin{pmatrix} \begin{pmatrix}0\\1\end{pmatrix} \\ \frac{\vec\sigma\cdot\vec p}{E+m}\begin{pmatrix}0\\1\end{pmatrix}\end{pmatrix}, \quad v_1 = N\begin{pmatrix} \frac{\vec\sigma\cdot\vec p}{E+m}\begin{pmatrix}0\\1\end{pmatrix} \\ \begin{pmatrix}0\\1\end{pmatrix}\end{pmatrix}, \quad v_2 = N\begin{pmatrix} \frac{\vec\sigma\cdot\vec p}{E+m}\begin{pmatrix}1\\0\end{pmatrix} \\ \begin{pmatrix}1\\0\end{pmatrix}\end{pmatrix}$$

... with the following new definition (also relativistically normalised with $N = \sqrt{E+m}$):

$$u_\uparrow = N\begin{pmatrix} \vec e_+ \\ \frac{\vec\sigma\cdot\vec p}{E+m}\vec e_+\end{pmatrix}, \quad u_\downarrow = N\begin{pmatrix} \vec e_- \\ \frac{\vec\sigma\cdot\vec p}{E+m}\vec e_-\end{pmatrix}, \quad v_\uparrow = N\begin{pmatrix} \frac{\vec\sigma\cdot\vec p}{E+m}\vec e_- \\ \vec e_-\end{pmatrix}, \quad v_\downarrow = N\begin{pmatrix} \frac{\vec\sigma\cdot\vec p}{E+m}\vec e_+ \\ \vec e_+\end{pmatrix}$$

which looks as follows if everything (except the N prefactor, due to lack of space!) is written out in full:

$$u_\uparrow \propto \begin{pmatrix} \cos\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2} \\ \frac{|\vec p|}{E+m}\cos\frac{\theta}{2} \\ \frac{|\vec p|}{E+m}e^{i\phi}\sin\frac{\theta}{2}\end{pmatrix}, u_\downarrow \propto \begin{pmatrix} -\sin\frac{\theta}{2} \\ e^{i\phi}\cos\frac{\theta}{2} \\ \frac{|\vec p|}{E+m}\sin\frac{\theta}{2} \\ -\frac{|\vec p|}{E+m}e^{i\phi}\cos\frac{\theta}{2}\end{pmatrix}, v_\uparrow \propto \begin{pmatrix} \frac{|\vec p|}{E+m}\sin\frac{\theta}{2} \\ -\frac{|\vec p|}{E+m}e^{i\phi}\cos\frac{\theta}{2} \\ -\sin\frac{\theta}{2} \\ e^{i\phi}\cos\frac{\theta}{2}\end{pmatrix}, v_\downarrow \propto \begin{pmatrix} \frac{|\vec p|}{E+m}\cos\frac{\theta}{2} \\ \frac{|\vec p|}{E+m}e^{i\phi}\sin\frac{\theta}{2} \\ \cos\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2}\end{pmatrix}.$$

Helicity Eigenstates: $\{u_\uparrow, u_\downarrow, v_\uparrow, v_\downarrow\}$. IV

Aside: for the reasons already given on page 93, the desired (physical) helicity eigenvalues will only be found for the anti-particle helicity states $v_\uparrow$ and $v_\downarrow$ if they are evaluated using $\hat{h}^{(v)} = -\hat{h}$ rather than $\hat{h}$. In other words, the arrows only make sense as follows:

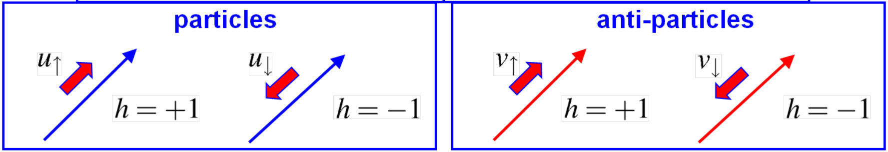
$$\hat{h} u_\uparrow = +u_\uparrow, \qquad \hat{h} u_\downarrow = -u_\downarrow \qquad \text{yet} \qquad \hat{h}^{(v)} v_\uparrow = +v_\uparrow, \qquad \hat{h}^{(v)} v_\downarrow = -v_\downarrow.$$

Summary of Helicity Eigenstates of the Dirac Equation:

- The particle and anti-particle helicity eigenstates states are:

$$u_{\uparrow} = N \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) \\ e^{i\phi}\sin\left(\frac{\theta}{2}\right) \\ \frac{|\vec{p}|}{E+m}\cos\left(\frac{\theta}{2}\right) \\ \frac{|\vec{p}|}{E+m}e^{i\phi}\sin\left(\frac{\theta}{2}\right) \end{pmatrix} \qquad u_{\downarrow} = N \begin{pmatrix} -\sin\left(\frac{\theta}{2}\right) \\ e^{i\phi}\cos\left(\frac{\theta}{2}\right) \\ \frac{|\vec{p}|}{E+m}\sin\left(\frac{\theta}{2}\right) \\ -\frac{|\vec{p}|}{E+m}e^{i\phi}\cos\left(\frac{\theta}{2}\right) \end{pmatrix}$$

$$v_{\uparrow} = N \begin{pmatrix} \frac{|\vec{p}|}{E+m}\sin\left(\frac{\theta}{2}\right) \\ -\frac{|\vec{p}|}{E+m}e^{i\phi}\cos\left(\frac{\theta}{2}\right) \\ -\sin\left(\frac{\theta}{2}\right) \\ e^{i\phi}\cos\left(\frac{\theta}{2}\right) \end{pmatrix} \qquad v_{\downarrow} = N \begin{pmatrix} \frac{|\vec{p}|}{E+m}\cos\left(\frac{\theta}{2}\right) \\ \frac{|\vec{p}|}{E+m}e^{i\phi}\sin\left(\frac{\theta}{2}\right) \\ \cos\left(\frac{\theta}{2}\right) \\ e^{i\phi}\sin\left(\frac{\theta}{2}\right) \end{pmatrix}$$



- For all four states, normalising to $2E$-per-unit-vol still needs $N = \sqrt{E+m}$. These helicity eigenstates will be used extensively in the calculations that follow.

Intrinsic Parity of Dirac Particles I

Before leaving the Dirac equation, consider parity. The parity operation is defined as spatial inversion through the origin:

$$x' \equiv -x; \quad y' \equiv -y; \quad z' \equiv -z; \quad t' \equiv t. \tag{44}$$

Consider a Dirac spinor, $\psi(x, y, z, t)$, which satisfies the Dirac equation

$$i\gamma^1 \frac{\partial \psi}{\partial x} + i\gamma^2 \frac{\partial \psi}{\partial y} + i\gamma^3 \frac{\partial \psi}{\partial z} - m\psi = -i\gamma^0 \frac{\partial \psi}{\partial t}. \tag{45}$$

Pre-multiplying (45) by γ^0 and making the unprimed to primed substitutions defined in (44) then results in:

$$\begin{aligned}
& -i\gamma^0\gamma^1 \frac{\partial \psi}{\partial x'} - i\gamma^0\gamma^2 \frac{\partial \psi}{\partial y'} - i\gamma^0\gamma^3 \frac{\partial \psi}{\partial z'} - m\gamma^0\psi = -i\gamma^0\gamma^0 \frac{\partial \psi}{\partial t'} \\
\Longrightarrow \quad & i\gamma^1 \frac{\partial \gamma^0\psi}{\partial x'} + i\gamma^2 \frac{\partial \gamma^0\psi}{\partial y'} + i\gamma^3 \frac{\partial \gamma^0\psi}{\partial z'} - m\gamma^0\psi = -i\gamma^0 \frac{\partial \gamma^0\psi}{\partial t'} \\
\Longrightarrow \quad & i\gamma^1 \frac{\partial \psi'}{\partial x'} + i\gamma^2 \frac{\partial \psi'}{\partial y'} + i\gamma^3 \frac{\partial \psi'}{\partial z'} - m\psi' = -i\gamma^0 \frac{\partial \psi'}{\partial t'} \qquad (46)
\end{aligned}$$

provided that in the last line we have introduced a new quantity $\psi'(x', y', z', t')$ defined by

$$\psi'(x', y', z', t') = \gamma^0\psi(x, y, z, t).$$

Intrinsic Parity of Dirac Particles II

Equations (45) and (46) are identical except that one is written with primed quantities and one without. Thus:

- We have found that the parity operator $\hat{P}$ which transforms spinors to their parity-conjugates must take the form: $\hat{P} = \lambda\gamma^0$ for some λ.
- If we further wish to have $(\hat{P})^2 = 1$ we are constrained to take $\lambda = \pm 1$. Either would work, but the most common convention in the literature is:

$$\boxed{\hat{P} = \gamma^0}.$$

- A basis for the spinors of stationary particles and anti-particles can be obtained from the $\vec{p} \to 0$ limit of the spinors on page 84 and yields:

$$u_1 = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \; u_2 = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \; v_1 = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \; v_2 = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}. \quad \text{Also:} \quad \gamma^0 = \begin{pmatrix}1 & 0 & 0 & 0\\0 & 1 & 0 & 0\\0 & 0 & -1 & 0\\0 & 0 & 0 & -1\end{pmatrix}$$

  and so in this convention fermions (u_1 and u_2) have positive parity and anti-fermions (v_1 and v_2) have negative parity when at rest.
- If we had used $\hat{P} = -\gamma^0$ then the parities for both would have reversed, but **it would still be the case that fermions and anti-fermions have opposite parity**. We will use this fact when we come to discuss mesons much later in the course.

Summary (1 or 2)

- The formulation of relativistic quantum mechanics starting from the linear Dirac equation

$$\hat{H}\psi = (\vec{\alpha}.\vec{p} + \beta m)\psi = i\frac{\partial\psi}{\partial t}$$

implied new degrees of freedom which were found to describe spin-$\frac{1}{2}$ particles and spin-$\frac{1}{2}$ anti-particles.

- In terms of (4×4) gamma-matrices the Dirac Equation was written:

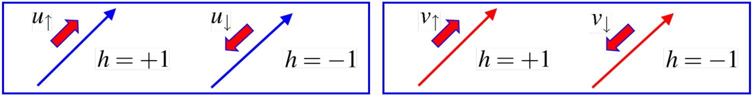
$$(i\gamma^{\mu}\partial_{\mu} - m)\psi = 0.$$

- We introduced a four-vector current and an adjoint spinor:

$$j^{\mu} = \psi^{\dagger}\gamma^{0}\gamma^{\mu}\psi = \overline{\psi}\gamma^{\mu}\psi$$

Summary (2 of 2)

- A useful helicity-ordered basis for particle and anti-particle spinors $\{u_\uparrow, u_\downarrow, v_\uparrow, v_\downarrow\}$ was summarised on page 106.



- In terms of 4-component spinors, the charge conjugation and parity operations were found (in our representation of the gamma-matrices) to be:

$$\psi \to \hat{C}\psi = i\gamma^2\psi^\dagger$$

$$\psi \to \hat{P}\psi = \gamma^0\psi$$

- Now that we have all we need to know about a relativistic description of the particles we scatter off each other, we can go on to discuss particle interactions and QED next!

Appendix III: Dimensions of the Dirac Matrices I

In a d-dimensional spacetime there will always be d gamma matrices, as one is associated with each spacetime derivative in the Hamiltonian. That is why in 4-dimensional spacetime we have four gamma matrices: γ_0, γ_1, γ_2 and γ_3.
But why does $d = 4$ force those matrices to be (4×4)-matrices ?
Rather than answer the above question, we instead state (and later prove) the more general result (47) linking the $(n \times n)$ size of gamma matrices to the number d of spacetime dimension with which they are associated:

$$n = 2^{\lfloor \frac{d}{2} \rfloor}. \tag{47}$$

The result (47) is a direct consequence of the gamma matrices having to satisfy (as we already saw in (27)) the defining property of a (so called) 'Clifford Algebra', namely that:

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu} 1_{n\times n}. \tag{48}$$

Warning: the proof we provide for the above statement relies on Schur's Lemma. This may be a source of dissatisfaction for some persons taking the course because **Schur's Lemma**, although stated in the Groups and Representations section of the Part IB Mathematics course within Natural Sciences Tripos, was stated in that coure without proof. If you find that annoying, you will have to find an alternative proof.

Appendix III: Dimensions of the Dirac Matrices II

Aside on size of Pauli matrices:

Although we are mainly interested in proving (47) to substantiate the claim that each γ^μ is a (4×4)-matrix, we note that the same result can be used to explain why the Pauli matrices are (2×2)-matrices. The reason is that the three $(d = 3)$ Pauli matrices satisfy their own equivalent of (48), namely: $\sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}$. Hence $n = 2^{\lfloor 3/2 \rfloor} = 2^1 = 2$.

We wish to prove the result stated in (47) is the relationship between the dimension d of spacetime and the dimension n of the (irreducible) $(n \times n)$ irreducible matrices γ_μ satisfying (48) with $\mu, \nu = 0, 1, \cdots, d-1$. Conveniently, the relationship (47) between n and d which we seek to prove does not depend on the signature of the metric since it is possible to convert a representation designed for one signature (say $g_{\mu\nu} = \mathrm{diag}(+,-,-,-)$) to another (say $g_{\mu\nu} = \mathrm{diag}(+,+,+,+)$) without changing n by multiplying appropriate γ-matrices by $i = \sqrt{-1}$.
Therefore, without loss of generality, we actually take as our start point the simplest possibility, namely:

$$\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2\delta_{\mu\nu} \cdot 1_{n\times n}. \tag{49}$$

We nonetheless demand that the γ-matrices are irreducible – i.e. that there is not a similarity transformation that would reduce them all to a (non-trivial) block diagonal form. We start by noting that with those assumptions:

Appendix III: Dimensions of the Dirac Matrices III

- **Every γ^μ is invertible**. [To prove this simply set $\mu = \nu$ in (49) and take the determinant of both sides.]
- **For the matrix $\gamma^* \equiv \gamma^0\gamma^1 \cdots \gamma^{d-1}$ we have**

$$\gamma^*\gamma^\mu = (-1)^{d-1}\gamma^\mu\gamma^*. \tag{50}$$

[Proof: When γ^μ commutes with γ^* it must pass $d-1$ dissimilar γ-matrices and a single 'identical' γ-matrix. Given (49) there are therefore $d-1$ anti-commutations and a single commutation. $\square$]

- **The matrix $\gamma^* \equiv \gamma^0\gamma^1 \cdots \gamma^{d-1}$ squares to either +1 or -1 depending on d.** [Proof: it takes $\frac{1}{2}(d-1)d$ flips of adjacent pairs to reverse the order of d objects, and since all the γ-matrices in γ^* are dissimilar and thus anti-commute we can deduce that

$$\gamma^* \equiv \gamma^0\gamma^1 \cdots \gamma^{d-1} = (-1)^{\frac{1}{2}(d-1)d} \cdot \gamma^{d-1} \cdots \gamma^1\gamma^0$$

Appendix III: Dimensions of the Dirac Matrices IV

and so

$$\begin{aligned}(\gamma^*)^2 &= (-1)^{\frac{1}{2}(d-1)d} \cdot (\gamma^{d-1} \cdots \gamma^1 \gamma^0) \cdot (\gamma^0 \gamma^1 \cdots \gamma^{d-1}) \\ &= (-1)^{\frac{1}{2}(d-1)d} \prod_{\mu=0}^{d-1} \delta^{\mu\mu} \\ &= (-1)^{\frac{1}{2}(d-1)d} \\ &= s(d) \end{aligned} \tag{51}$$

in which $s(d) \equiv (-1)^{\frac{1}{2}(d-1)d}$ is a d-dependent sign in $\{+1, -1\}$.]

- **If $d > 1$ then n must be even.** [To prove this, consider $\mu \neq \nu$ (which requires $d > 1$) in (49). In this case (49) becomes $\gamma^\mu \gamma^\nu = -\gamma^\mu \gamma^\mu$ which implies that $\det\{\gamma^\mu\} \det\{\gamma^\nu\} = (-1)^n \det\{\gamma^\nu\} \det\{\gamma^\mu\}$ which (since every γ^μ is invertible) implies that $1 = (-1)^n$ and thus that n is even.]

Appendix III: Dimensions of the Dirac Matrices V

- **Theorem A: Any product of any number of γ-matrices may (up to a sign) be written as a product of at most d gamma matrices in strictly ascending order of their indices.** [This is because (49) states that dissimilar γ-matrices anti-commute, and that individual γ-matrices square to ± 1'. Therefore, an arbitrary product of γ-matrices can always have its γ-matrices permuted into numerical order (with a sign change if an odd number of permutations is required) leaving at most one copy of each γ-matrix as repeats will disappear (up to a sign) on account of the squaring property.]

The last result above motivates the following definition.

Definition

If A is any integer whose binary representation modulo 2^d is $\vec{A}$, i.e. if (A mod 2^d) $= \sum_{i=0}^{d-1} A_i \cdot 2^i$ with each $A_i \in \{0,1\}$, then define Γ_A by

$$\Gamma_A = \prod_{i=0}^{d-1} \left\{ \begin{matrix} \gamma_i & \text{if } A_i = 1 \\ 1 & \text{otherwise} \end{matrix} \right\}. \tag{52}$$

For example, this definition would make $\Gamma_{13} = \gamma_0\gamma_2\gamma_3$ since $13 = 1 \cdot 2^0 + 0 \cdot 2^1 + 1 \cdot 2^2 + 1 \cdot 2^3$.

Appendix III: Dimensions of the Dirac Matrices VI

On account of the modulo 2^d part of the definition, any continuous range of indices of length 2^d would suffice to include every such Γ-matrix. Without loss of generality will always take indices A to be in the set

$$\mathcal{A} = \{1, 2, \cdots, 2^d\},$$

and mapped into that range, if necessary, by an implicit modulo 2^d operation. We therefore define a complete list, L, of Γ-matrices as follows:

$$L = (\Gamma_1, \Gamma_2, \ldots, \Gamma_{2^d}) = (\Gamma_A \mid A \in \mathcal{A}). \tag{53}$$

Note that although we have defined 2^d quantities Γ_A in the list L we have not shown that they are all unique. In other words, we cannot assume '$(A \neq B) \implies (\Gamma_A \neq \Gamma_B)$' or '$(\Gamma_A = \Gamma_B) \implies (A = B)$' unless later proved.

We now state and prove two important properties of the Γ-matrices:
The most general form of this Lemma is

$$\mathrm{Tr}\left[\Gamma_A\right] = \begin{cases} n & \text{if } A = 0 \mod 2^n \\ 0 & \text{if } (A \neq 0 \mod 2^n) \text{ and } (d \text{ is even or } \sum_{i=1}^{d} A_i \text{ is even}) \\ Tr[\Gamma_A] & \text{otherwise.} \end{cases} \tag{54}$$

Appendix III: Dimensions of the Dirac Matrices VII

Alternatively, a narrower form could be stated as follows

$$\text{When } d \text{ is even:} \qquad \mathrm{Tr}\left[\Gamma_A\right] = \begin{cases} n & \text{if } A = 0 \mod 2^n \\ 0 & \text{otherwise.} \end{cases} \tag{55}$$

The trace of Γ_0 is always trivially n as $\Gamma_0 = 1_{n\times n}$. Every other Γ_A is the product of one or more dissimilar γ-matrices. We split the remainder of the proof into two parts: part (i) shows that traces of products are zero where the remaining products contain an **even** number of γ-matrices, while part (ii) shows the same for products containing any **odd** number of γ-matrices. Note the subtle differences between these two parts of of the proof: the first needs to assume that the multiplied gammas are **distinct** but does not need to worry about whether d is even or odd. In contrast the second does not care about distinctness in the gammas but **needs to assume that d is even**. If k is an integer

Appendix III: Dimensions of the Dirac Matrices VIII

greater than zero, and if $a_1, a_2, \ldots, a_k$ are k **distinct** integers in $[0, d-1]$ and if $T = \mathrm{Tr}\left[\gamma_{a_1}\gamma_{a_2}\cdots\gamma_{a_{k-1}}\gamma_{a_k}\right]$ then

$$\begin{aligned} T &= \mathrm{Tr}\left[\gamma_{a_1}\gamma_{a_2}\cdots\gamma_{a_{k-1}}\gamma_{a_k}\right] \\ &= (-1)^{k-1}\cdot\mathrm{Tr}\left[\gamma_{a_k}\gamma_{a_1}\gamma_{a_2}\cdots\gamma_{a_{k-1}}\right] \\ &\qquad\qquad\text{(after } k-1 \text{ anti-commutations using (49) and } k>0\text{)} \\ &= (-1)^{k-1}\cdot\mathrm{Tr}\left[\gamma_{a_1}\gamma_{a_2}\cdots\gamma_{a_{k-1}}\gamma_{a_k}\right] \qquad \text{(trace cyclicity)} \\ &= (-1)^{k-1}\cdot T \end{aligned}$$

therefore:

$$\text{“The trace of the product of an \textbf{even} number of \textbf{distinct} } \gamma\text{-matrices} \ldots$$
$$\ldots\text{is zero provided the even number is greater than or equal to two”}. \qquad (56)$$

Appendix III: Dimensions of the Dirac Matrices IX

Not examinable

If k is an integer greater than zero, and if $a_1, a_2, \ldots, a_k$ are k integers in $[0, d-1]$ and if $T = \mathrm{Tr}\left[\gamma_{a_1}\gamma_{a_2}\cdots\gamma_{a_{k-1}}\gamma_{a_k}\right]$ then

$$
\begin{aligned}
& & T &= \mathrm{Tr}\left[\gamma_{a_1}\gamma_{a_2}\cdots\gamma_{a_{k-1}}\gamma_{a_k}\right] & \\
\Longrightarrow & & s(d)\cdot T &= \mathrm{Tr}\left[(\gamma^*\gamma^*)\gamma_{a_1}\gamma_{a_2}\cdots\gamma_{a_{k-1}}\gamma_{a_k}\right] & \text{(by (51))} \\
\Longrightarrow & & s(d)\cdot T &= \mathrm{Tr}\left[\gamma^*\gamma_{a_1}\gamma_{a_2}\cdots\gamma_{a_{k-1}}\gamma_{a_k}\gamma^*\right] & \text{(trace cyclicity)} \\
\Longrightarrow & & s(d)\cdot T &= ((-1)^{d-1})^k\cdot\mathrm{Tr}\left[\gamma^*\gamma^*\gamma_{a_k}\gamma_{a_1}\gamma_{a_2}\cdots\gamma_{a_{k-1}}\right] & \text{(after } k \text{ uses of (50))} \\
\Longrightarrow & & \cdot T &= (-1)^{k(d-1)}\cdot\mathrm{Tr}\left[\gamma_{a_k}\gamma_{a_1}\gamma_{a_2}\cdots\gamma_{a_{k-1}}\right] & \text{(by (51) again)} \\
\Longrightarrow & & \cdot T &= (-1)^{k(d-1)}\cdot T &
\end{aligned}
$$

therefore:

"when d is even, the trace of the product of an odd number of γ-matrices is zero". (57)

This conclude our proof of Lemma 1. □

$$\Gamma_A\Gamma_B = s(A,B)\cdot\Gamma_{A\oplus B} \tag{58}$$

Not examinable

Appendix III: Dimensions of the Dirac Matrices X

Not examinable

in which '$\oplus$' represents 'BITWISE EXCLUSIVE OR' and $s(A, B)$ is a function mapping pairs of indices to the set $\{+1, -1\}$.

$$\Gamma_A \Gamma_B = \prod_{i=0}^{d-1} \begin{Bmatrix} \gamma_i & \text{if } A_i = 1 \\ 1 & \text{otherwise} \end{Bmatrix} \prod_{i=0}^{d-1} \begin{Bmatrix} \gamma_i & \text{if } B_i = 1 \\ 1 & \text{otherwise} \end{Bmatrix}$$

$$= s_1(A, B) \prod_{i=0}^{d-1} \left(\begin{Bmatrix} \gamma_i & \text{if } A_i = 1 \\ 1 & \text{otherwise} \end{Bmatrix} \begin{Bmatrix} \gamma_i & \text{if } B_i = 1 \\ 1 & \text{otherwise} \end{Bmatrix} \right)$$

Not examinable

Appendix III: Dimensions of the Dirac Matrices XI

where $s_1(A,B) \in \{+1,-1\}$ is a sign which will depend on how many anti-commutations deriving from (49) were needed to re-order the matrices, and so

$$\Gamma_A \Gamma_B = s_1(A,B) \prod_{i=0}^{d-1} \left\{ \begin{matrix} (\gamma_i)^2 & \text{if } A_i = B_i = 1 \\ \gamma_i & \text{if } A_i \oplus B_i = 1 \\ 1 & \text{otherwise} \end{matrix} \right\}$$

$$= s_1(A,B) \prod_{i=0}^{d-1} \left\{ \begin{matrix} g_{ii} \quad (\text{no sum } i) & \text{if } A_i = B_i = 1 \\ \gamma_i & \text{if } A_i \oplus B_i = 1 \\ 1 & \text{otherwise} \end{matrix} \right\} \quad (\text{by (49)})$$

$$= s(A,B) \prod_{i=0}^{d-1} \left\{ \begin{matrix} 1 & \text{if } A_i = B_i = 1 \\ \gamma_i & \text{if } A_i \oplus B_i = 1 \\ 1 & \text{otherwise} \end{matrix} \right\}$$

Appendix III: Dimensions of the Dirac Matrices XII

where $s(A, B)$ is a new sign function that accounts for our having replaced g_{ii} with 1, and so

$$\begin{aligned}\Gamma_A\Gamma_B &= s(A,B)\prod_{i=0}^{d-1}\left\{\begin{matrix}\gamma_i & \text{if } A_i \oplus B_i = 1 \\ 1 & \text{otherwise}\end{matrix}\right\} \\ &= s(A,B)\Gamma_{A\oplus B} \qquad \square.\end{aligned}$$

A corollary of (58) is that every Γ-matrix is invertible. [Proof: setting B equal to A in (58) tells us that $(\Gamma_A)^2 = s(A,A)\cdot\Gamma_0 = s(A,A)\cdot 1_{n\times n} = \pm 1_{n\times n}$ and so

$$(\Gamma_A)^{-1} \text{ is either } \Gamma_A \text{ or } -\Gamma_A. \qquad (59)$$

]
Perhaps we can do better. Suppose A has a ones in its binary representation (i.e. $a = \sum_{i=0}^{d-1} A_i$ so that Γ_A is a product of a gamma matrices in ascending order of index). If we then square Γ_A we could attempt to permute adjacent gamma matrices within the product so as to annihilate every identical pairing, leaving behind only a sign. This process would require $a-1$ anticommutations to annihilate the first pair, $a-2$ the

Appendix III: Dimensions of the Dirac Matrices XIII

second, *etc*, and none for the last. This is a total of $\frac{1}{2}(a-1)a$ anticommutations, and so we can make the very specific claim that

$$(\Gamma_A)^2 = (-1)^{\frac{1}{2}(a-1)a} \tag{60}$$

or equivalently

$$(\Gamma_A)^{-1} = (-1)^{\frac{1}{2}(a-1)a} \cdot \Gamma_A. \tag{61}$$

Indeed, we see that the already derived result (51) could be viewed with hindsight as a simple corollary of (60).
Knowing that the Γ-matrices are all invertible we may define a matrix S as follows:

$$S = \sum_{X\in\mathcal{A}} (\Gamma_X)^{-1} \cdot Y \cdot \Gamma_X \tag{62}$$

Appendix III: Dimensions of the Dirac Matrices XIV

where Y is an arbitrary $(n \times n)$-matrix whose value we will fix later. It is follows that for any integer A (not summed) in the usual range $\mathcal{A}$:

$$\begin{aligned}
(\Gamma_A)^{-1} \cdot S \cdot \Gamma_A &= \sum_{X \in \mathcal{A}} (\Gamma_X \Gamma_A)^{-1} \cdot Y \cdot (\Gamma_X \Gamma_A) \\
&= \sum_{X \in \mathcal{A}} (s_X \Gamma_{A \oplus X})^{-1} \cdot Y \cdot (s_X \Gamma_{A \oplus X}) \qquad \text{(using (58))} \\
&= \sum_{X \in \mathcal{A}} (\Gamma_{A \oplus X})^{-1} \cdot Y \cdot (\Gamma_{A \oplus X}) \\
&= \sum_{X \in A \oplus \mathcal{A}} (\Gamma_X)^{-1} \cdot Y \cdot (\Gamma_X) \\
&= \sum_{X \in \mathcal{A}} (\Gamma_X)^{-1} \cdot Y \cdot (\Gamma_X) \qquad \text{(since } A \oplus \mathcal{A} \equiv \{A \oplus B, B \in \mathcal{A}\} = \mathcal{A}) \\
&= S
\end{aligned}$$

and thus $S \cdot \Gamma_A = \Gamma_A \cdot S$.
Having found a matrix S which commutes with every element Γ_A of a list L of matrices, one might hope to use Schur's Lemma to claim that S is some multiple of $1_{n \times n}$. However, a precondition of the only version of Schur's Lemma which I understand and which also

Appendix III: Dimensions of the Dirac Matrices XV

allows that conclusion to be drawn requires the elements of L to form an irreducible representation of some group G. Not only have we not yet shown that this precondition is satisfied, it actually looks likely to be false! For example, for the usual γ-matrices in $d = 4$ dimensions we would have $\Gamma_1\Gamma_2 = \gamma_1\gamma_2 = -\gamma_2\gamma_1 = -\Gamma_2\Gamma_1$ and so for L to be closed under multiplication it would need to contain both $+\Gamma_2\Gamma_1$ and $-\Gamma_2\Gamma_1$. This seems unlikely as we did not set up L to contain negated copies of every element. It therefore seems unlikely that L is closed under multiplication and so it seems unlikely that L represents a group.
It could be argued that the source of the problem is the annoying sign $s(A, B)$ in (58). If that pesky sign were not there and the constant '$+1$' were always in its place, products of Γ-matrices would be closed. We cannot arbitrarily dispose of that pesky sign, but it does suggest a resolution: we could double the length of our list L by adding to it another copy of itself but with the sign of every matrix reversed in the second half. The elements of this list will then be closed under multiplication, which is would be a requirement for them to be any kind of representation. We shall call the set containing all those elements G:

$$G = \{+\Gamma_A \mid A \in \mathcal{A}\} \cup \{-\Gamma_A \mid A \in \mathcal{A}\}. \tag{63}$$

This set of matrices is: (i) closed under multiplication, (ii) contains the identit $\Gamma_{2^d} = 1_{n\times n}$, (iii) contains an inverse for every element (see proof in (59)). Finally (iv) matrix multiplication is associative. Therefore G together with the operation of matrix multiplication forms a group. As it is a finite matrix group it is also representation of

Appendix III: Dimensions of the Dirac Matrices XVI

itself. This representation must be irreducible since the representation contains elements which are copies of the original γ-matrices (e.g. $\Gamma_1 = \gamma_0$, $\Gamma_2 = \gamma_1$, ... $\Gamma_{2^d} = \gamma_d$), and those original γ-matrices were taken to be be irreducible at the outset by assumption (see paragraph containing (49)). Although we have increased the number of elements in G relative to L, we can be sure that our old S will commute with every element of the new G because

$$([S, +\Gamma_A] = 0) \iff ([S, -\Gamma_A] = 0).$$

We have thus established all the preconditions necessary to allow us to use Schur's Lemma to state that S is a multiple of the identity, or more specifically:

$$\lambda \cdot 1_{n\times n} = \sum_{X\in\mathcal{A}} (\Gamma_A)^{-1} \cdot Y \cdot \Gamma_A \tag{64}$$

Appendix III: Dimensions of the Dirac Matrices XVII

for some scalar λ that will depend on Y. Taking the trace of both sides of (64) and using the cyclicity of the trace gives us:

$$\begin{aligned} n\lambda &= \sum_{X\in\mathcal{A}} \operatorname{Tr}\left[(\Gamma_A)^{-1}\cdot Y\cdot\Gamma_A\right] \\ &= \sum_{X\in\mathcal{A}} \operatorname{Tr}\left[Y\cdot\Gamma_A\cdot(\Gamma_A)^{-1}\right] \\ &= \sum_{X\in\mathcal{A}} \operatorname{Tr} Y \\ &= 2^d\cdot\operatorname{Tr} Y \end{aligned}$$

and thus

$$\lambda = \frac{2^d}{n}\cdot\operatorname{Tr} Y. \qquad (65)$$

Putting this value for λ back into (64) yields

$$\frac{2^d}{n}\cdot\operatorname{Tr} Y\cdot 1_{n\times n} = \sum_{X\in\mathcal{A}}(\Gamma_A)^{-1}\cdot Y\cdot\Gamma_A. \qquad (66)$$

Appendix III: Dimensions of the Dirac Matrices XVIII

Not examinable

We now exercise our remaining freedom to choose Y to be any $(n \times n)$-matrix we wish, deciding to let

$$[Y]_{ij} = \delta_{is}\delta_{jt}$$

where s and t are integers in $[1, n]$ which we may choose to fix later. With that choice in mind, and with i and j being other arbitrary integers also in $[1, n]$, (66) can be expanded as:

$$\left[\frac{2^d}{n} \cdot \mathrm{Tr}\, Y \cdot 1_{n\times n}\right]_{ij} = \left[\sum_{X\in\mathcal{A}} (\Gamma_A)^{-1} \cdot Y \cdot \Gamma_A\right]_{ij}$$

or equivalently

$$\frac{2^d}{n} \cdot (\delta_{ms}\delta_{mt}) \cdot \delta_{ij} = \sum_{X\in\mathcal{A}} ((\Gamma_A)^{-1})_{im} \cdot (\delta_{ms}\delta_{nt}) \cdot (\Gamma_A)_{nj}$$

which simplifies to

$$\frac{2^d}{n} \cdot \delta_{st} \cdot \delta_{ij} = \sum_{X\in\mathcal{A}} ((\Gamma_A)^{-1})_{is} \cdot (\Gamma_A)_{tj}. \tag{67}$$

Not examinable

Appendix III: Dimensions of the Dirac Matrices XIX

Since (67) is true for any i,j,s,t in $[1, n]$, let us set $s \to i$ and $t \to j$ and then sum over i and j. Making use of the summation convention over i and j we find that:

$$\frac{2^d}{n} \cdot \delta_{ij} \cdot \delta_{ij} = \sum_{A \in \mathcal{A}} ((\Gamma_A)^{-1})_{ii} \cdot (\Gamma_A)_{jj}$$

which simplifies to

$$\frac{2^d}{n} \cdot n = \sum_{A \in \mathcal{A}} \mathrm{Tr}\left[(\Gamma_A)^{-1}\right] \cdot \mathrm{Tr}[\Gamma_A]$$

or

$$2^d = \sum_{A \in \mathcal{A}} \mathrm{Tr}\left[(\Gamma_A)^{-1}\right] \cdot \mathrm{Tr}[\Gamma_A]. \qquad (68)$$

Appendix III: Dimensions of the Dirac Matrices XX

For the case that d is even we may now use (55) to simplify (68) to

$$\begin{aligned} 2^d &= \mathrm{Tr}\left[(\Gamma_0)^{-1}\right] \cdot \mathrm{Tr}[\Gamma_0] \\ &= \mathrm{Tr}\left[(1_{n\times n})^{-1}\right] \cdot \mathrm{Tr}[1_{n\times n}] \\ &= \mathrm{Tr}[1_{n\times n}] \cdot \mathrm{Tr}[1_{n\times n}] \\ &= n \cdot n = n^2 \\ \Longrightarrow \qquad n &= 2^{d/2} \qquad \text{(but only for } d \text{ even!).} \end{aligned} \tag{69}$$

This is a bit of a trick. One may always generate an irreducible representation of the gamma matrices for an **odd** spacetime dimension $d+1$ from an irreducible representation valid for an **even** number of spacetime dimensions d. The way to do this is surprisingly simple: if

$$\left\{\gamma^0, \gamma^1, \ldots, \gamma^{d-1}\right\}$$

is an irrep of (49) for an **even** number of spacetime dimensions d, and if we define

$$\gamma^* \equiv \gamma^0 \gamma^1 \cdots \gamma^{d-1}$$

Appendix III: Dimensions of the Dirac Matrices XXI

and if we recall the definition of $s(d)$ from (51), then

$$\left\{\gamma^0, \gamma^1, \ldots, \gamma^{d-1}\right\} \cup \left\{\sqrt{s(d)} \cdot \gamma^*\right\} \tag{70}$$

will be an irrep of (49) valid for dimension $d+1$ spacetime dimensions. That (70) is the irrep it is claimed to be is a consequence of three things: (i) γ^* was proved in (50) to anticommute with all the other gamma matrices when d is **even** and this anti-commutation is the property enforced/required by (49) whenever $\mu \neq \nu$, (ii) that $\sqrt{s(d)}\gamma^*$ squares to 1 was proved in (51), and this is the property enforced/required by (49) whenever $\mu = \nu$, and (iii) the representation (70) is an irrep as the first d gammas formed an irrep by themselves (i.e. as there was no transformation which could 'reduce' them, there cannot be an irrep that could 'reduce' both then and γ^*). It may be observed that this argument cannot be used to grow irreps without limit, since once an irrep for even d is grown to an irrep for odd d, the 'next' γ^* would fail to anticommute as desired. Nonetheless, the clear message is that the dimension of the gamma matrices for odd spacetime dimension d is always the same as the even dimension $d-1$, and so (69) now informs us that

$$n = 2^{(d-1)/2} \qquad \text{(but only when } d \text{ is odd!).} \tag{71}$$

Appendix III: Dimensions of the Dirac Matrices XXII

There result (69) for even d can be merged with the result (71) for odd d into a single expression valid for any d:

$$\begin{aligned} & n = \begin{cases} 2^{d/2} & \text{(when } d \text{ is even)} \\ 2^{(d-1)/2} & \text{(when } d \text{ is odd)} \end{cases} \\ \implies \quad & n = 2^{\lfloor d/2 \rfloor} \qquad \text{(for any } d\text{).} \end{aligned} \tag{72}$$

This concludes the proof of (47) which is also a proof of the lesser claim that Dirac Spinors have four components in the usual 4-dimensional spacetime.

Appendix IV: Magnetic Moment I

- In the part II Relativity and Electrodynamics course it was shown that the motion of a charged particle in an electromagnetic field $A^\mu = (\phi, \vec{A})$ can be obtained by making the minimal substitution $\vec{p} \to \vec{p} - q\vec{A}; \quad E \to E - q\phi$
- Applying this to the equations in (**??**)

$$\begin{aligned} (\vec{\sigma}.\vec{p} - q\vec{\sigma}.\vec{A})u_B &= (E - m - q\phi)u_A \\ (\vec{\sigma}.\vec{p} - q\vec{\sigma}.\vec{A})u_A &= (E + m - q\phi)u_B \end{aligned} \tag{73}$$

Multiplying (73) by $(E + m - q\phi)$

$$\begin{aligned} (\vec{\sigma}.\vec{p} - q\vec{\sigma}.\vec{A})u_B &= (E - m - q\phi)u_A \\ (\vec{\sigma}.\vec{p} - q\vec{\sigma}.\vec{A})u_A &= (E + m - q\phi)u_B \end{aligned} \tag{74}$$

where kinetic energy $T = E - m$

- In the non-relativistic limit $T \ll m$ (74) becomes

$$\begin{aligned} (\vec{\sigma}.\vec{p} - q\vec{\sigma}.\vec{A})(\vec{\sigma}.\vec{p} - q\vec{\sigma}.\vec{A})u_A &\approx 2m(T - q\phi)u_A \\ \left[(\vec{\sigma}.\vec{p})^2 - q(\vec{\sigma}.\vec{A})(\vec{\sigma}.\vec{p}) - q(\vec{\sigma}.\vec{p})(\vec{\sigma}.\vec{A}) + q^2(\vec{\sigma}.\vec{A})^2\right] u_A &\approx 2m(T - q\phi)u_A \end{aligned} \tag{75}$$

Appendix IV: Magnetic Moment II

- Now $\vec{\sigma}.\vec{A} = \begin{pmatrix} A_z & A_x - iA_y \\ A_x + iA_y & -A_z \end{pmatrix}$; $\vec{\sigma}.\vec{B} = \begin{pmatrix} B_z & B_x - iB_y \\ B_x + iB_y & -B_z \end{pmatrix}$; which leads to $(\vec{\sigma}.\vec{A})(\vec{\sigma}.\vec{B}) = \vec{A}.\vec{B} + i\vec{\sigma}.(\vec{A} \wedge \vec{B})$
  and $(\vec{\sigma}.\vec{A})^2 = |\vec{A}|^2$
- The operator on the LHS of (75):

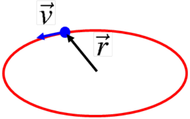
$$\begin{aligned}
&= \vec{p}^2 - q\left[\vec{A}.\vec{p} + i\vec{\sigma}.\vec{A} \wedge \vec{p} + \vec{p}.\vec{A} + i\vec{\sigma}.\vec{p} \wedge \vec{A}\right] + q^2\vec{A}^2 \\
&= (\vec{p} - q\vec{A})^2 - iq\vec{\sigma}.\left[\vec{A} \wedge \vec{p} + \vec{p} \wedge \vec{A}\right] \\
&= (\vec{p} - q\vec{A})^2 - q^2\vec{\sigma}.\left[\vec{A}.\vec{\nabla} + \vec{\nabla}.\vec{A}\right] \qquad (\text{since } \vec{p} = -i\vec{\nabla}) \\
&= (\vec{p} - q\vec{A})^2 - q\vec{\sigma}.(\vec{\nabla} \wedge \vec{A}) \qquad (\text{since } (\vec{\nabla} \wedge \vec{A})\psi = \vec{\nabla} \wedge (\vec{A}\psi) + \vec{A} \wedge (\vec{\nabla}\psi)) \\
&= (\vec{p} - q\vec{A})^2 - q\vec{\sigma}.\vec{B} \qquad (\text{since } \vec{B} = \vec{\nabla} \wedge \vec{A})
\end{aligned}$$

Substituting back into (75) gives the Schrödinger-Pauli equation for the motion of a non-relativisitic spin $\frac{1}{2}$ particle in an EM field:

$$\left[\frac{1}{2m}(\vec{p} - q\vec{A})^2 - \frac{q}{2m}\vec{\sigma}.\vec{B} + q\phi\right] u_A = Tu_A.$$

Appendix IV: Magnetic Moment III

$$\left[\frac{1}{2m}(\vec{p} - q\vec{A})^2 - \frac{q}{2m}\vec{\sigma}.\vec{B} + q\phi\right] u_A = T u_A$$



- Since the energy of a magnetic moment in a field is we can identify the intrinsic magnetic moment of a spin $\frac{1}{2}$ particle to be:

$$\vec{\mu} = \frac{q}{2m}\vec{\sigma}$$

In terms of the spin: $\vec{S} = \frac{1}{2}\vec{\sigma}$

$$\vec{\mu} = \frac{q}{m}\vec{S}$$

- Classically, for a charged particle current loop

$$\mu = \frac{q}{2m}\vec{L}$$

- The intrinsic magnetic moment of a spin half Dirac particle is twice that expected from classical physics. This is often expressed in terms of the gyromagnetic ratio is g=2.

$$\vec{\mu} = g\frac{q}{}\vec{S}$$

Appendix V: Generators of Lorentz Transformations I

It will shortly be seen that the quantities

$$(M^{\alpha\beta})^{\mu\nu} = g^{\mu\alpha} g^{\nu\beta} - g^{\nu\alpha} g^{\mu\beta} \tag{76}$$

or the equivalent (but less symmetric) quantities

$$(M^{\alpha\beta})^{\mu}{}_{\nu} = g^{\mu\alpha} \delta^{\beta}_{\nu} - \delta^{\alpha}_{\nu} g^{\mu\beta} \tag{77}$$

are generators of Lorentz Transformations. The indices $\alpha\beta$ choose between generators $M^{\alpha\beta}$, while ${}^{\mu}{}_{\nu}$ in $(M^{\alpha\beta})^{\mu}{}_{\nu}$ are there to act on vector indices. Evident antisymmetry in the $\alpha\beta$ of (76) means that there are only six independent non-zero generators. Suppressing

Appendix V: Generators of Lorentz Transformations II

the vector indices (taken to be ${}^{\mu}{}_{\nu}$) and taking $g^{\mu\nu} = \mathrm{diag}(+,-,-,-)$ the six independent generators are:

$$K_1 = M^{01} = -M^{10} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$K_2 = M^{02} = -M^{20} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$K_3 = M^{03} = -M^{30} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Appendix V: Generators of Lorentz Transformations III

and

$$J_1 = M^{23} = -M^{32} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & +1 & 0 \end{pmatrix}$$

$$J_2 = M^{31} = -M^{13} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & +1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

$$J_3 = M^{12} = -M^{21} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

or, for short:

$$J_i = \frac{1}{2}\epsilon_{ijk} M^{jk}$$
$$K_i = M^{0i}.$$

Appendix V: Generators of Lorentz Transformations IV

[Aside: The generators obey commutation relations

$$[J_i, J_j] = \epsilon_{ijk} J_k, \qquad [J_i, K_j] = \epsilon_{ijk} K_k, \qquad [K_i, K_j] = -\epsilon_{ijk} J_k.$$

The first of these says that the J's generate rotations in three-dimensional space and fixes the overall sign of the Js. The second says the Ks transform as a vector under rotations. End of aside]

With above definition[1] one could represent and arbitrary Lorentz transformation (boost, rotation or both) as

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu}$$

with

$$\Lambda^{\mu}{}_{\nu} = \left(\exp\left[\frac{1}{2} w_{\alpha\beta} (M^{\alpha\beta})^{\bullet}{}_{\bullet} \right] \right)^{\mu}{}_{\nu} \tag{78}$$

$$= \delta^{\mu}_{\nu} + \frac{1}{2} \omega_{\alpha\beta} (M^{\alpha\beta})^{\mu}{}_{\nu} + O(\omega^2) \tag{79}$$

using a set of parameters $w_{\alpha\beta}$ which may as well be antisymmetric in $\alpha\beta$ (since any symmetric part would not participate in (79) on account of the $(\alpha \leftrightarrow \beta)$-antisymmetry of $M^{\alpha\beta}$) and so contain six independent degrees of freedom (controlling three boosts and

Appendix V: Generators of Lorentz Transformations V

Not examinable

three rotations) as required. In most of the proofs which follow we use the infinitesimal transformations to first order in ω since if some properties can be proved for infinitesimal transformations then it is always be possible to generalise that result to the exponential form for a finite transformation.

Not examinable

[1] Compare to similar but slightly different sign/index conventions in http://www.phys.ufl.edu/~fry/6607/lorentz.pdf.

Appendix V Why do $(M^{\alpha\beta})^{\mu}{}_{\nu}$ generate Lorentz transformations? I

Not examinable

Lorentz transformations should be continuously connected to the identity (which (79) is, when $\omega_{\alpha\beta} = 0$) and should preserve inner products. The transformation in Eq. (79) preserves inner products because:

$$
\begin{aligned}
x' \cdot y' &= g_{\mu\nu} x'^{\mu} y'^{\nu} \\
&= g_{\mu\nu} (\Lambda^{\mu}{}_{\sigma} x^{\sigma})(\Lambda^{\nu}{}_{\tau} y^{\tau}) \\
&= g_{\mu\nu} (\delta^{\mu}_{\sigma} + \frac{1}{2}\omega_{\alpha\beta}(M^{\alpha\beta})^{\mu}{}_{\sigma})(\delta^{\nu}_{\tau} + \frac{1}{2}\omega_{\bar{\alpha}\bar{\beta}}(M^{\bar{\alpha}\bar{\beta}})^{\nu}{}_{\tau}) x^{\sigma} y^{\tau} + O(\omega)^2 \\
&= \left[g_{\sigma\tau} + \frac{1}{2}\left(\omega_{\alpha\beta}(M^{\alpha\beta})_{\tau\sigma} + \omega_{\bar{\alpha}\bar{\beta}}(M^{\bar{\alpha}\bar{\beta}})_{\sigma\tau} \right) \right] x^{\sigma} y^{\tau} + O(\omega^2) \\
&= \left[g_{\sigma\tau} + \frac{1}{2}\left(\omega_{\alpha\beta}(M^{\alpha\beta})_{\tau\sigma} + \omega_{\alpha\beta}(M^{\alpha\beta})_{\sigma\tau} \right) \right] x^{\sigma} y^{\tau} + O(\omega^2) \qquad \text{relabelling} \\
&= \left[g_{\sigma\tau} + \frac{1}{2}\left(\omega_{\alpha\beta}(M^{\alpha\beta})_{\tau\sigma} - \omega_{\alpha\beta}(M^{\alpha\beta})_{\tau\sigma} \right) \right] x^{\sigma} y^{\tau} + O(\omega^2) \qquad \text{antisymmetry of } M \\
&= g_{\sigma\tau} x^{\sigma} y^{\tau} + O(\omega^2) \\
&= x \cdot y + O(\omega^2).
\end{aligned}
$$

Not examinable

Appendix V Why do $(M^{\alpha\beta})^{\mu}{}_{\nu}$ generate Lorentz transformations? II

Not examinable

If the above argument seems too abstract, a more concrete way of checking that we have generators of Lorentz transformations might instead be to compute

$$\exp\{(\eta K_1)\} = \begin{pmatrix} \cosh\eta & \sinh\eta & 0 & 0 \\ \sinh\eta & \cosh\eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{80}$$

as this will be recognised by some as a boost in the positive x-direction with rapidity η (that is with $\cosh\eta = \gamma$ and $\sinh\eta = \beta\gamma$) while

$$\exp\{(\theta J_1)\} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta & -\sin\theta \\ 0 & 0 & \sin\theta & \cos\theta \end{pmatrix} \tag{81}$$

will be recognised by most as a rotation by an angle θ about the x-axis.

Not examinable

Appendix V: Lorentz covariance of the Dirac equation I

If the Dirac Equation:

$$i\gamma^{\mu}\partial_{\mu}\psi = m\psi \tag{82}$$

is to be Lorentz covariant, there would have to exist a matrix $S(\Lambda)$ such that $\psi' = S(\Lambda)\psi$ is the solution of the Lorentz transformed Dirac Equation

$$i\gamma^{\mu}\partial'_{\mu}\psi' = m\psi'. \tag{83}$$

Equation (83) implies

$$i\gamma_{\mu}\partial'^{\mu}\psi' = m\psi' \tag{84}$$

and so

$$i\gamma_{\mu}\Lambda^{\mu}{}_{\nu}\partial^{\nu}S(\Lambda)\psi = mS(\Lambda)\psi \tag{85}$$

and so since $S(\Lambda)$ is independent of position

$$i\gamma_{\mu}S(\Lambda)\Lambda^{\mu}{}_{\nu}\partial^{\nu}\psi = S(\Lambda)m\psi \tag{86}$$

which using (82) becomes

$$i\gamma_{\mu}S(\Lambda)\Lambda^{\mu}{}_{\nu}\partial^{\nu}\psi = S(\Lambda)i\gamma^{\mu}\partial_{\mu}\psi$$

Appendix V: Lorentz covariance of the Dirac equation II

and hence

$$i\gamma^{\mu} S(\Lambda)\Lambda_{\mu}{}^{\nu}\partial_{\nu}\psi = S(\Lambda)i\gamma^{\nu}\partial_{\nu}\psi$$

or

$$i\left[\gamma^{\mu} S(\Lambda)\Lambda_{\mu}{}^{\nu} - S(\Lambda)\gamma^{\nu}\right]\partial_{\nu}\psi = 0. \qquad (87)$$

Therefore, if we can show that there exists a matrix $S(\Lambda)$ satisfying

$$\gamma^{\mu} S(\Lambda)\Lambda_{\mu}{}^{\nu} = S(\Lambda)\gamma^{\nu} \qquad (88)$$

we will have found a solution to (87) and thus will have found that the Dirac Equation is Lorentz covariant as desired. Thought it would be entirely possible to work directly with (88) it is perhaps nicer to bring both S matrices to the left hand side

$$S^{-1}(\Lambda)\gamma^{\mu} S(\Lambda)\Lambda_{\mu}{}^{\nu} = \gamma^{\nu}$$

and then use the identity

$$\Lambda_{\mu}{}^{\nu}\Lambda^{\sigma}{}_{\nu} \equiv \delta^{\sigma}_{\mu} \qquad (89)$$

Appendix V: Lorentz covariance of the Dirac equation III

so that (88) ends up being written in the more common and (perhaps) more suggestive and useful form:

$$S^{-1}(\Lambda)\gamma^{\sigma}S(\Lambda) = \Lambda^{\sigma}{}_{\nu}\gamma^{\nu}. \qquad (90)$$

[Aside: Here is (for infinitesimal Lorentz transformations) a proof of the identity (89):

$$\begin{aligned}
\Lambda_{\mu}{}^{\nu}\Lambda^{\sigma}{}_{\nu} &= \left(g_{\mu}{}^{\nu} + \frac{1}{2}\omega_{\alpha\beta}(M^{\alpha\beta})_{\mu}{}^{\nu}\right)\left(g^{\sigma}{}_{\nu} + \frac{1}{2}\omega_{\bar\alpha\bar\beta}(M^{\bar\alpha\bar\beta})^{\sigma}{}_{\nu}\right) + O(\omega^2) \\
&= \delta^{\sigma}_{\mu} + \frac{1}{2}\left[\omega_{\alpha\beta}(M^{\alpha\beta})_{\mu}{}^{\sigma} + \omega_{\bar\alpha\bar\beta}(M^{\bar\alpha\bar\beta})^{\sigma}{}_{\mu}\right] + O(\omega^2) \\
&= \delta^{\sigma}_{\mu} + \frac{1}{2}\left[\omega_{\alpha\beta}(M^{\alpha\beta})_{\mu}{}^{\sigma} + \omega_{\alpha\beta}(M^{\alpha\beta})^{\sigma}{}_{\mu}\right] + O(\omega^2) \qquad \text{(relabelling)} \\
&= \delta^{\sigma}_{\mu} + \frac{1}{2}\omega_{\alpha\beta}\left[(M^{\alpha\beta})_{\mu}{}^{\sigma} + (M^{\alpha\beta})^{\sigma}{}_{\mu}\right] + O(\omega^2) \qquad \text{(factorising)} \\
&= \delta^{\sigma}_{\mu} + \frac{1}{2}\omega_{\alpha\beta}\left[(M^{\alpha\beta})^{\tau\sigma} + (M^{\alpha\beta})^{\sigma\tau}\right]g_{\mu\tau} + O(\omega^2) \qquad \text{(tidying)} \\
&= \delta^{\sigma}_{\mu} + \frac{1}{2}\omega_{\alpha\beta}\left[(M^{\alpha\beta})^{\tau\sigma} - (M^{\alpha\beta})^{\tau\sigma}\right]g_{\mu\tau} + O(\omega^2) \qquad \text{(antisymmetry of } M\text{)} \\
&= \delta^{\sigma}_{\mu} + O(\omega^2).
\end{aligned}$$

Appendix V: Lorentz covariance of the Dirac equation IV

End of aside]

Lemma

A valid choice of $S(\Lambda)$ (for an infinitesimal Lorentz transformation) is given by:

$$S(\Lambda) = 1 + \frac{1}{4}\omega_{\alpha\beta}\gamma^{\alpha}\gamma^{\beta} + O(\omega^2). \tag{91}$$

Appendix V: Lorentz covariance of the Dirac equation V

Proof.

$$
\begin{aligned}
S^{-1}(\Lambda)\gamma^\sigma S(\Lambda) &= \left(1 - \frac{1}{4}\omega_{\alpha\beta}\gamma^\alpha\gamma^\beta\right)\gamma^\sigma\left(1 + \frac{1}{4}\omega_{\bar\alpha\bar\beta}\gamma^{\bar\alpha}\gamma^{\bar\beta}\right) + O(\omega^2) \\
&= \gamma^\sigma + \frac{1}{4}\left(\omega_{\bar\alpha\bar\beta}\gamma^\sigma\gamma^{\bar\alpha}\gamma^{\bar\beta} - \omega_{\alpha\beta}\gamma^\alpha\gamma^\beta\gamma^\sigma\right) + O(\omega^2) \\
&= \gamma^\sigma + \frac{1}{4}\omega_{\alpha\beta}\left(\gamma^\sigma\gamma^\alpha\gamma^\beta - \gamma^\alpha\gamma^\beta\gamma^\sigma\right) + O(\omega^2) \\
&= \gamma^\sigma + \frac{1}{4}\omega_{\alpha\beta}\left((\gamma^\sigma\gamma^\alpha + \gamma^\alpha\gamma^\sigma)\gamma^\beta - \gamma^\alpha(\gamma^\sigma\gamma^\beta + \gamma^\beta\gamma^\sigma)\right) + O(\omega^2) \\
&= \gamma^\sigma + \frac{1}{4}\omega_{\alpha\beta}\left(2g^{\sigma\alpha}\gamma^\beta - \gamma^\alpha 2g^{\sigma\beta}\right) + O(\omega^2) \qquad \text{since } \{\gamma^\mu,\gamma^\nu\} \equiv 2g^{\mu\nu} \\
&= \left(\delta^\sigma_\nu + \frac{1}{2}\omega_{\alpha\beta}\left(g^{\sigma\alpha}\delta^\beta_\nu - \delta^\alpha_\nu g^{\sigma\beta}\right)\right)\gamma^\nu + O(\omega^2) \\
&= \left(\delta^\sigma_\nu + \frac{1}{2}\omega_{\alpha\beta}(M^{\alpha\beta})^\sigma{}_\nu\right)\gamma^\nu + O(\omega^2) \qquad \text{using (77)} \\
&= \Lambda^\sigma{}_\nu\gamma^\nu + O(\omega^2) \qquad \text{using (79).}
\end{aligned}
$$

□

Appendix V: Lorentz covariance of the Dirac equation VI

[Aside: Since $\gamma^\alpha\gamma^\beta = \frac{1}{2}\{\gamma^\alpha,\gamma^\beta\} + \frac{1}{2}[\gamma^\alpha,\gamma^\beta]$ we can also rewrite (91) in the more frequently seen (conventional) form:

$$S(\Lambda) = 1 + \frac{1}{8}\omega_{\alpha\beta}[\gamma^\alpha,\gamma^\beta] + O(\omega^2). \qquad (92)$$

End of aside]

Appendix V: Transformation properties of $\overline{\phi}\psi$, $\overline{\phi}\gamma^\mu\psi$ and $\overline{\phi}\gamma^\mu\gamma^\nu\psi$. I

Each of the expressions $\overline{\phi}\psi$, $\overline{\phi}\gamma^\mu\psi$ and $\overline{\phi}\gamma^\mu\gamma^\nu\psi$ is of the form $\overline{\phi}\gamma^\mu\gamma^\nu\cdots\gamma^\tau\psi$. To understand how any of them is affected by a Lorentz transformation it is therefore interesting to consider the following set of manipulations:[2]

$$\begin{aligned}
\overline{\phi'}\gamma^\mu\gamma^\nu\cdots\gamma^\tau\psi' &= \overline{(S(\Lambda)\phi)}[\gamma^\mu\gamma^\nu\cdots\gamma^\tau](S(\Lambda)\psi) \\
&= \phi^\dagger S^\dagger(\Lambda)\gamma^0[\gamma^\mu S(\Lambda)S^{-1}(\Lambda)\gamma^\nu S(\Lambda)\cdots S^{-1}(\Lambda)\gamma^\tau]S(\Lambda)\psi \\
&= \phi^\dagger S^\dagger(\Lambda)\gamma^0 S(\Lambda)(S^{-1}(\Lambda)\gamma^\mu S(\Lambda))(S^{-1}(\Lambda)\gamma^\nu S(\Lambda))\cdots(S^{-1}(\Lambda)\gamma^\tau S(\Lambda))\psi \\
&= \phi^\dagger S^\dagger(\Lambda)\gamma^0 S(\Lambda)(\Lambda^\mu{}_\alpha\gamma^\alpha)(\Lambda^\nu{}_\beta\gamma^\gamma)\cdots(\Lambda^\tau{}_\lambda\gamma^\lambda)\psi \qquad \text{using (90)}
\end{aligned}$$

which itself suggests that if we can show that

$$S^\dagger(\Lambda)\gamma^0 S(\Lambda) = \gamma^0 \tag{93}$$

then we will have proved that

$$\overline{\phi'}\gamma^\mu\gamma^\nu\cdots\gamma^\tau\psi' = \overline{\phi}(\Lambda^\mu{}_\alpha\gamma^\alpha)(\Lambda^\nu{}_\beta\gamma^\gamma)\cdots(\Lambda^\tau{}_\lambda\gamma^\lambda)\psi$$

which will itself have showed that each of the expressions under consideration transforms like a tensor of the appropriate rank.

Appendix V: Transformation properties of $\overline{\phi}\psi$, $\overline{\phi}\gamma^\mu\psi$ and $\overline{\phi}\gamma^\mu\gamma^\nu\psi$. II

Not examinable

We must therefore prove (93). To do so is a two-stage process. First we compute $S^\dagger(\Lambda)$. Then we combine it with $\gamma^0 S(\Lambda)$. Starting with (91):

$$\begin{aligned}
S^\dagger(\Lambda) &= \left[1 + \frac{1}{4}\omega_{\alpha\beta}\gamma^\alpha\gamma^\beta\right]^\dagger + O(\omega^2) \\
&= 1 + \frac{1}{4}\omega_{\alpha\beta}(\gamma^\alpha\gamma^\beta)^\dagger + O(\omega^2) \qquad (\omega_{\alpha\beta} \text{ are real}) \\
&= 1 + \frac{1}{4}\omega_{\alpha\beta}(\gamma^\beta)^\dagger(\gamma^\alpha)^\dagger + O(\omega^2) \\
&= 1 + \frac{1}{4}\omega_{\alpha\beta}(\gamma^0\gamma^\beta\gamma^0)(\gamma^0\gamma^\alpha\gamma^0) + O(\omega^2) \\
&= 1 + \frac{1}{4}\omega_{\alpha\beta}\gamma^0\gamma^\beta\gamma^\alpha\gamma^0 + O(\omega^2) \qquad (94)
\end{aligned}$$

Not examinable

Appendix V: Transformation properties of $\overline{\phi}\psi$, $\overline{\phi}\gamma^\mu\psi$ and $\overline{\phi}\gamma^\mu\gamma^\nu\psi$. III

Not examinable

from which we can deduce (using (91)) that

$$\begin{aligned}
S^\dagger(\Lambda)\gamma^0 S(\Lambda) &= \left(1+\frac{1}{4}\omega_{\alpha\beta}\gamma^0\gamma^\beta\gamma^\alpha\gamma^0\right)\gamma^0\left(1+\frac{1}{4}\omega_{\bar\alpha\bar\beta}\gamma^{\bar\alpha}\gamma^{\bar\beta}\right)+O(\omega^2)\\
&= \gamma^0+\frac{1}{4}\left(\omega_{\alpha\beta}\gamma^0\gamma^\beta\gamma^\alpha\gamma^0\gamma^0+\omega_{\bar\alpha\bar\beta}\gamma^0\gamma^{\bar\alpha}\gamma^{\bar\beta}\right)+O(\omega^2)\\
&= \gamma^0\left[1+\frac{1}{4}\left(\omega_{\alpha\beta}\gamma^\beta\gamma^\alpha+\omega_{\beta\alpha}\gamma^\beta\gamma^\alpha\right)\right]+O(\omega^2) \qquad ((\bar\alpha,\bar\beta)\to(\beta,\alpha))\\
&= \gamma^0\left[1+0\right]\psi+O(\omega^2) \qquad (\omega_{\alpha\beta}=-\omega_{\beta\alpha})\\
&= \gamma^0+O(\omega^2)
\end{aligned}$$

verifying (93) as required. This completes our proof that:

- $\overline{\phi}\psi$ is Lorentz invariant scalar,
- $\overline{\phi}\gamma^\mu\psi$ transforms as a Lorentz vector, and
- $\overline{\phi}\gamma^\mu\gamma^\nu\psi$ transforms as a second-rank tensor, *etc.*

Not examinable

[2] These manipulations may look complex but they really only consist of inserting lots of 'ones' in form '$S(\Lambda)S^{-1}(\Lambda)$' at the right places, using $\overline{\phi}\equiv\phi^\dagger\gamma^0$ and using (90) many times.

References I

[1] Paul Adrien Maurice Dirac.
The quantum theory of the electron.
Proc. Roy. Soc. A, 117:610–628, 2 1928.

[2] P. A. M. Dirac.
Note on exchange phenomena in the thomas atom.
Mathematical Proceedings of the Cambridge Philosophical Society, 26(3):376–385, 1930.

[3] W. Pauli and V. F. Weisskopf.
On Quantization of the Scalar Relativistic Wave Equation. (In German).
Helv. Phys. Acta, 7:709–731, 1934.