Kern- und Teilchenphysik I
Lecture 7: Dirac Equation

(adapted from the Handout of Prof. Mark Thomson)

Prof. Nico Serra
Dr. Patrick Owen, Mr. Davide Lancierini

http://www.physik.uzh.ch/de/lehre/PHY211/HS2017.html
Non relativistic QM

- 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:
  \[ \vec{p} \rightarrow -i\vec{\nabla}, \quad E \rightarrow i\frac{\partial}{\partial t} \]
which gives the time dependent Schrödinger equation (take V=0 for simplicity)

\[-\frac{1}{2m} \nabla^2 \psi = i\frac{\partial \psi}{\partial t} \quad \text{(S1)}\]

with plane wave solutions:
\[ \psi = Ne^{i(\vec{p}.\vec{r} - Et)} \]

- 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

\[ -\frac{1}{2m} \nabla^2 \psi^* = -i\frac{\partial \psi^*}{\partial t} \quad \text{(S2)} \]
\[ \psi^* \times (S1) - \psi \times (S2) : \quad -\frac{1}{2m} (\psi^* \nabla^2 \psi - \psi \nabla^2 \psi^*) = i \left( \psi^* \frac{\partial \psi}{\partial t} + \psi \frac{\partial \psi^*}{\partial t} \right) \]

\[ -\frac{1}{2m} \nabla \cdot (\psi^* \nabla \psi - \psi \nabla \psi^*) = i \frac{\partial}{\partial t} (\psi^* \psi) \]

- Which by comparison with the continuity equation

\[ \nabla \cdot j + \frac{\partial \rho}{\partial t} = 0 \]

leads to the following expressions for probability density and current:

\[ \rho = \psi^* \psi = |\psi|^2 \quad \text{and} \quad j = \frac{1}{2mi} \left( \psi^* \nabla \psi - \psi \nabla \psi^* \right) \]

- For a plane wave

\[ \psi = Ne^{i(p \cdot r - Et)} \]

\[ \rho = |N|^2 \quad \text{and} \quad j = |N|^2 \frac{\vec{p}}{m} = |N|^2 \vec{v} \]

- 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 \(j\) is a vector in the particle’s direction with magnitude equal to the flux.
Applying \( \vec{p} \rightarrow -i\vec{\nabla}, \ E \rightarrow i\partial / \partial t \) to the relativistic equation for energy:

\[
E^2 = |\vec{p}|^2 + m^2
\]

(KG1)

gives the Klein-Gordon equation:

\[
\frac{\partial^2 \psi}{\partial t^2} = \vec{\nabla}^2 \psi - m^2 \psi
\]

(KG2)

Using \( \partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \left( \frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \)

\[
\partial_\mu \partial_\nu = \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}
\]

KG can be expressed compactly as

\[
(\partial_\mu \partial_\nu + m^2) \psi = 0
\]

(KG3)

For plane wave solutions, \( \psi = Ne^{i(\vec{p} \cdot \vec{r} - Et)} \) the KG equation gives:

\[
-E^2 \psi = -|\vec{p}|^2 \psi - m^2 \psi
\]

\[
\Rightarrow \quad E = \pm \sqrt{|\vec{p}|^2 + m^2}
\]

Not surprisingly, the KG equation has negative energy solutions – this is just what we started with in eq. KG1.

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:

\[
\psi^* \frac{\partial^2 \psi^*}{\partial t^2} = \nabla^2 \psi^* - m^2 \psi^* \tag{KG4}\]

\[
\psi^* \times \text{(KG2)} - \psi \times \text{(KG4)} : \]

\[
\psi^* \frac{\partial^2 \psi}{\partial t^2} - \psi \frac{\partial^2 \psi^*}{\partial t^2} = \psi^*(\nabla^2 \psi - m^2 \psi) - \psi(\nabla^2 \psi^* - m^2 \psi^*)
\]

\[
\frac{\partial}{\partial t} \left( \psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t} \right) = \nabla \cdot (\psi^* \nabla \psi - \psi \nabla \psi^*)
\]

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

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

For a plane wave

\[
\psi = N e^{i(\vec{p} \cdot \vec{r} - Et)}
\]

\[
\rho = 2E |N|^2 \quad \text{and} \quad \vec{j} = |N|^2 \vec{p}
\]

Particle densities are proportional to \(E\). We might have anticipated this from the previous discussion of Lorentz invariant phase space (i.e. density of \(1/V\) in the particles rest frame will appear as \(E/V\) in a frame where the particle has energy \(E\) due to length contraction).
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 are overcome and the KG equation is used to describe spin-0 particles (inherently single particle description → 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.
- The resulting wave equation had solutions which not only solved this problem but also fully describe the intrinsic spin and magnetic moment of the electron!
Dirac Equation

- **Schrödinger eqn:**
  \[ -\frac{1}{2m} \nabla^2 \psi = i \frac{\partial \psi}{\partial t} \]
  \( 1^{\text{st}} \) order in \( \partial / \partial t \)

- **Klein-Gordon eqn:**
  \[ (\partial^\mu \partial_\mu + m^2) \psi = 0 \]
  \( 2^{\text{nd}} \) order throughout

- **Dirac looked for an alternative which was** \( 1^{\text{st}} \) **order throughout:**
  \[ \hat{H} \psi = (\vec{\alpha} \cdot \vec{p} + \beta m) \psi = i \frac{\partial \psi}{\partial t} \]

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

  \( \text{(D1)} \)

- **Writing (D1) 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…**
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:

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

- \((\alpha_x \alpha_y + \alpha_y \alpha_x) \frac{\partial^2 \psi}{\partial x \partial y} - (\alpha_y \alpha_z + \alpha_z \alpha_y) \frac{\partial^2 \psi}{\partial y \partial z} - (\alpha_z \alpha_x + \alpha_x \alpha_z) \frac{\partial^2 \psi}{\partial z \partial x}\)

- \((\alpha_x \beta + \beta \alpha_x) m \frac{\partial \psi}{\partial x} - (\alpha_y \beta + \beta \alpha_y) m \frac{\partial \psi}{\partial y} - (\alpha_z \beta + \beta \alpha_z) m \frac{\partial \psi}{\partial z}\)

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

\( \alpha_x^2 = \alpha_y^2 = \alpha_z^2 = \beta^2 = 1 \)  \( \text{(D2)} \)

\( \alpha_j \beta + \beta \alpha_j = 0 \)  \( \text{(D3)} \)

\( \alpha_j \alpha_k + \alpha_k \alpha_j = 0 \)  \( (j \neq k) \)  \( \text{(D4)} \)

\* Immediately we see that the \( \alpha_j \) and \( \beta \) cannot be numbers. Require 4 mutually anti-commuting matrices.

\* Must be (at least) 4x4 matrices.
Consequently the wave-function must be a four-component Dirac Spinor

\[
\psi = \begin{pmatrix}
\psi_1 \\
\psi_2 \\
\psi_3 \\
\psi_4
\end{pmatrix}
\]

A consequence of introducing an equation that is 1st order in time/space derivatives is that the wave-function has new degrees of freedom!

For the Hamiltonian \( \hat{H} \psi = (\vec{\alpha} \cdot \vec{p} + \beta m) \psi = i\partial \psi / \partial t \) to be Hermitian requires

\[
\alpha_x = \alpha_x^\dagger; \quad \alpha_y = \alpha_y^\dagger; \quad \alpha_z = \alpha_z^\dagger; \quad \beta = \beta^\dagger;
\]

i.e. the require 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 is in the commutation relations.

A convenient choice is based on the Pauli spin matrices:

\[
\beta = \begin{pmatrix}
I & 0 \\
0 & -I
\end{pmatrix}, \quad \alpha_j = \begin{pmatrix}
0 & \sigma_j \\
\sigma_j & 0
\end{pmatrix}
\]

with

\[
I = \begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}, \quad \sigma_x = \begin{pmatrix}
0 & 1 \\
1 & 0
\end{pmatrix}, \quad \sigma_y = \begin{pmatrix}
0 & -i \\
i & 0
\end{pmatrix}, \quad \sigma_z = \begin{pmatrix}
1 & 0 \\
0 & -1
\end{pmatrix}
\]

The matrices are Hermitian and anti-commute with each other
Pauli Matrices

\[
S_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}
\]
\[
S_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}
\]
\[
S_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}
\]

\[
\{\sigma_a, \sigma_b\} = 2\delta_{ab} I
\]
\[
[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k
\]
\[
[\sigma^2, \sigma_j] = 0
\]

\[
\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = -i\sigma_1\sigma_2\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I
\]

\[
\sigma^2 = \sigma_x^2 + \sigma_y^2 + \sigma_z^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}
\]
Probability Density

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

- Start with the Dirac equation

\[
- i \alpha_x \frac{\partial \psi}{\partial x} - i \alpha_y \frac{\partial \psi}{\partial y} - i \alpha_z \frac{\partial \psi}{\partial z} + m \beta \psi = i \frac{\partial \psi}{\partial t} \tag{D6}
\]

and its Hermitian conjugate

\[
+ i \frac{\partial \psi^\dagger}{\partial x} \alpha_x^\dagger + i \frac{\partial \psi^\dagger}{\partial y} \alpha_y^\dagger + i \frac{\partial \psi^\dagger}{\partial z} \alpha_z^\dagger + m \psi^\dagger \beta^\dagger = -i \frac{\partial \psi^\dagger}{\partial t} \tag{D7}
\]

- Consider \( \psi^\dagger \times (D6) - (D7) \times \psi \) remembering \( \alpha, \beta \) are Hermitian

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

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

- Now using the identity:

\[
\psi^\dagger \alpha_x \frac{\partial \psi}{\partial x} + \frac{\partial \psi^\dagger}{\partial x} \alpha_x \psi \equiv \frac{\partial (\psi^\dagger \alpha_x \psi)}{\partial x}
\]
where $\psi^\dagger = (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*)$

• The probability density and current can be identified as:

\[
\rho = \psi^\dagger \psi \quad \text{and} \quad \vec{j} = \psi^\dagger \vec{\alpha} \psi
\]

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 components naturally give rise to the property of intrinsic spin.

• It can be shown that Dirac spinors represent spin-half particles with an intrinsic magnetic moment of

\[
\vec{\mu} = \frac{q}{m} \vec{S}
\]
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 (D6) by \( \beta \):

\[
\begin{align*}
&i\beta \alpha_x \frac{\partial \psi}{\partial x} + i\beta \alpha_y \frac{\partial \psi}{\partial y} + i\beta \alpha_z \frac{\partial \psi}{\partial z} - \beta^2 m \psi = -i\beta \frac{\partial \psi}{\partial t} \\
\rightarrow &\quad i\gamma^1 \frac{\partial \psi}{\partial x} + i\gamma^2 \frac{\partial \psi}{\partial y} + i\gamma^3 \frac{\partial \psi}{\partial z} - m \psi = -i\gamma^0 \frac{\partial \psi}{\partial t}
\end{align*}
\]

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

\[
\left( i \gamma^\mu \partial_\mu - m \right) \psi = 0 \quad \text{(D9)}
\]

★ **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.
Properties of \( \gamma \) matrices

- From the properties of the \( \alpha \) and \( \beta \) matrices (D2)-(D4) immediately obtain:

\[
(\gamma^0)^2 = \beta^2 = 1 \quad \text{and} \quad (\gamma^1)^2 = \beta \alpha_x \beta \alpha_x = -\alpha_x \beta \beta \alpha_x = -\alpha_x^2 = -1
\]

- The full set of relations is

\[
\begin{align*}
(\gamma^0)^2 &= 1 \\
(\gamma^1)^2 &= (\gamma^2)^2 = (\gamma^3)^2 = -1 \\
\gamma^0 \gamma^j + \gamma^j \gamma^0 &= 0 \\
\gamma^j \gamma^k + \gamma^k \gamma^j &= 0 \quad (j \neq k)
\end{align*}
\]

which can be expressed as:

\[
\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu \nu} \quad \text{(defines the algebra)}
\]

- Are the gamma matrices Hermitian?
  - \( \beta \) is Hermitian so \( \gamma^0 \) is Hermitian.
  - The \( \alpha \) 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 \( \gamma^1, \gamma^2, \gamma^3 \) are anti-Hermitian
• 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

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

• 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 \( j^\mu \) is the four-vector current.

• 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 is defined as

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

i.e.

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

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

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...