Chapter 5

Towards the Dirac Equation

Dirac tried to find a partial differential equation with a positive probability density. He realized that the problem in (4.13) resulted from the fact that one started from the non-relativistic definition \( \tilde{j}_{NR} \) in (4.9). This equation cannot determine a definite sign. For solving the problem one could start from \( \rho_{NR} \) but then one would need a different equation of 1st order since the possibility of defining \( \rho = \psi^* \psi \) is related to the fact that the Schrödinger equation contains only a first derivative with respect to the time. Thus one faces the following problem; that one needs a differential equation (1st order) for \( \psi(x) \) which leads to the correct energy-momentum relation

\[
p_{\mu}p^\mu = \frac{E^2}{c^2} - \vec{p}^2 = m^2 c^2.
\]

This problem cannot be solved for a scalar function \( \psi \) and the crucial point of Dirac’s idea was to introduce a ”vector” quantity

\[
\psi(x) = \begin{pmatrix}
\psi_1(x) \\
\psi_2(x) \\
\vdots \\
\psi_n(x)
\end{pmatrix}
\]

and thus a positive definite density

\[
\rho(x) := \sum_{\alpha=1}^{n} \psi_\alpha^*(x) \psi_\alpha(x).
\]
The most general form of a first-order differential equation for this \( \psi(x) \) is given by

\[
\sum_\beta [i \gamma^\mu_\alpha \partial_\mu - m \Gamma_{\alpha\beta}] \psi_\beta(x) = 0
\]  

(5.4)

for \( \alpha = 1, 2, \ldots, n \) (notation consistent with the one of Bjorken/Drell).

Here \( \gamma^\mu_\alpha \) and \( \Gamma_{\alpha\beta} (\mu = 0, 1, 2, 3 \ ; \alpha, \beta = 1, 2, \ldots, n) \) are coefficients which we assume to be complex constants since we do not consider forces. Note that \( i \partial_\mu \) is the operator for the four-momentum. With the matrix notation \( \gamma^\mu := (\gamma^\mu_\alpha) \) and \( \Gamma := (\Gamma_{\alpha\beta}) \Gamma \) Eq. (5.4) can be written as

\[
(i \gamma^\mu \partial_\mu - m \Gamma) \psi(x) = 0
\]

(5.5)

so that energy and momentum fulfill the relation (3.28). Dirac required that each component \( \psi_\alpha(x) \) of the wave function separately fulfills the Klein-Gordon equation (3.32):

\[
(\Box + m^2) \psi_\alpha(x) = 0
\]

(5.6)

for \( \alpha = 1, 2, \ldots, n \). The requirement is compatible with (5.5) and leads to relations for the \( \gamma^\mu \) and \( \Gamma \).

To illustrate this further we want to mention that a similar formal relation occurs when considering electromagnetic waves. There are free Maxwell equations

\[
\begin{align*}
\text{div } \vec{E} &= 0 & \text{rot } \vec{B} - \frac{1}{c} \ddot{\vec{E}} &= 0 \\
\text{div } \vec{B} &= 0 & \text{rot } \vec{E} + \frac{1}{c} \ddot{\vec{B}} &= 0
\end{align*}
\]

(5.7)

represent a first-order differential equation for \( \left( \begin{array}{c} \vec{E} \\ \vec{B} \end{array} \right) \). It can be recast into the form (5.4) where the \( \gamma^\mu_\alpha \) turn to six-dimensional matrices and \( \Gamma_{\alpha\beta} \) vanish. On the other hand follows from Maxwell’s equations

\[
\Box \left( \begin{array}{c} \vec{E} \\ \vec{B} \end{array} \right) = 0
\]

(5.8)
i.e. the same wave equation is valid for each component. We proceeded the other way. Starting from an equation

\[(\Box + m^2)\psi(x) = 0\]  \hspace{1cm} (5.9)

for the components we want to determine the matrices \(\gamma^\mu\) and \(\Gamma\). For this we apply onto (5.5) the 'conjugate' operator \((-i\gamma^{\nu} \partial_\gamma - m\Gamma)\):

\[
(-i\gamma^\nu \partial_\nu - m\Gamma)(i\gamma^\mu \partial_\mu - m\Gamma) \psi = 0
\]

\[
[\gamma^{\nu} \gamma^\mu \partial_{\nu\mu}^2 + im(\gamma^{\nu} \Gamma - \Gamma \gamma^\nu) \partial_\nu + m^2 \Gamma^2] \psi = 0
\]  \hspace{1cm} (5.10)

where

\[
\partial_{\nu\mu}^2 := \partial_\nu \partial_\mu = \frac{\partial^2}{\partial x^\nu \partial x^\mu}
\]  \hspace{1cm} (5.11)

and \(\nu\) being the common summation index. A comparison with (5.6) gives

\[
\gamma^{\nu} \gamma^\nu \partial_{\nu\mu}^2 = \Box \mathbf{1}
\]

\[
\gamma^{\nu} \Gamma - \Gamma \gamma^\nu = 0
\]

\[
\Gamma^2 = \mathbf{1}
\]  \hspace{1cm} (5.12)

These equations determine five matrices \(\gamma^0, \gamma^1, \gamma^2, \gamma^3, \Gamma\) uniquely up to similarity transformations.

Together with \(\gamma^\mu, \Gamma\) the quantities

\[
\hat{\gamma}^\mu = S\gamma^\mu S^{-1}
\]

\[
\hat{\Gamma} = S\Gamma S^{-1}
\]  \hspace{1cm} (5.13)

fulfill equations (5.12) where \(S\) is an arbitrary non-singular matrix. Next one has to explicitly construct the \(\gamma\)-matrices. First from (5.12) follows that \(\Gamma\) is a non-singular matrix:

\[
\Gamma^{-1} = \Gamma,
\]  \hspace{1cm} (5.14)

thus multiplying (5.5) with \(\Gamma^{-1}\) gives

34
(i\gamma^\mu \partial_\mu - m \mathbf{1}) \psi(x) = 0 \quad \text{(5.15)}

Without losing generality we can choose \( \Gamma = \mathbf{1} \Gamma \) and thus the second equation (5.12) is fulfilled.

Thus the Dirac equation takes the form

\[(i\gamma^\mu \partial_\mu - m) \psi(x) = 0 \quad \text{(5.16)}\]

or when introducing \( \hbar \) or \( c \) explicitly

\[(i\hbar \gamma^\mu \partial_\mu - mc) \psi(x) = 0 \quad \text{(5.17)}\]

The latter can be verified by dimensional considerations:

According to the first Eq. (5.12) the \( \gamma \)-matrices are dimensionless. \( i\hbar \partial_m u \) stands for the four-momentum and thus the second term in (5.17) has to contain the momentum \( mc \).

Considering the first Eq. (5.12) shows that due to the symmetry \( \partial^2_{\mu\nu} = \partial^2_{\nu\mu} \)

\[\gamma^\mu \gamma^\nu \partial^2_{\mu\nu} = \frac{1}{2} (\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu) \partial^2_{\mu\nu}. \quad \text{(5.18)}\]

On the other hand

\[\Box = \partial^\nu \partial_\nu = g^\mu\nu \partial^2_{\mu\nu}, \quad \text{(5.19)}\]

from which follows

\[\{\gamma^\mu, \gamma^\nu\} := \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^\mu\nu. \quad \text{(5.20)}\]

These equations define the Dirac \( \gamma \)-matrices. The Dirac problem is solved as soon as we find matrices fulfilling (5.20).

**In Summary:** With the help of the \( \gamma \)-matrices it is possible to explicitly take the square root of the d’Alembert operator:

\[\sqrt{-\Box} = i\gamma^\mu \partial_\mu. \quad \text{(5.21)}\]
In fact:

\[
(i \gamma^\mu \partial_\mu)^2 = -\gamma^\mu \gamma^\nu \partial_{\mu\nu}^2 \\
= -\frac{1}{2} (\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu) \partial_{\mu\nu}^2 \\
= -g^{\mu\nu} \partial_{\mu\nu}^2 = -\Box. \tag{5.22}
\]

This last calculation clearly stresses the importance of the anti-commutation relation (5.20).