Chapter 6

Relativistic quantum mechanics

The Schrödinger equation for a free particle in the coordinate representation,

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi,$$

is manifestly *not Lorentz constant* since time and space derivatives are of different order. This implies that it changes its structure under a transformation from one inertial system to another.

By using the correspondence principle, Schrödinger, Gordon, and Klein formulated in 1926-1927 the so-called *Klein-Gordon equation*, which was a scalar relativistic wave equation of second order (see for instance: E. Schrödinger, Ann Physik **81**, 109 (1926); W.Gordon, Z. Physik **40**, 117 (1926); O.Klein, Z. Physik **41**, 407 (1927)). This equation was, though, initially dismissed because it led to negative probability densities.

Dirac (P.A.M. Dirac, Proc. Roy. Soc. (London) **A117**, 610 (1928)) published in 1928 the *Dirac equation*, which is a relativistic equation for spin-1/2 particles. This equation has, like the Klein-Gordon equation, solutions with *negative energy*. In order to avoid transitions of an electron to states of negative energy, Dirac postulated in 1930 that the states of negative energy should all be occupied and the particles that are missing in these occupied states represent particles of opposite charge (antiparticles). Pauli and Weisskopf, using this concept, also interpreted the Klein-Gordon equation as such that could be used to describe f.i. mesons with spin zero (π-mesons).

6.1   The Klein-Gordon equation

6.1.1   Derivation

We shall consider the correspondence principle in order to derive the Klein- Gordon equation.

Following the correspondence principle, we can replace classical quantities by operators

as

$$\text{energy} \quad E \;\rightarrow\; i\hbar\frac{\partial}{\partial t}$$

$$\text{momentum} \quad \vec{p} \;\rightarrow\; -i\hbar\vec{\nabla}$$

and obtain from the non-relativistic energy of a free particle

$$E = \frac{\vec{p}^2}{2m},$$

its time-dependent Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\Psi = -\frac{\hbar^2\nabla^2}{2m}\Psi$$

which is not Lorentz invariant.

In order to introduce the relativistic formulation, we first define our notation according to the special theory of relativity:

1 - components of space-time four-vectors are denoted by Greek indices;

2 - components of spacial three-vectors are denoted by Latin indices or cartesian coordinates (x, y, z);

3 - we use Einstein's summation convention: Greek or Latin indices that appear twice, one contravariant and one covariant, are summed over;

4 - notation for the *Lorentz transformation* is as follows.

(i) The contravariant and covariant components of the position vector are

$$
\begin{array}{llllll}
x^\mu: & x^0 = ct & x^1 = x & x^2 = y & x^3 = z & \text{(contravariant)} \\
x_\mu: & x_0 = ct & x_1 = -x & x_2 = -y & x_3 = -z & \text{(covariant)}
\end{array}
$$

The metric tensor g,

$$g = g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \text{(Minkowski metric)},$$

relates covariant and contravariant components as

$$x_\mu = g_{\mu\nu}x^\nu, \quad x^\mu = g^{\mu\nu}x_\nu.$$

For g_ν^μ we have that

$$g_\nu^\mu = g^{\mu\sigma}g_{\sigma\nu} \equiv \delta_\nu^\mu$$

and

$$g^\mu_\nu = \delta^\mu_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

For instance, the d'Alembert operator $\square$ is defined in terms of $g_{\mu\nu}$ as

$$\square = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \sum_{i=1}^{3}\frac{\partial^2}{\partial(x^i)^2} = \partial_\mu\partial^\mu = g_{\mu\nu}\partial^\mu\partial^\nu,$$

where the *covariant* vector

$$\partial_\mu = \frac{\partial}{\partial x^\mu}$$

is the four-dimensional generalization of the gradient vector. Remember that in electrodynamics the d'Alembert operator $\square \equiv \partial_\mu\partial^\mu$ is invariant under Lorentz transformations.

(ii) *Inertial frames* are frames of reference in which, in the absence of forces, particles move uniformly. The Lorentz transformations tell us how the coordinates of two inertial frames transform into one another. The coordinates of two reference systems in uniform motion must be related to one another by a *linear transformation*. Therefore, the *inhomogenious Lorentz transformation* has the form

$$\boxed{x'^\mu = \Lambda^\mu_\nu x^\nu + a^\mu},$$

where Λ^μ_ν and a^μ are real.

(iii) *World line* of an object is the sequence of *space-time* events corresponding to the history of the object. In an internal reference frame, the world line of a particle is a straight line.

(iv) *Principle of relativity*: the laws of nature are the same in all inertial frames. For instance, the requirement that the d'Alambert operator be invariant under Lorentz transformation yields

$$\boxed{\Lambda^\nu_\mu g^{\mu\nu}\Lambda^\rho_\nu = g^{\lambda\rho}} \tag{6.1}$$

or, in matrix form,

$$\boxed{\Lambda g\Lambda^T = g}. \tag{6.2}$$

Relations (6.1) and (6.2) define the Lorentz transformations. A Lorentz transformation can be denoted by the form (Λ, a) with

→ translation group $(1, a)$

→ rotation group $(\Lambda, 0)$

The set of Lorentz transformations form the Lorentz group. The set of transformations that also includes translations is described as the Poincaré group.

(v) In relativity, *proper time* is the *time* between events occuring at the same place as where the single clock that measures it is located. The proper time depends on the events and on the motion of the clock between the events. An *accelerating* clock will measure a *shorter* proper time between two events than a *non-accelerating* (inertial) clock would do for the same events.

5 -
$$x^\mu(s) = (ct, \vec{x})$$
is the contravariant four-vector representation of the world line as a function of proper time s. The differential of the proper time, ds, is related to dx^0 via

$$ds = \sqrt{1 - (v/c)^2}\, dx^0,$$

where $\vec{v} = c\frac{d\vec{x}}{dx^0}$ is the velocity. One can also show that

$$\dot{x}^\mu(s) = \frac{dx}{ds} = \frac{dx}{dt} \cdot \frac{dt}{ds} = \frac{v/c}{\sqrt{1 - (v/c)^2}}.$$

6 - for the four-momentum we have

$$p^\mu = m\,c\,\dot{x}^\mu(s) = \frac{1}{\sqrt{1 - (v/c)^2}} \begin{pmatrix} mc \\ m\vec{v} \end{pmatrix} = \begin{pmatrix} E/c \\ \vec{p} \end{pmatrix},$$

with the definition

$$E/c = p^0 = \frac{m\,c}{\sqrt{1 - (v/c)^2}} \quad \text{(kinetic energy of the particle)}$$

Therefore, according to the special theory of relativity, the energy E and the momentum (p_x, p_y, p_z) transform as the components of a *contravariant four-vector*:

$$\boxed{p^\mu = (p^0, p^1, p^2, p^3) = \left(\frac{E}{c}, p_x, p_y, p_z \right)}.$$

7 - with the metric tensor $g_{\mu\nu}$, we obtain the *covariant* components of the four-momentum:

$$p_\mu = g_{\mu\nu}p^\nu = \left(\frac{E}{c}, -\vec{p} \right).$$

The *invariant* scalar product of the four-momentum is given by

$$p_\mu p^\mu = \frac{E^2}{c^2} - \vec{p}^2 = m^2 c^2,$$

where m is the rest mass and c is the speed of light. By using special relativity, we obtain the energy-momentum relation for a free particle:

$$\boxed{E = \sqrt{\vec{p}^2 c^2 + m^2 c^4}}. \tag{6.3}$$

We apply now the *correspondence principle* to eq. (6.3) and obtain the wave equation

$$i\hbar\frac{\partial}{\partial t}\Psi = \sqrt{-\hbar^2 c^2 \nabla^2 + m^2 c^4}\,\Psi,$$

in which time and space do not occur symmetrically because a Taylor expansion of the square root leads to infinitely high derivatives in space. But if, instead of eq. (6.3), we start from the squared equation,

$$E^2 = \vec{p}^{\,2}c^2 + m^2 c^4,$$

we arrive at

$$-\hbar^2\frac{\partial^2}{\partial t^2}\Psi = \left(-\hbar^2 c^2 \vec{\nabla}^2 + m^2 c^4\right)\Psi,$$

which can be written in the following *Lorentz covariant form*:

$$\left(\partial_\mu\partial^\mu + \left(\frac{mc}{\hbar}\right)^2\right)\Psi = 0, \tag{6.4}$$

with $x^\mu = (x^0 = ct, \vec{x})$ being the space-time position vector.

Lorentz covariant form of an equation: an equation is *Lorentz covariant* if it can be written in terms of Lorentz covariant quantities (i.e. quantities that transform under a given representation of the Lorentz group). Such an equation holds then in any *inertial frame*, which is a requirement of the principle of relativity.

Eq. (6.4) is the Klein-Gordon equation. Let us now consider its properties.

6.1.2 The continuity equation

We first use the Klein-Gordon equation to derive the continuity equation.

Multiplying the Klein-Gordon equation by Ψ^*

$$\Psi^*\left(\partial_\mu\partial^\mu + \left(\frac{mc}{\hbar}\right)^2\right)\Psi = 0 \tag{6.5}$$

and subtracting from (6.5) its complex conjugate

$$\Psi\left(\partial_\mu\partial^\mu + \left(\frac{mc}{\hbar}\right)^2\right)\Psi^* = 0,$$

we get

$$\Psi^*\partial_\mu\partial^\mu\Psi - \Psi\partial_\mu\partial^\mu\Psi^* = 0$$

and then

$$\partial_\mu\left(\Psi^*\partial^\mu\Psi - \Psi\partial^\mu\Psi^*\right) = 0. \tag{6.6}$$

We multiply now (6.6) by $\frac{\hbar}{2mi}$ and obtain (writing explicitly)

$$\frac{\partial}{\partial t}\left(\frac{i\hbar}{2mc^2}\left(\Psi^*\frac{\partial\Psi}{\partial t} - \Psi\frac{\partial\Psi^*}{\partial t}\right)\right) + \vec{\nabla}\cdot\frac{\hbar}{2mi}\left[\Psi^*\nabla\Psi - \Psi\nabla\Psi^*\right] = 0,$$

which has the form of the *continuity equation*

$$\boxed{\dot{\rho} + \vec{\nabla}\cdot\vec{j} = 0}\,,$$

with the density

$$\boxed{\rho = \frac{i\hbar}{2mc^2}\left(\Psi^*\frac{\partial\Psi}{\partial t} - \Psi\frac{\partial\Psi^*}{\partial t}\right)}$$

and the current density

$$\boxed{\vec{j} = \frac{\hbar}{2mi}\left(\Psi^*\vec{\nabla}\Psi - \Psi\vec{\nabla}\Psi^*\right)}.$$

Note:

1) ρ is not positive-defined and therefore cannot be directly interpreted as a probability density;

2) the Klein-Gordon equation is a second-order differential equation in t and therefore the initial values of Ψ and $\frac{\partial\Psi}{\partial t}$ are chosen independently, so that ρ as a function of $\vec{x}$ can be positive and negative.

6.1.3   Free solutions of the Klein-Gordon equation

The Klein-Gordon equation has two free solutions

$$\Psi(\vec{x},t) = e^{i(Et-\vec{p}\cdot\vec{x})/\hbar},$$

with

$$E = \pm\sqrt{\vec{p}^2c^2 + m^2c^4}.$$

Both solutions with positive and negative energies are possible, and the energy is *not* bounded from below, which does not make sense for a free particle. Moreover, this is a scalar theory, which doesn't contain spin and therefore can only describe spin-zero particles.

For these reasons, the Klein-Gordon equation was rejected. One year later Dirac proposed his equation for spin-1/2 fermions. He interpreted the negative energy solutions by assuming that the unoccupied states of negative energy describe antiparticles. The Klein-Gordon equation, though, is still suitable for describing mesons.

6.2 Dirac equation

6.2.1 Derivation of the Dirac equation

Dirac postulated a differential equation of first order such that

1 - the density

$$\rho = \left(\Psi^* \frac{\partial}{\partial t} \Psi + c.c. \right),$$

which contains the first time derivative, is *positive definite*;

2 - the equation is *relativistic covariant*. This means that its *spatial derivatives* are of the same order as the time derivatives, that is, of first order.

We are thus seeking for a wave equation of the form

$$\boxed{ i\hbar \frac{\partial}{\partial t} \Psi = \left(\frac{\hbar c}{i} \alpha^k \partial_k + \beta mc^2 \right) \Psi \equiv H\Psi }, \tag{6.7}$$

where α^k, β have to be yet defined.

The *Dirac Hamiltonian* (6.7) is linear in the momentum operator and in the energy at rest. Therefore, α^k and β cannot be scalar numbers. If they were, then the equation would not be form invariant with respect to spatial rotations. α^k and β must be hermitian matrices for H to be hermitian.

Let α^k and β be $N \times N$ matrices and

$$\Psi = \left(\begin{array}{c} \Psi_1 \\ \vdots \\ \Psi_N \end{array} \right)$$

be an N-component column vector. N is not yet determined.
We then impose the following requirements on eq. (6.7):

(i) The components of Ψ must satisfy the Klein-Gordon equation, so that plane waves fulfill the relativistic energy-momentum relation

$$E^2 = p^2 c^2 + m^2 c^4.$$

(ii) There exists a conserved four-component current j^μ whose zeroth component is a *positive density*.

(iii) The equation must be *Lorentz covariant*, which means that it must have the same form in all reference frames that are connected by a Lorentz (Poincaré) transformation. This requirement is fulfilled since (6.7) is of the first order.

In order to fulfill condition (i), we consider the second derivative of eq. (6.7)

$$
\begin{aligned}
-\hbar^2 \frac{\partial^2}{\partial t^2} \Psi &= -\hbar^2 c^2 \sum_{ij} \frac{1}{2} (\alpha^i \alpha^j + \alpha^j \alpha^i) \partial_i \partial_j \Psi \\
&+ \frac{\hbar m c^3}{i} \sum_{i=1}^{3} (\alpha^i \beta + \beta \alpha^i) \partial_i \Psi + \beta^2 m^2 c^4 \Psi,
\end{aligned}
\tag{6.8}
$$

where we used $\partial_i \partial_j = \partial_j \partial_i$. Comparing eq. (6.8) with the Klein-Gordon equation

$$
\left[\partial_\mu \partial^\mu + \left(\frac{mc}{\hbar} \right)^2 \right] \Psi = 0 ,
$$

we obtain three conditions

$$
\begin{aligned}
\alpha^i \alpha^j + \alpha^j \alpha^i &= 2\delta^{ij} \mathbf{1} \\
\alpha^i \beta + \beta \alpha^i &= 0 \\
(\alpha^i)^2 &= \beta^2 = \mathbf{1}
\end{aligned}
\tag{6.9}
$$

In order to fulfill condition (ii), let us consider the continuity equation.

6.2.2 Continuity equation

The adjoint of Ψ is

$$
\Psi^\dagger = (\Psi_1^*, \Psi_2^*, \ldots, \Psi_N^*).
$$

We multiply the Dirac equation (6.7) from the left by $\Psi^\dagger$

$$
i\hbar \Psi^\dagger \frac{\partial \Psi}{\partial t} = \frac{\hbar c}{i} \Psi^\dagger \alpha^i \partial_i \Psi + mc^2 \Psi^\dagger \beta \Psi.
$$

The corresponding complex conjugate relation is

$$
-i\hbar \frac{\partial}{\partial t} \Psi^\dagger \Psi = -\frac{\hbar c}{i} (\partial_i \Psi^\dagger) \alpha^{i\dagger} \Psi + mc^2 \Psi^\dagger \beta^\dagger \Psi,
$$

and the difference between these two relations is then

$$
\begin{aligned}
\frac{\partial}{\partial t} (\Psi^\dagger \Psi) &= -c \left((\partial_i \Psi^\dagger) \alpha^{i\dagger} \Psi + \Psi^\dagger \alpha^i \partial_i \Psi \right) \\
&+ \frac{imc^2}{\hbar} (\Psi^\dagger \beta^\dagger \Psi - \Psi^\dagger \beta \Psi).
\end{aligned}
$$

For this equation to resemble a continuity equation we need the matrices α and β to be hermitian, i.e.

$$
\alpha^{i\dagger} = \alpha^i, \quad \beta^\dagger = \beta .
$$

With these conditions, the *density*

$$
\rho \equiv \Psi^\dagger \Psi = \sum_{\alpha=1}^{N} \Psi_\alpha^* \Psi_\alpha
\tag{6.10}
$$

and the *current density*

$$\boxed{j^k \equiv c\Psi^\dagger \alpha^k \Psi}$$

satisfy the *continuity equation*

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

The zeroth component of j^μ is

$$j^0 = c\rho,$$

and the *four-current-density* is defined as

$$j^\mu \equiv (j^0, j^k).$$

With this, the continuity equation can be written as

$$\boxed{\partial_\mu j^\mu = \frac{1}{c}\frac{\partial}{\partial t}j^0 + \frac{\partial}{\partial x^k}j^k = 0},$$

which is a nice compact form. The density defined in eq. (6.10) is positive definite and can be given the interpretation of a *probabiliy density.*

6.2.3   Properties of the Dirac matrices α^k and β

1 - The matrices α^k and β anticommute (see eq. (6.9)).

2 - From $(\alpha^k)^2 = \beta^2 = \mathbf{1}$ follows that the matrices α^k and β have only eigenvalues ± 1.

3 - We consider eq. (6.9)

$$\alpha^k\beta + \beta\alpha^k = 0.$$

For

$$\alpha^k = -\beta\alpha^k\beta$$

we see, using the cyclic invariance of the trace, that

$$Tr\alpha^k = -Tr\beta\alpha^k\beta = -Tr\alpha^k\beta^2 = -Tr\alpha^k.$$

Analogously,

$$Tr\beta = -Tr\beta.$$

So,

$$\boxed{Tr\alpha^k = Tr\beta = 0},$$

which means that the number of *positive* and *negative* eigenvalues must be equal, and therefore N (the dimension of the matrix) must be even.

$N = 2$ is not sufficient since among the 2×2 matrices that are known, $\mathbf{1}$, σ_x, σ_y, and σ_z, only three are mutually anticomuting; the $\mathbf{1}$ commutes with the rest. Thus, $N = 4$ is the smallest dimension in which it is possible to fulfill (6.7)-(6.9).

A concrete representation of the matrices is

$$\alpha^i = \begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}, \tag{6.11}$$

with σ^i being the Pauli matrices

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

and $\mathbf{1}$ being the 2-dimensional identity. α^i and β defined in this way satisfy the condition (6.9)

$$\alpha^i \beta + \beta \alpha^i = \begin{pmatrix} 0 & -\sigma^i \\ \sigma^i & 0 \end{pmatrix} + \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} = 0.$$

→ The Dirac equation (6.7) together with the definitions of α^i and β (6.11) is called "standart representation" of the Dirac equation.

→ Ψ is a four-spinor, or bispinor, when it is represented by two-component spinors.

→ $\Psi^\dagger$ is the hermitian adjoint spinor.

6.2.4   Covariant form of the Dirac equation

We multiply the Dirac equation (6.7) by $\frac{\beta}{c}$

$$-i\,\hbar\,\beta\,\partial_0\Psi - i\,\hbar\,\beta\,\alpha^i\,\partial_i\Psi + mc\Psi = 0$$

and introduce the *Dirac matrices* γ^0 and γ^i as

$$\boxed{\gamma^0 \equiv \beta \quad \text{and} \quad \gamma^i \equiv \beta\alpha^i}.$$

γ^0 and γ^i have the following properties.

1 - γ^0 is hermitian and $(\gamma^0)^2 = \mathbf{1}$.

2 - γ^k is anti-hermitian:
$$(\gamma^k)^\dagger = -\gamma^k \quad \text{and} \quad (\gamma^k)^2 = -\mathbf{1}.$$

   Proof:

$$\begin{aligned} (\gamma^k)^\dagger &= \alpha^k \beta = -\beta \alpha^k = -\gamma^k; \\ (\gamma^k)^2 &= \beta \alpha^k \beta \alpha^k = -\mathbf{1}. \end{aligned}$$

3 - $\gamma^0 \gamma^k + \gamma^k \gamma^0 = \beta\beta\alpha^k + \beta\alpha^k\beta = 0.$

4 - $\gamma^k \gamma^l + \gamma^l \gamma^k = \beta\alpha^k \beta\alpha^l + \beta\alpha^l \beta\alpha^k = 0 \quad$ for $k \neq l.$

Properties 1-4 can be summarized as

$$\boxed{\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}\mathbf{1}}.$$

Having introduced γ^i and γ^0, the *Dirac equation* (6.7) has now the form

$$\boxed{\left(-i\gamma^\mu \partial_\mu + \frac{mc}{\hbar}\right)\Psi = 0} . \tag{6.12}$$

This equation can be written in a more compact form using the following *Feynman notation*:

$$\rlap{/}{v} \equiv \gamma v \equiv \gamma^\mu v_\mu = \gamma_\mu v^\mu = \gamma^0 v^0 - \vec{\gamma} \cdot \vec{v} \quad (v^\mu \text{ denotes any vector}).$$

The *Feynman slash* indicates multiplication by the matrix γ_μ. Eq. (6.12) thus becomes

$$\boxed{\left(-i\rlap{/}{\partial} + \frac{mc}{\hbar}\right)\Psi = 0},$$

with

$$\gamma^0 = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \quad \text{and} \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}.$$

Note: one can obtain other representations of the γ matrices by considering a non-singular transformation M:

$$\gamma \rightarrow M\gamma M^{-1}.$$

For instance, the so-called Majorana representation of γ or the Chiral representation of γ are also used.

6.2.5   Non-relativistic limit of the Dirac equation

Particle at rest

Consider eq. (6.7),

$$i\hbar \frac{\partial \Psi}{\partial t} = \left(\frac{\hbar c}{i} \alpha^l \partial_l + \beta mc^2\right)\Psi,$$

for a free particle *at rest*, i.e. with wavevector $\vec{k} = 0$. In this case, the spatial derivatives vanish, and the equation becomes

$$i\hbar \frac{\partial \Psi}{\partial t} = \beta mc^2 \Psi.$$

The solutions of this equation are

$$\Psi_1^{(+)} = e^{-\frac{imc^2}{\hbar}t} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad \Psi_2^{(+)} = e^{-\frac{imc^2}{\hbar}t} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

$$\Psi_1^{(-)} = e^{\frac{imc^2}{\hbar}t} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \Psi_2^{(-)} = e^{\frac{imc^2}{\hbar}t} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

$\Psi_1^{(+)}$ and $\Psi_2^{(+)}$ are positive-energy solutions, and $\Psi_1^{(-)}$ and $\Psi_2^{(-)}$ are negative-energy solutions. We examine now the positive-energy solutions. The negative-energy solutions will be discussed later. It is experimentally shown that the negative solutions are *physical solutions* and describe antiparticles (positrons).

Coupling to the electromagnetic field

The coupling to the electromagnetic field is included into consideration by

- replacing the canonical momentum $\vec{p}$ by the kinetic momentum $\left(\vec{p} - \frac{e}{c}\vec{A}\right)$ and

- augmenting the rest energy in the Dirac Hamiltonian by the scalar electric field $e\Phi$,

so that (6.7) becomes

$$\boxed{i\hbar\frac{\partial\Psi}{\partial t} = \left(c\vec{\alpha}\cdot\left(\vec{p} - \frac{e}{c}\vec{A}\right) + \beta mc^2 + e\Phi\right)\Psi}. \qquad (6.13)$$

Here, e is the charge of the particle, with $e = -e_0$ for the electron.

Non-relativistic limit: the Pauli equation

In order to derive the *non-relativistic limit* of the Dirac equation, we consider the representation of the Dirac matrices in (6.11) and decompose the four-spinor into 2-component vectors $\tilde{\varphi}$ and $\tilde{\chi}$

$$\Psi \equiv \begin{pmatrix} \tilde{\varphi} \\ \tilde{\chi} \end{pmatrix},$$

with

$$\boxed{i\hbar\frac{\partial}{\partial t}\begin{pmatrix} \tilde{\varphi} \\ \tilde{\chi} \end{pmatrix} = c\begin{pmatrix} \vec{\sigma}\cdot\vec{\pi}\,\tilde{\chi} \\ \vec{\sigma}\cdot\vec{\pi}\,\tilde{\varphi} \end{pmatrix} + e\Phi\begin{pmatrix} \tilde{\varphi} \\ \tilde{\chi} \end{pmatrix} + mc^2\begin{pmatrix} \tilde{\varphi} \\ -\tilde{\chi} \end{pmatrix}},$$

where

$$\vec{\pi} = \vec{p} - \frac{e}{c}\vec{A}$$

is the kinetic momentum operator.

In the non-relativistic limit, the rest energy mc^2 is the largest energy involved. Therefore, we take as *positive-energy solutions* wavefunctions of the form

$$\begin{pmatrix} \tilde{\varphi} \\ \tilde{\chi} \end{pmatrix} = e^{-\frac{imc^2}{\hbar}t} \begin{pmatrix} \varphi \\ \chi \end{pmatrix},$$

where

$$\begin{pmatrix} \varphi \\ \chi \end{pmatrix}$$

are considered to vary slowly with time and satisfy the equation

$$\boxed{i\hbar\frac{\partial}{\partial t} \begin{pmatrix} \varphi \\ \chi \end{pmatrix} = c \begin{pmatrix} \vec{\sigma}\cdot\vec{\pi}\,\chi \\ \vec{\sigma}\cdot\vec{\pi}\,\varphi \end{pmatrix} + e\Phi \begin{pmatrix} \varphi \\ \chi \end{pmatrix} - 2mc^2 \begin{pmatrix} 0 \\ \chi \end{pmatrix}}. \tag{6.14}$$

For the equation for χ, one may neglect $\hbar\frac{\partial}{\partial t}\chi$ and $e\Phi\chi$ in comparison to $2mc^2\chi$ and get

$$c\,\vec{\sigma}\cdot\vec{\pi}\,\varphi = 2mc^2\chi,$$

$$\boxed{\chi = \frac{\vec{\sigma}\cdot\vec{\pi}}{2mc}\varphi}. \tag{6.15}$$

We see from this equation that χ is a factor $\sim \frac{v}{c}$ smaller than φ. Therefore,

φ is the large component of the spinor and

χ is the small component of the spinor.

We insert (6.15) in the first equation of (6.14) and obtain

$$\boxed{i\hbar\frac{\partial\varphi}{\partial t} = \left(\frac{1}{2m}(\vec{\sigma}\cdot\vec{\pi})(\vec{\sigma}\cdot\vec{\pi}) + e\Phi \right)\varphi}. \tag{6.16}$$

Then, using the identity

$$(\vec{\sigma}\cdot\vec{a})(\vec{\sigma}\cdot\vec{b}) = \vec{a}\cdot\vec{b} + i\vec{\sigma}(\vec{a}\times\vec{b}),$$

with

$$\sigma^i\sigma^j = \delta_{ij} + i\epsilon^{ijk}\sigma^k,$$

we get

$$(\vec{\sigma}\cdot\vec{\pi})(\vec{\sigma}\cdot\vec{\pi}) = \vec{\pi}^2 + i\vec{\sigma}(\vec{\pi}\times\vec{\pi}) = \boxed{\vec{\pi}^2 - \frac{e\hbar}{c}\vec{\sigma}\cdot\vec{B}}. \tag{6.17}$$

Note that we used

$$\begin{aligned} (\vec{\pi}\times\vec{\pi})^i\,\varphi &= -i\hbar\left(\frac{-e}{c}\right)\epsilon^{ijk}\left(\partial_j A^k - A^k\partial_j\right)\varphi \\ &= i\frac{\hbar e}{c}\epsilon^{ijk}\left(\partial_j A^k\right)\varphi = i\frac{\hbar e}{c}B^i\varphi, \end{aligned}$$

with

$$\boxed{B^i = \epsilon^{ijk}\partial_j A^k}.$$

We introduce now eq. (6.17) into eq. (6.16):

$$\boxed{i\hbar\frac{\partial}{\partial t}\varphi = \left[\frac{1}{2m}\left(\vec{p} - \frac{e}{c}\vec{A}\right)^2 - \frac{e\hbar}{2mc}\vec{\sigma}\cdot\vec{B} + e\Phi\right]\varphi}, \tag{6.18}$$

which is the *Pauli equation* for the Pauli spinor φ as is known from non-relativistic quantum mechanics. The two components of φ describe the spin of the electron.

In this derivation, one obtains the correct gyromagnetic ratio $g = 2$ for the electron.

Proof: We assume a homogeneous magnetic field $\vec{B}$ that can be represented by the vector potential $\vec{A}$:

$$\vec{B} = curl\vec{A}, \qquad \vec{A} = \frac{1}{2}\vec{B}\times\vec{x}.$$

Introducing the orbital angular momentum $\vec{L}$ and spin $\vec{S}$,

$$\vec{L} = \vec{x}\times\vec{p}, \qquad \vec{S} = \frac{1}{2}\hbar\vec{\sigma},$$

into eq. (6.18) and using the fact that

$$\left(\vec{p}\cdot\vec{A}\right) = \frac{\hbar}{i}\left(\vec{\nabla}\cdot\vec{A}\right) = 0$$

is the Coulomb gauge and

$$\begin{aligned}
-\vec{p}\cdot\vec{A} - \vec{A}\cdot\vec{p} &=& -2\vec{A}\cdot\vec{p} = -2\frac{1}{2}\left(\vec{B}\times\vec{x}\right)\cdot\vec{p} \\
&=& -\left(\vec{x}\times\vec{p}\right)\cdot\vec{B} = -\vec{L}\cdot\vec{B},
\end{aligned}$$

we get

$$\boxed{i\hbar\frac{\partial\varphi}{\partial t} = \left(\frac{\vec{p}^2}{2m} - \frac{e}{2mc}(\vec{L} + 2\vec{S})\cdot\vec{B} + \frac{e^2}{2mc^2}\vec{A}^2 + e\Phi\right)\varphi}. \tag{6.19}$$

The eigenvalues of the spin operator $\vec{S}$ are $\pm\hbar/2$. The interaction with the electromagnetic field is, following eq. (6.19),

$$H_{int} = -\vec{\mu}\cdot\vec{B} + \frac{e^2}{2mc^2}\vec{A}^2 + e\Phi,$$

with the magnetic moment

$$\vec{\mu} = \vec{\mu}_{orbit} + \vec{\mu}_{spin} = \frac{e}{2mc}\left(\vec{L} + 2\vec{S}\right). \tag{6.20}$$

From eq. (6.20) it follows that the spin-induced magnetic moment is

$$\boxed{\vec{\mu}_{spin} = g\frac{e}{2mc}\vec{S}} \quad \Rightarrow \quad g = 2 \quad \text{(gyromagnetic, or Landé, factor)}.$$

For the electron,

$$\frac{e}{2mc} = -\frac{\mu_B}{\hbar},$$

where μ_B is the Bohr magneton

$$\mu_B = \frac{e_0\hbar}{2mc} = 0.927 \times 10^{-20} \text{ erg/G}.$$

The non-relativistic limit approximation is good for atoms with small atomic number Z.

Coupling to electromagnetic field. Compact form

Here we will show a more compact way to express eq. (6.14), the coupling of the Dirac equation to the electromagnetic field. For this purpose, we consider the covariant and contravariant forms of the momentum operator,

$$p_\mu = i\hbar\partial_\mu \quad \text{and} \quad p^\mu = i\hbar\partial^\mu,$$

with

$$\partial_\mu = \frac{\partial}{\partial x^\mu} \quad \text{and} \quad \partial^\mu = \frac{\partial}{\partial x_\mu}.$$

Then,

$$p^0 = p_0 = i\hbar\frac{\partial}{\partial(ct)}, \qquad p^1 = -p_1 = i\hbar\frac{\partial}{\partial x_1} = \frac{\hbar}{i}\frac{\partial}{\partial x^1}.$$

Coupling to the electromagnetic field is now realized via

$$p_\mu \to p_\mu - \frac{e}{c}A_\mu,$$

with $A^\mu = (\Phi, \vec{A})$ being the four-potential. This scheme is called *minimal coupling*. This implies

$$i\hbar\frac{\partial}{\partial x^\mu} \to i\hbar\frac{\partial}{\partial x^\mu} - \frac{e}{c}A_\mu$$

or, explicitly,

$$\begin{cases} i\hbar\frac{\partial}{\partial t} & \to \quad i\hbar\frac{\partial}{\partial t} - e\Phi \\ \frac{\hbar}{i}\frac{\partial}{\partial x^i} & \to \quad \frac{\hbar}{i}\frac{\partial}{\partial x^i} + \frac{e}{c}A_i = \frac{\hbar}{i}\frac{\partial}{\partial x^i} - \frac{e}{c}A^i \end{cases}$$

For the spacial components, this is identical to the replacement

$$\frac{\hbar}{i}\vec{\nabla} \to \frac{\hbar}{i}\vec{\nabla} - \frac{e}{c}\vec{A}.$$

Inserting this replacement into the Dirac equation, we obtain

$$\boxed{\left(-\gamma^\mu\left(i\hbar\partial_\mu - \frac{e}{c}A_\mu\right) + mc\right)\Psi = 0} \quad \to \quad \begin{array}{l} \text{Dirac equation in a relativisic covariant} \\ \text{form in the presence of an electromagnetic field} \end{array}$$

6.2.6   Physical interpretation of the solution to the Dirac equation

The Dirac equation possesses negative-energy solutions. The kinetic energy in these states is *negative*, which means that the particle moves in the opposite direction to the one occupying the usual state of positive energy. Therefore, such a particle, carrying the charge of an electron, is repelled by the field of a proton.

One cannot exclude these negative-energy states since the positive-energy states do not represent a complete set of solutions. The physical consequence of these two sets of solutions is that, when an external perturbation (measurement) causes an electron to enter a certain state, this state will be a combination of positive- and negative-energy states. In the case when the electron is confined in a region that is smaller than its Compton wavelength, this interference between positive- and negative- energy components produces an oscillatory motion known as "Zitterbewegung". In order to see that let us solve the Dirac equation in the Heisenberg representation. The Heisenberg operators $O(t)$ fulfill the equation of motion:

$$\frac{dO(t)}{dt} = \frac{1}{i\hbar}\left[O(t), H\right].$$

We consider free particles, and therefore the momentum $\vec{p}$ conmutes with the Hamiltonian

$$\boxed{H = c\vec{\alpha}\cdot\vec{p} + \beta mc^2} \tag{6.21}$$

and therefore

$$\frac{d\vec{p}(t)}{dt} = 0$$

what leads to $\vec{p}(t) = \vec{p} = constant$.
Note also that

$$\vec{v}(t) = \frac{d\vec{x}(t)}{dt} = \frac{1}{i\hbar}\left[\vec{x}(t), H\right] = c\vec{\alpha}(t)$$

and

$$\frac{d\vec{\alpha}}{dt} = \frac{1}{i\hbar}\left[\vec{\alpha}(t), H\right] = \frac{2}{i\hbar}(c\vec{p} - H\vec{\alpha}(t))$$

Since H is time independent, we can integrate the previous equation:

$$\vec{v}(t) = c\vec{\alpha}(t) = cH^{-1}\vec{p} + e^{\frac{2iHt}{\hbar}}(\vec{\alpha}(0) - cH^{-1}\vec{p})$$

and integrating again we obtain:

$$\vec{x}(t) = \vec{x}(0) + \frac{c^2\vec{p}}{H}t + \frac{\hbar c}{2iH}(e^{\frac{2iHt}{\hbar}} - 1)(\vec{\alpha}(0) - \frac{c\vec{p}}{H}) \tag{6.22}$$

Since

$$\vec{\alpha}H + H\vec{\alpha} = 2c\vec{p}$$

then

$$(\vec{\alpha} - \frac{c\vec{p}}{H})H + H(\vec{\alpha} - \frac{c\vec{p}}{H}) = 0 \tag{6.23}$$

We observe that the solution Eq. 6.22 contains a linear term in time which corresponds to the group velocity motion, and an oscillating term which describes the "Zitterbewegung" of the free particle. Since the vanishing of the anticonmutator Eq. 6.23 implies that the energies must be of opposite sign, the "Zitterbewegung" is the result of interference between positive and negative energy states.

Discussion about hole theory in class.

BIBLIOGRAPHY

- J. J. Sakurai: " Modern Quantum Mechanics", Addison-Wesley Publishing Company 1994.

- W. Nolting: "Grundkurs Theoretische Physik 5/2. Quantenmechanik-Methoden und Anwendungen", Springer 2004.

- F. Schwabl: "Advanced Quantum Mechanics", Springer 2000.

- C. Cohen-Tannoudji, B. Diu, F. Laloke: "Quantum Mechanics II", Wiley-VCH 2005.

- M. Tinkham: "Group Theory and Quantum Mechanics", McGraw-Hill, Inc. 1964.

- L. H. Ryder: "Quantum Field Theory", Cambridge University Press , Inc. 1985.