

# Relativistic Quantum Mechanics

## Lecture notes

László Szunyogh and Bendegúz Nyári

27th November 2022

## Contents

|          |                                                                       |           |
|----------|-----------------------------------------------------------------------|-----------|
| <b>1</b> | <b>Special relativity in a nutshell</b>                               | <b>3</b>  |
| 1.1      | Minkowski vectors and Lorentz transformations . . . . .               | 3         |
| 1.2      | Relativistic mechanics . . . . .                                      | 6         |
| 1.3      | Electromagnetic fields . . . . .                                      | 7         |
| <b>2</b> | <b>First steps towards Relativistic Quantum Mechanics</b>             | <b>9</b>  |
| 2.1      | Operators in coordinate representation . . . . .                      | 9         |
| 2.2      | The Klein–Gordon equation . . . . .                                   | 10        |
| <b>3</b> | <b>The Dirac equation</b>                                             | <b>13</b> |
| 3.1      | The Dirac equation for free particles . . . . .                       | 13        |
| 3.1.1    | The Dirac matrices . . . . .                                          | 14        |
| 3.2      | The Dirac equation in the presence of electromagnetic field . . . . . | 16        |
| 3.2.1    | Gauge transformation . . . . .                                        | 16        |
| 3.2.2    | The Dirac Hamiltonian . . . . .                                       | 17        |
| 3.2.3    | The spectrum of free particles and the positron . . . . .             | 18        |
| 3.2.4    | Continuity equation . . . . .                                         | 19        |
| 3.2.5    | Conjugate spinor and the conjugate Dirac equation . . . . .           | 21        |
| 3.3      | Lorentz invariance of the Dirac equation . . . . .                    | 22        |
| 3.3.1    | The four-current density . . . . .                                    | 24        |
| 3.3.2    | Rotations and the spin . . . . .                                      | 25        |
| 3.3.3    | The total angular momentum . . . . .                                  | 26        |
| 3.3.4    | Spatial reflection (parity) . . . . .                                 | 28        |
| 3.4      | Non-relativistic approximation and relativistic corrections . . . . . | 29        |
| 3.4.1    | Relativistic kinetic energy correction . . . . .                      | 30        |

|       |                                                          |    |
|-------|----------------------------------------------------------|----|
| 3.4.2 | Spin-orbit interaction . . . . .                         | 32 |
| 3.4.3 | Darwin term . . . . .                                    | 34 |
| 3.4.4 | Spin magnetization probability current density . . . . . | 37 |

# 1 Special relativity in a nutshell

## 1.1 Minkowski vectors and Lorentz transformations

Let  $\mathbf{e}_\mu$  ( $\mu = 0, 1, 2, 3$ ) be a basis in the four dimensional Minkovski space  $\mathcal{M}$ . Any vector,  $\mathbf{x} \in \mathcal{M}$ , can uniquely be expressed as a linear combination of the basis vectors,

$$\mathbf{x} = x^\mu \mathbf{e}_\mu , \quad (1)$$

where  $x^\mu$  are called the *contravariant coordinates* labelled by upper indices:

$$x = \{x^\mu\} = (x^0, x^1, x^2, x^3) = (ct, \vec{x}) . \quad (2)$$

Based on the invariant Minkovski distance,

$$ds^2 = c^2 dt^2 - d\vec{x}^2 , \quad (3)$$

we define the inner product in the Minkovski space:  $\mathbf{x}, \mathbf{y} \in \mathcal{M}$ ,

$$(\mathbf{x}, \mathbf{y}) = (x^\mu \mathbf{e}_\mu, y^\nu \mathbf{e}_\nu) = x^\mu g_{\mu\nu} y^\nu , \quad (4)$$

with the metric tensor,

$$g_{\mu\nu} \equiv (\mathbf{e}_\mu, \mathbf{e}_\nu) = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} = \text{diag}(1, -1, -1, -1) . \quad (5)$$

Note that the metric tensor is by definition symmetric. Since the metric tensor is not positive definite, the Minkowski space is a pseudo-Euclidean space. The Minkowski space can be decomposed into *timelike* vectors with positive norm, *spacelike* vectors with negative norm and *lighlike* vectors with zero norm represented with the light cone (world-lines).

The inverse of the metric tensor  $g^{\mu\nu}$  is defined as

$$g^{\mu\lambda} g_{\lambda\nu} = \delta^\mu_\nu , \quad (6)$$

$$g_{\mu\lambda} g^{\lambda\nu} = \delta_\mu^\nu , \quad (7)$$

where

$$\delta^\mu_\nu = \delta_\mu^\nu = \begin{cases} 1 & \text{if } \mu = \nu \\ 0 & \text{if } \mu \neq \nu \end{cases} . \quad (8)$$

For the Minkowski metric the metric tensor and its inverse are obviously *numerically equivalent*,

$$g^{\mu\nu} = g_{\mu\nu} . \quad (9)$$

It is convenient to introduce the *covariant coordinates*,

$$x_\cdot = \{x_\mu\} = (x_0, x_1, x_2, x_3) , \quad (10)$$

such that

$$x_\mu \equiv g_{\mu\nu} x^\nu = (ct, -\vec{x}) . \quad (11)$$

The relationship between the contravariant and covariant coordinates can be written briefly as

$$x_{\cdot} = g x^{\cdot} . \quad (12)$$

The inverse relation reads

$$x^{\mu} = \delta^{\mu}_{\cdot\nu} x^{\nu} = g^{\mu\lambda} g_{\lambda\nu} x^{\nu} = g^{\mu\lambda} x_{\lambda} , \quad (13)$$

or

$$x^{\cdot} = g^{-1} x_{\cdot} . \quad (14)$$

The inner product in the Minkovski space can then be simply expressed as

$$(\mathbf{x}, \mathbf{y}) = x^{\mu} y_{\mu} = x_{\mu} y^{\mu} . \quad (15)$$

We demand that the Minkovski distance should be invariant on the choice of the inertial system. Different inertial systems are related to each other by real valued linear transformations,

$$x'^{\mu} = \Lambda^{\mu}_{\cdot\nu} x^{\nu} + a^{\mu} \quad (\Lambda^{\mu}_{\cdot\nu}, a^{\mu} \in \mathbb{R}) \quad (16)$$

or in matrix-vector notation,

$$\Lambda = \{\Lambda^{\mu}_{\cdot\nu}\} \implies x' = \Lambda x + a , \quad (17)$$

called *Poincaré transformations*. Invariance means,

$$dx'^{\mu} g_{\mu\tau} dx'^{\tau} = \Lambda^{\mu}_{\cdot\nu} dx^{\nu} g_{\mu\tau} \Lambda^{\tau}_{\cdot\lambda} dx^{\lambda} = dx^{\nu} \Lambda^{\mu}_{\cdot\nu} g_{\mu\tau} \Lambda^{\tau}_{\cdot\lambda} dx^{\lambda} = dx^{\nu} g_{\nu\lambda} dy^{\lambda} , \quad (18)$$

implying the identity

$$\Lambda^{\mu}_{\cdot\nu} g_{\mu\tau} \Lambda^{\tau}_{\cdot\lambda} = g_{\nu\lambda} . \quad (19)$$

Introducing the transpose of  $\Lambda$ ,

$$(\Lambda^T)^{\cdot\mu}_{\nu} \equiv \Lambda^{\mu}_{\cdot\nu} , \quad (20)$$

the previous identity can be written as

$$(\Lambda^T)^{\cdot\mu}_{\nu} g_{\mu\tau} \Lambda^{\tau}_{\cdot\lambda} = g_{\nu\lambda} . \quad (21)$$

The above derivation can be repeated compactly as,

$$dx'^{\cdot} g dx'^{\cdot} = \Lambda dx^{\cdot} g \Lambda dx^{\cdot} = dx^{\cdot} \Lambda^T g \Lambda dx^{\cdot} \rightarrow \Lambda^T g \Lambda = g . \quad (22)$$

The  $\Lambda$  matrices, i.e. the solutions of Eq. (22) have the following properties:

1.  $\Lambda^{\mu}_{\cdot\nu} = \delta^{\mu}_{\cdot\nu}$  is a solution:  $\delta^{\cdot\tau}_{\mu} g_{\tau\sigma} \delta^{\sigma}_{\cdot\nu} = g_{\mu\nu}$
2.  $\det \Lambda^T g \Lambda = (\det \Lambda)^2 \det g = \det g \implies \det \Lambda = \pm 1$ . This implies that any solution is invertible.
3. If  $\Lambda$  is a solution, then  $\Lambda^{-1}$  is also solution:  $g = (\Lambda^T)^{-1} g \Lambda^{-1}$
4. If  $\Lambda_1$  and  $\Lambda_2$  are solutions, then  $\Lambda_1 \Lambda_2$  is solution:  $(\Lambda_1 \Lambda_2)^T g \Lambda_1 \Lambda_2 = \Lambda_2^T \Lambda_1^T g \Lambda_1 \Lambda_2 = \Lambda_2^T g \Lambda_2 = g$

From the above properties it follows that the real valued matrices  $\Lambda$  satisfying Eq. (22) form a group (Lorentz group). As what follows we deal with the *Lorentz transformations* for which  $a^\mu = 0$ .

The infinitesimal Lorentz transformation  $\omega$  is defined as

$$\delta_{\cdot\nu}^\mu + \omega_{\cdot\nu}^\mu \leftrightarrow I + \omega \quad (|\omega| \ll 1). \quad (23)$$

Eq. (22) then implies

$$(I + \omega^T) g (I + \omega) = g \implies \omega^T g = -g \omega. \quad (24)$$

Since

$$(\omega^T)_{\mu}^{\cdot\tau} g_{\tau\nu} = \omega_{\cdot\mu}^{\tau} g_{\tau\nu} = g_{\nu\tau} \omega_{\cdot\mu}^{\tau} = \omega_{\nu\mu} \quad \text{and} \quad g_{\mu\tau} \omega_{\cdot\nu}^{\tau} = \omega_{\mu\nu} \quad (25)$$

$$\Downarrow$$

$$\omega_{\mu\nu} = -\omega_{\nu\mu}, \quad (26)$$

i.e., the matrix  $\omega_{\mu\nu}$  is antisymmetric.

*Statement:* The matrix

$$e^\omega = \sum_n \frac{\omega^n}{n!} \quad (27)$$

is solution of Eq. (22).

*Proof:* First we prove that

$$(\omega^n)^T g = (-1)^n g \omega^n, \quad (28)$$

which is satisfied for  $n = 1$ . Suppose that  $(\omega^{n-1})^T g = (-1)^{n-1} g \omega^{n-1}$ . Then

$$(\omega^n)^T g = (\omega^T)^n g = (\omega^T)^{n-1} \omega^T g = -(\omega^T)^{n-1} g \omega = (-1)^n g \omega^n. \quad (29)$$

We thus conclude that

$$(e^\omega)^T g = \sum_n \frac{(\omega^n)^T g}{n!} = g \sum_n \frac{(-\omega)^n}{n!} = g e^{-\omega} \rightarrow (e^\omega)^T g e^\omega = g. \quad (30)$$

The Lorentz transform of the covariant vectors can be obtained as

$$x'_\mu = g_{\mu\nu} x'^\nu = g_{\mu\nu} \Lambda_{\cdot\tau}^\nu x^\tau = g_{\mu\nu} \Lambda_{\cdot\tau}^\nu g^{\tau\sigma} x_\sigma = \bar{\Lambda}_\mu^{\cdot\sigma} x_\sigma, \quad (31)$$

where

$$\bar{\Lambda}_\mu^{\cdot\sigma} \equiv g_{\mu\nu} \Lambda_{\cdot\tau}^\nu g^{\tau\sigma} \quad (32)$$

or

$$x' = g x'^\cdot = g \Lambda x^\cdot = g \Lambda g^{-1} x^\cdot \rightarrow \bar{\Lambda} = g \Lambda g^{-1}. \quad (33)$$

The matrix of the Lorentz transformation for the covariant vectors  $\bar{\Lambda}$  can even more directly be related to  $\Lambda$ :

$$(\Lambda^T)_{\mu}^{\cdot\tau} \bar{\Lambda}_\tau^{\cdot\nu} = \Lambda_{\cdot\mu}^\tau g_{\tau\sigma} \Lambda_{\cdot\rho}^\sigma g^{\rho\nu} = g_{\mu\rho} g^{\rho\nu} = \delta_{\mu}^{\cdot\nu}, \quad (34)$$

i.e.

$$\Lambda^T \bar{\Lambda} = \Lambda^T g \Lambda g^{-1} = g g^{-1} = I, \quad (35)$$

from which we conclude

$$\bar{\Lambda} = (\Lambda^{-1})^T. \quad (36)$$

We can show that the four-derivative,

$$\partial_\mu = \frac{\partial}{\partial x^\mu} \Rightarrow \partial_\mu x^\nu = \delta_\mu^\nu \quad (37)$$

$$\partial_\mu = \left( \frac{1}{c} \partial_t, \vec{\nabla} \right) \quad (38)$$

is a covariant vector:

$$x'^\mu = \Lambda^\mu_{\cdot\nu} x^\nu \rightarrow x^\nu = (\Lambda^{-1})^\nu_{\cdot\mu} x'^\mu \quad (39)$$

$\Downarrow$

$$\partial'_\mu = \frac{\partial}{\partial x'^\mu} = \frac{\partial}{\partial x^\nu} \frac{\partial x^\nu}{\partial x'^\mu} = \partial_\nu (\Lambda^{-1})^\nu_{\cdot\mu} = \left[ (\Lambda^{-1})^T \right]^\nu_{\cdot\mu} \partial_\nu = \bar{\Lambda}^\nu_{\cdot\mu} \partial_\nu, \quad (40)$$

or

$$x^\cdot = \Lambda^{-1} x'^\cdot \rightarrow \frac{\partial x^\cdot}{\partial x'^\cdot} = \Lambda^{-1} \rightarrow \partial'_\cdot = \partial_\cdot \frac{\partial x^\cdot}{\partial x'^\cdot} = \partial_\cdot \Lambda^{-1} = (\Lambda^{-1})^T \partial_\cdot = \bar{\Lambda} \partial_\cdot. \quad (41)$$

The contravariant derivative can be expressed as

$$\partial^\mu = \left( \frac{1}{c} \partial_t, -\vec{\nabla} \right), \quad (42)$$

while the d'Alembert operator,

$$\square \equiv \partial_\mu \partial^\mu = \frac{1}{c^2} \partial_t^2 - \Delta, \quad (43)$$

is invariant under Lorentz transformations.

## 1.2 Relativistic mechanics

The four-velocity is defined as

$$u^\mu = \frac{dx^\mu}{d\tau} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} (c, \vec{v}), \quad (44)$$

where  $\tau$  is the proper time,

$$d\tau = \frac{ds}{c} = dt \sqrt{1 - \frac{v^2}{c^2}}, \quad (45)$$

with

$$\vec{v} = \frac{d\vec{x}(t)}{dt}. \quad (46)$$

It is easy to see that the norm of the four-velocity is Lorentz invariant,

$$u^\mu g_{\mu\nu} u^\nu = u^\mu u_\mu = \frac{c^2 - v^2}{1 - \frac{v^2}{c^2}} = c^2. \quad (47)$$

The four-momentum is defined as

$$p^\mu = m u^\mu = \left( \frac{E}{c}, \vec{p} \right), \quad (48)$$

where  $m$  is the rest mass of the particle. The momentum and the energy are then given by

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (49)$$

and

$$E = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} , \quad (50)$$

respectively. Obviously, the length of the four-momentum is Lorentz invariant,

$$p^\mu p_\mu = m^2 c^2 , \quad (51)$$

which can also be expressed as the relativistic energy-momentum dispersion relation,

$$E^2 = (mc^2)^2 + c^2 \vec{p}^2 . \quad (52)$$

### 1.3 Electromagnetic fields

The contravariant four-potential is composed of the scalar and vector potentials as,

$$A^\mu = \left( \frac{\phi}{c}, \vec{A} \right) , \quad (53)$$

The four-current is defined as

$$j^\mu = (c\rho, \vec{j}) , \quad (54)$$

satisfying the Lorentz invariant continuity equation,

$$\partial_\mu j^\mu = 0 . \quad (55)$$

which is equivalent with the usual form,

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0 . \quad (56)$$

Introducing the field strength tensor,

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu , \quad (57)$$

$$F^{i0} = \frac{\mathcal{E}_i}{c} , \quad F^{ij} = -\epsilon_{ijk} B_k , \quad (58)$$

with  $\vec{\mathcal{E}}$  and  $\vec{B}$  being the electric and magnetic fields, respectively, the Maxwell equations can be summarized as

$$\partial_\mu F^{\mu\nu} = \mu_0 j^\nu , \quad (59)$$

$$\epsilon_{\lambda\tau\mu\nu} \partial^\tau F^{\mu\nu} = 0 , \quad (60)$$

where  $\epsilon_{\lambda\tau\mu\nu}$  is the totally antisymmetric four-index tensor with the normalization  $\epsilon_{0123} = 1$ .

In the presence of electromagnetic fields the four velocity is related to the kinetic momentum rather than to the canonical momentum (*minimal substitution*),

$$p^\mu \rightarrow K^\mu = p^\mu - qA^\mu = mu^\mu \quad (61)$$

$\Downarrow$

$$K^\mu = \left( \frac{E - q\phi}{c}, \vec{p} - q\vec{A} \right) \quad (62)$$

$$K_\mu = \left( \frac{E - q\phi}{c}, -(\vec{p} - q\vec{A}) \right), \quad (63)$$

the length of which is Lorentz invariant,

$$K^\mu K_\mu = m^2 u^\mu u_\mu = m^2 c^2. \quad (64)$$

This expression is equivalent with the energy-momentum dispersion:

$$\frac{(E - q\phi)^2}{c^2} - (\vec{p} - q\vec{A})^2 = m^2 c^2 \quad (65)$$

$\Downarrow$

$$E = \sqrt{m^2 c^4 + c^2 (\vec{p} - q\vec{A})^2} + q\phi. \quad (66)$$

## 2 First steps towards Relativistic Quantum Mechanics

### 2.1 Operators in coordinate representation

Relativistic quantum mechanics should provide with a Lorentz invariant wave equation reflecting the dispersion relation (52). Moreover, the axioms of quantum mechanics imposed for the wavefunction and the operators (e.g. probabilistic interpretation, commutation relations) should be preserved.

Within nonrelativistic quantum mechanics the coordinate and momentum operators are defined in coordinate representation as

$$x_i \rightarrow x_i \cdot \quad p_i \rightarrow \frac{\hbar}{i} \partial_i = -i\hbar \partial_i. \quad (67)$$

Since  $x_i$  and  $-\partial_i$  are the spatial coordinates of the contravariant space-time vector,  $x^\mu$ , and the contravariant four-derivative  $\partial^\mu$ , respectively, it makes sense to extend the above definitions as

$$x^\mu \rightarrow x^\mu \cdot \quad p^\mu \rightarrow i\hbar \partial^\mu = \left( \frac{i\hbar \partial_t}{c}, -i\hbar \vec{\nabla} \right) = (p^0, \vec{p}) . \quad (68)$$

The corresponding covariant four-vector operators are

$$x_\mu \rightarrow x_\mu \cdot \quad p_\mu \rightarrow i\hbar \partial_\mu = \left( \frac{i\hbar \partial_t}{c}, i\hbar \vec{\nabla} \right) = (p_0, -\vec{p}) . \quad (69)$$

Obviously, the following commutation relations apply,

$$[p_\mu, x^\nu] = i\hbar [\partial_\mu, x^\nu] = i\hbar \left[ \frac{\partial}{\partial x^\mu}, x^\nu \right] = i\hbar \delta_\mu^\nu, \quad (70)$$

which can be transformed into covariant form,

$$[p^\mu, x^\nu] = i\hbar g^{\mu\nu}, \quad (71)$$

or

$$[p_\mu, x_\nu] = i\hbar g_{\mu\nu}, \quad (72)$$

Obviously, the above expressions are a proper extension of the commutators postulated in non-relativistic quantum mechanics,

$$[p_i, x_j] = \frac{\hbar}{i} \delta_{ij}. \quad (73)$$

As an excersize, let's prove that the commutation relations (71) are Lorentz invariant:

$$C^{\mu\nu} \equiv [p^\mu, x^\nu] \rightarrow \mathbb{C} = i\hbar g^{-1} \quad (74)$$

$$C'^{\mu\nu} = [p'^\mu, x'^\nu] = \Lambda_{\cdot\tau}^\mu \Lambda_{\cdot\lambda}^\nu [p^\tau, x^\lambda] = \Lambda_{\cdot\tau}^\mu [p^\tau, x^\lambda] (\Lambda^T)_{\lambda}^{\cdot\nu} = \Lambda_{\cdot\tau}^\mu C^{\tau\lambda} (\Lambda^T)_{\lambda}^{\cdot\nu}, \quad (75)$$

consequently,

$$\begin{aligned} \Lambda^T g \Lambda &= g \rightarrow g^{-1} \Lambda^T = \Lambda^{-1} g^{-1} \\ C' &= \Lambda C \Lambda^T = i\hbar \Lambda g^{-1} \Lambda^T = i\hbar \Lambda \Lambda^{-1} g^{-1} = i\hbar g^{-1}, \end{aligned} \quad (76)$$

where we used Eq. (22).

We can easily verify the commutator of the timelike coordinates,

$$[p^0, x^0] = i\hbar [\partial_t, t] = i\hbar. \quad (77)$$

Given that for a relativistic particle  $p_0 = p^0 = \frac{E}{c}$ , the energy should be associated with the derivation with respect to time,

$$E = i\hbar \frac{\partial}{\partial t}. \quad (78)$$

The kinetic momentum operators can be expressed in coordinate representation as follows

$$K^\mu = p^\mu - qA^\mu = i\hbar\partial^\mu - qA^\mu = \left( \frac{1}{c} (i\hbar\partial_t - q\phi), -i\hbar\vec{\nabla} - q\vec{A} \right), \quad (79)$$

$$K_\mu = p_\mu - qA_\mu = i\hbar\partial_\mu - qA_\mu = \left( \frac{1}{c} (i\hbar\partial_t - q\phi), i\hbar\vec{\nabla} + q\vec{A} \right). \quad (80)$$

## 2.2 The Klein–Gordon equation

If we substitute the operators (79) and (80) into the Lorentz invariant expression (64) and let this operator act on the wavefunction, we arrive at the *Klein–Gordon equation*,

$$[K^\mu K_\mu - m^2 c^2] \psi(\vec{r}, t) = 0 \quad (81)$$

$\Downarrow$

$$\left[ \frac{1}{c^2} (i\hbar\partial_t - q\phi)^2 - \left( i\hbar\vec{\nabla} + q\vec{A} \right)^2 \right] \psi(\vec{r}, t) = m^2 c^2 \psi(\vec{r}, t). \quad (82)$$

In case of  $\vec{A}(\vec{r}, t) = 0$  and  $\phi(\vec{r}, t) = 0$ , i.e. for a free particle, the above equation reduces to,

$$[\square + \kappa^2] \psi(\vec{r}, t) = 0 \quad (83)$$

where  $\kappa = \frac{mc}{\hbar}$  is the Compton wavelength and  $\square = \frac{1}{c^2} \partial_t^2 - \Delta$  is the d’Alambert operator as we noted already.

For time independent vector and scalar potentials, the time dependence of the wave function can be separated in the usual way,

$$\psi(\vec{r}, t) = \psi(\vec{r}) e^{-\frac{i}{\hbar} Et}, \quad (84)$$

leading to the *stationary Klein–Gordon equation*,

$$\left[ \left( i\hbar\vec{\nabla} + q\vec{A} \right)^2 - \frac{(E - q\phi)^2}{c^2} + m^2 c^2 \right] \psi(\vec{r}) = 0 \quad . \quad (85)$$

For a free particle, the solution of the equation,

$$\left[ -\hbar^2 \Delta - \frac{E^2}{c^2} + m^2 c^2 \right] \psi(\vec{r}) = 0, \quad (86)$$

is a plane wave,

$$\psi(\vec{r}) = A e^{\frac{i}{\hbar} \vec{p} \cdot \vec{r}} \quad (87)$$

which by substitution yields

$$c^2 p^2 + m^2 c^4 - E^2 = 0 , \quad (88)$$

in agreement with Eq. (52). The energy of the free particle with momentum  $\vec{p}$  can then take

$$E = \pm \sqrt{m^2 c^4 + c^2 p^2} \quad (89)$$

allowing for  $E \geq mc^2$  or  $E \leq -mc^2$ . The appearance of the negative energy solutions is a new feature as compared to the classical relativistic mechanics. The negative energy solutions should not be overlooked, since the positive and the negative energy solutions together form a complete set of the Hilbert space.

The stationary Klein-Gordon equation can be solved for the H atom ( $\vec{A}(\vec{r}) = 0$ ,  $\phi(\vec{r}) = -Zke^2/r$ ). Expanding the eigenenergies up to first order in  $1/c^2$ , we obtain

$$E_{n\ell} \simeq mc^2 - \frac{m(Zke^2)^2}{2\hbar^2 n^2} + \frac{m(Zke^2)^4}{4\hbar^4 n^4 c^2} \left( \frac{3}{2} - \frac{4n}{2\ell+1} \right) + \dots , \quad (90)$$

where  $n$  and  $\ell$  are the well known principal and orbital quantum numbers of the non-relativistic spectrum of the H atom. Clearly, the second term on the righthand side is identical with the the nonrelativistic eigenenergies that accurately reproduce the main structure of the emission spectrum of the H atom (Lyman, Balmer, Paschen ... series), but the third term is inconsistent with the fine structure of the measured spectrum.

There occurs an even more serious problem if we derive the continuity equation needed for the probabilistic interpretation. Let's start with the covariant form of the Klein-Gordon equation (81):

$$\begin{aligned} & \left( (\partial_\mu - \frac{qA_\mu}{i\hbar})(\partial^\mu - \frac{qA^\mu}{i\hbar}) + \kappa^2 \right) \psi = \\ & \left( \partial_\mu \partial^\mu - \frac{q}{i\hbar} (\partial_\mu A^\mu + A^\mu \partial_\mu) - \frac{q^2}{\hbar^2} A_\mu A^\mu + \kappa^2 \right) \psi = 0 . \end{aligned} \quad (91)$$

Conjugating the above equation yields

$$\left( \partial_\mu \partial^\mu + \frac{q}{i\hbar} (\partial_\mu A^\mu + A^\mu \partial_\mu) - \frac{q^2}{\hbar^2} A_\mu A^\mu + \kappa^2 \right) \psi^* = 0 , \quad (92)$$

while combining the two equations we obtain

$$\frac{1}{c^2} (\psi^* \partial_\mu \partial^\mu \psi - \psi \partial_\mu \partial^\mu \psi^*) - \frac{q}{i\hbar} (\psi^* \partial_\mu A^\mu \psi + \psi^* A^\mu \partial_\mu \psi + \psi \partial_\mu A^\mu \psi^* + \psi A^\mu \partial_\mu \psi^*) = 0 . \quad (93)$$

By using the below manipulations,

$$\partial_\mu (\psi^* \partial^\mu \psi - \psi \partial^\mu \psi^*) = \underbrace{(\partial_\mu \psi^*) (\partial^\mu \psi) - (\partial_\mu \psi) (\partial^\mu \psi^*)}_{=0} + (\psi^* \partial_\mu \partial^\mu \psi - \psi \partial_\mu \partial^\mu \psi^*) ,$$

$$\begin{aligned} & \psi^* \partial_\mu A^\mu \psi + \psi^* A^\mu \partial_\mu \psi + \psi \partial_\mu A^\mu \psi^* + \psi A^\mu \partial_\mu \psi^* \\ & = 2 (\partial_\mu A^\mu) \psi^* \psi + 2 \psi^* A^\mu \partial_\mu \psi + 2 \psi A^\mu \partial_\mu \psi^* \\ & = 2 \partial_\mu (A^\mu \psi^* \psi) , \end{aligned}$$

and multiplying the equation by  $\frac{i\hbar}{2m}$ , we arrive at the Lorentz invariant continuity equation,

$$\partial_\mu j^\mu = 0, \quad (94)$$

with the probability four-current density,

$$\begin{aligned} j^\mu &= \frac{i\hbar}{2m} (\psi^* \partial^\mu \psi - \psi \partial^\mu \psi^*) - \frac{q}{m} A^\mu \psi^* \psi \\ &= \frac{1}{2m} (\psi^* p^\mu \psi - \psi p^\mu \psi^*) - \frac{q}{m} A^\mu \psi^* \psi \\ &= \frac{1}{m} \text{Re} (\psi^* p^\mu \psi) - \frac{q}{m} A^\mu \psi^* \psi, \end{aligned} \quad (95)$$

which is the contravariant extension of the probability current density we obtained in non-relativistic Quantum Mechanics. While this seems to be a reassuring result, the problem is that the finding probability density,

$$\rho = \frac{j^0}{c} = \frac{i\hbar}{2mc^2} (\psi^* \partial_t \psi - \psi \partial_t \psi^*) - \frac{q\phi}{mc^2} \psi^* \psi, \quad (96)$$

is not positive definite. In case of time independent potentials with the stationary wavefunction (84),

$$\rho(\vec{r}) = \frac{E - q\phi(\vec{r})}{mc^2} |\psi(\vec{r})|^2. \quad (97)$$

This obviously gives a negative probability density for free particles with negative energy,  $E = -\sqrt{m^2 c^4 + p^2 c^2}$ . The fundamental problem arises from the fact that the probability density (96) contains not only the wavefunction  $\psi(\vec{r}, t)$  but also its time derivative  $\partial_t \psi(\vec{r}, t)$  which is a consequence that the Klein-Gordon equation is of second order in the time derivative.

Furthermore, in the Klein-Gordon equation we deal with a scalar wavefunction what excludes the interpretation of the spin. This, however, doesn't mean that the Klein-Gordon equation has no significance: within quantum field theory it provides the field equation of the zero spin particles like  $\pi$  mesons.

### 3 The Dirac equation

#### 3.1 The Dirac equation for free particles

We have seen that the probabilistic interpretation fails within the framework of the Klein–Gordon equation because it contains the second time derivative of the wavefunction. As proposed by Dirac (1928), the wave equation of free particles should be written as a linear form of the four-momentum  $p_\mu$ ,

$$(\gamma^\mu p_\mu - M) \psi = 0 , \quad (98)$$

where  $M$  is a scalar and the  $\gamma^\mu$  operators commute with the  $p_\mu$  operators. Consequently, the  $\psi$  wavefunction can be interpreted as the element of a tensor product Hilbert space composed of the Hilbert space of the square integrable functions where  $p_\mu$  act and another Hilbert space where the  $\gamma^\mu$  act.

We demand that the solutions of Eq. (98) should also be the solutions of the Klein-Gordon equation, which for free particles leads to the required relativistic energy-momentum dispersion. To this end, we act on equation (98) with the operator  $\gamma^\nu p_\nu + M$ ,

$$(\gamma^\nu p_\nu + M) (\gamma^\mu p_\mu - M) \psi = (\gamma^\nu \gamma^\mu p_\nu p_\mu - M^2) \psi = 0 . \quad (99)$$

Making use that the operators  $p_\mu$  commute with each other,

$$\gamma^\nu \gamma^\mu p_\mu p_\nu = \frac{1}{2} (\gamma^\nu \gamma^\mu p_\mu p_\nu + \gamma^\nu \gamma^\mu p_\nu p_\mu) \quad (100)$$

$$= \frac{1}{2} (\gamma^\nu \gamma^\mu + \gamma^\mu \gamma^\nu) p_\mu p_\nu . \quad (101)$$

Thus, the Klein-Gordon equation is recovered by fixing the constant  $M$  as

$$M = mc \quad (102)$$

and by demanding

$$\frac{1}{2} (\gamma^\nu \gamma^\mu + \gamma^\mu \gamma^\nu) p_\mu p_\nu = p^\mu p_\mu . \quad (103)$$

The latter condition implies the following anticommutator relations for  $\gamma^\mu$ ,

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu} , \quad (104)$$

as written out in details,

$$(\gamma^0)^2 = I \quad (\gamma^i)^2 = -I \quad (105)$$

$$\begin{aligned} \gamma^0 \gamma^i + \gamma^i \gamma^0 &= 0 & \gamma^i \gamma^j + \gamma^j \gamma^i &= -2\delta_{ij} \\ (i, j &= 1, 2, 3) . \end{aligned} \quad (106)$$

Later on we will see that in order to define a hermitian Hamilton operator the  $\gamma^\mu$  operators should possess the below properties,

$$(\gamma^0)^\dagger = \gamma^0 \quad \text{and} \quad (\gamma^0 \gamma^i)^\dagger = \gamma^0 \gamma^i . \quad (107)$$

From these relationships we deduce

$$\gamma^0 \gamma^i = (\gamma^i)^\dagger \gamma^0 \implies (\gamma^i)^\dagger = \gamma^0 \gamma^i \gamma^0 . \quad (108)$$

Since  $(\gamma^0)^2 = I$ , we conclude that in general

$$(\gamma^\mu)^+ = \gamma^0 \gamma^\mu \gamma^0. \quad (109)$$

Substituting  $p_\mu = i\hbar\partial_\mu$  into Eq. (98) we get Dirac's famous equation for free particles,

$$(i\hbar\gamma^\mu\partial_\mu - mc)\psi = 0 \quad (110)$$

$\Downarrow$

$$(\gamma^\mu\partial_\mu + i\kappa)\psi = 0, \quad (111)$$

where (as in the Klein–Gordon equation for a free particle)  $\kappa = \frac{mc}{\hbar}$  is the Compton wavelength.

### 3.1.1 The Dirac matrices

*Statement:* For the  $n$ -dimensional matrix representations of the operators  $\gamma^\mu$ ,  $n$  must be even.

*Proof:*

(i) First we show that the trace of the operators  $\gamma^\mu$  is zero. Namely,

$$\text{Tr}\gamma^0 = -\text{Tr}[\gamma^0 (\gamma^i)^2] = -\text{Tr}[\gamma^i \gamma^0 \gamma^i] = \text{Tr}[(\gamma^i)^2 \gamma^0] = -\text{Tr}\gamma^0, \quad (112)$$

$$\text{Tr}\gamma^j = \text{Tr}[\gamma^j (\gamma^0)^2] = \text{Tr}[\gamma^0 \gamma^j \gamma^0] = -\text{Tr}[(\gamma^0)^2 \gamma^j] = -\text{Tr}\gamma^j, \quad (113)$$

where we made use of the anticommutation relations (104) and the invariance of the trace under cyclic permutations of the operators.

(ii) Since  $(\gamma^0)^2 = 1$ , the eigenvalues of  $\gamma^0$  are  $\pm 1$ , while because of  $(\gamma^j)^2 = -1$  the eigenvalues of  $\gamma^j$  are  $\pm i$ .

Let us suppose that there are  $m$  eigenvalues of 1 (or  $i$ ) and  $n - m$  eigenvalues are  $-1$  (or  $-i$ ). Then,

$$\text{Tr}\gamma^0 = m - (n - m) = 2m - n = 0, \quad (114)$$

$$\text{Tr}\gamma^j = [m - (n - m)]i = (2m - n)i = 0, \quad (115)$$

from which it follows that  $n$  is an even integer.

*Statement:* In the space of  $2 \times 2$  matrices there cannot exist four linearly independent  $\gamma^\mu$  that satisfy the anticommutator relations (104).

*Proof:* Based on the anticommutator relations of the Pauli matrices,

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij},$$

it is easy to show that

$$\gamma^0 = \sigma_x, \gamma^1 = i\sigma_y, \gamma^2 = i\sigma_z,$$

satisfy (104). Any  $2 \times 2$  matrix like  $\gamma^3$  can be written as

$$\gamma^3 = aI_2 + \sum_{\mu=0,1,2} b_\mu \gamma^\mu,$$

where  $I_2$  is the unit matrix and  $a, b_\mu \in \mathbb{C}$ . The anticommutator of  $\gamma^\nu$  ( $\nu = 0, 1, 2$ ) and  $\gamma^3$ ,

$$\{\gamma^\nu, \gamma^3\} = 2a\gamma^\nu + \sum_{\mu=0,1,2} \{\gamma^\nu, \gamma^\mu\} b_\mu = 2a\gamma^\nu + (2 \sum_{\mu=0,1,2} g^{\nu\mu} b_\mu) I_2,$$

must vanish, which implies

$$a = 0$$

and

$$\sum_{\mu=0,1,2} g^{\nu\mu} b_{\mu} = 0.$$

Since the metric tensor is invertible, the last condition is fulfilled only for  $b_{\mu} = 0$ . This means that necessarily  $\gamma^3 = 0$ , which is not a solution, since  $(\gamma^3)^2 = 0 \neq -I_2$ .

It is easy to prove that the operators (or matrix representations)  $\gamma^{\mu}$  ( $\mu = 0, 1, 2, 3$ ) generate a group upon multiplying them in arbitrary way. This is the Dirac group which has 32 elements. It can be shown that the Dirac group has 17 irreducible representations: either 16 one-dimensional and one four-dimensional or 12 one-dimensional and 5 two-dimensional irreducible representations. Since, as we have shown above, the Dirac group has no two-dimensional representations, only the first case applies. Indeed, the four-dimensional representations satisfy the relationships (104) for  $\mu = 0, 1, 2, 3$ . Thus, as what follows we will represent the operators  $\gamma^{\mu}$  by  $4 \times 4$  matrices.

In particular, we will use the so-called *standard (Dirac) representation*,

$$\gamma^0 = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix} \quad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}, \quad (116)$$

where  $I_2$  is the  $2 \times 2$  unit matrix and  $\sigma_i$  ( $i = 1, 2, 3$ ) are the Pauli matrices. We can easily check that the required anticommutator relations apply:

$$\{\gamma^i, \gamma^k\} = - \begin{pmatrix} \{\sigma_i, \sigma_k\} & 0 \\ 0 & \{\sigma_i, \sigma_k\} \end{pmatrix} = -2\delta_{ik} I_4 = 2g^{ik} I_4, \quad (117)$$

$$\{\gamma^0, \gamma^i\} = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} + \begin{pmatrix} 0 & -\sigma_i \\ -\sigma_i & 0 \end{pmatrix} = 0, \quad (118)$$

and  $(\gamma^0)^2 = I_4$ . It is worth noting that the Dirac matrix  $\gamma^5 = \gamma^0 \gamma^1 \gamma^2 \gamma^3$ ,

$$\gamma^5 = \underbrace{\begin{pmatrix} 1_2 & 0 \\ 0 & -1_2 \end{pmatrix} \begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix}}_{\begin{pmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{pmatrix}} \underbrace{\begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{pmatrix}}_{\begin{pmatrix} -i\sigma_1 & 0 \\ 0 & -i\sigma_1 \end{pmatrix}} = -i \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix}. \quad (119)$$

anticommutes with the  $\gamma^{\mu}$  matrices ( $\mu = 0, 1, 2, 3$ ). According to Eq. (104), e.g.

$$\begin{aligned} \gamma^0 \gamma^5 &= \gamma^1 \gamma^2 \gamma^3 = \gamma^1 \gamma^2 \gamma^3 \gamma^0 \gamma^0 = -\gamma^1 \gamma^2 \gamma^0 \gamma^3 \gamma^0 \\ &= \gamma^1 \gamma^0 \gamma^2 \gamma^3 \gamma^0 = -\gamma^0 \gamma^1 \gamma^2 \gamma^3 \gamma^0 = -\gamma^5 \gamma^0, \end{aligned} \quad (120)$$

while

$$\begin{aligned} (\gamma^5)^2 &= \gamma^0 \gamma^1 \gamma^2 \gamma^3 \gamma^0 \gamma^1 \gamma^2 \gamma^3 = -\gamma^1 \gamma^2 \gamma^3 \gamma^1 \gamma^2 \gamma^3 \\ &= \gamma^2 \gamma^3 \gamma^2 \gamma^3 = \gamma^3 \gamma^3 = -I. \end{aligned} \quad (121)$$

Consequently, within Dirac's theory the states of the relativistic particles are described by *four-component wavefunctions*,

$$\psi(\vec{r}, t) = \begin{pmatrix} \psi_1(\vec{r}, t) \\ \psi_2(\vec{r}, t) \\ \psi_3(\vec{r}, t) \\ \psi_4(\vec{r}, t) \end{pmatrix}, \quad (122)$$

which we call *Dirac spinors*.

### 3.2 The Dirac equation in the presence of electromagnetic field

Let's postulate that in the presence of electromagnetic field the canonical momentum  $p_\mu$  should be substituted by the kinetic momentum  $K_\mu$  (*minimal coupling/substitution*),

$$(\gamma^\mu K_\mu - mc) \psi = 0 \quad (123)$$

where

$$K_\mu = p_\mu - qA_\mu = i\hbar\partial_\mu - qA_\mu \quad (124)$$

$\Downarrow$

$$(\gamma^\mu (i\hbar\partial_\mu - qA_\mu) - mc) \psi = 0 \quad (125)$$

or

$$\left( \gamma^\mu \left( \partial_\mu + \frac{iq}{\hbar} A_\mu \right) + i\kappa \right) \psi = 0, \quad (126)$$

where  $A_\mu = (\frac{q\phi}{c}, -\vec{A})$  is the covariant four-potential.

#### 3.2.1 Gauge transformation

We consider the *gauge transformation*,

$$\vec{A}' = \vec{A} + \vec{\nabla}\Lambda, \quad \phi' = \phi - \partial_t\Lambda \Rightarrow A'_\mu = \left( \frac{\phi - \partial_t\Lambda}{c}, -\vec{A} - \vec{\nabla}\Lambda \right) = A_\mu - \partial_\mu\Lambda. \quad (127)$$

The Dirac equation (126) then transforms to

$$\left( \gamma^\mu \left( \partial_\mu + \frac{iq}{\hbar} A'_\mu \right) + i\kappa \right) \psi' = 0 \quad (128)$$

$\Downarrow$

$$\left[ \gamma_\mu \left( \partial_\mu + \frac{iq}{\hbar} A_\mu - \frac{iq}{\hbar} \partial_\mu\Lambda \right) + i\kappa \right] \psi' = 0. \quad (129)$$

It is easy to prove that, as in the nonrelativistic case, the solution of the above equation is

$$\psi' = \psi e^{\frac{iq}{\hbar}\Lambda}. \quad (130)$$

This relies on the below equalities,

$$\begin{aligned} \partial_\mu(\psi e^{\frac{iq}{\hbar}\Lambda}) &= e^{\frac{iq}{\hbar}\Lambda} \partial_\mu\psi + \frac{iq}{\hbar} (\partial_\mu\Lambda) \psi e^{\frac{iq}{\hbar}\Lambda} \\ &= e^{\frac{iq}{\hbar}\Lambda} \left( \partial_\mu\psi + \frac{iq}{\hbar} (\partial_\mu\Lambda) \psi \right), \end{aligned} \quad (131)$$

which should be substituted into (129). The term  $-\frac{iq}{\hbar}(\partial_\mu\Lambda)\psi$  is then cancelled and dividing the equation by the factor  $e^{\frac{iq}{\hbar}\Lambda}$  we obtain equation (126).

### 3.2.2 The Dirac Hamiltonian

Substituting the coordinate representation of  $K_\mu$ ,

$$K_\mu = \left( \frac{1}{c} (i\hbar\partial_t - q\phi), -(\vec{p} - q\vec{A}) \right), \quad (132)$$

into Eq. (123), the Dirac equation reads as

$$\left[ \frac{1}{c} \gamma^0 (i\hbar\partial_t - q\phi) - \vec{\gamma}(\vec{p} - q\vec{A}) - mc \right] \psi = 0. \quad (133)$$

Multiplying by  $-c\gamma^0$  from the left,

$$\left[ (-i\hbar\partial_t + q\phi) + c\gamma^0\vec{\gamma}(\vec{p} - q\vec{A}) + \gamma^0 mc^2 \right] \psi = 0, \quad (134)$$

we arrive at the following equation,

$$i\hbar\partial_t\psi = \left[ c\gamma^0\vec{\gamma}(\vec{p} - q\vec{A}) + q\phi + \gamma^0 mc^2 \right] \psi. \quad (135)$$

Let us introduce the matrices  $\vec{\alpha} \equiv \gamma^0\vec{\gamma}$  and  $\beta \equiv \gamma^0$ . In standard representation the matrices  $\alpha_i$  take the form,

$$\alpha_i = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix} \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}, \quad (136)$$

while the matrix  $\beta$  is

$$\beta = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}. \quad (137)$$

With these matrices the Dirac equation can be written as

$$i\hbar\partial_t\psi = \left[ c\vec{\alpha}(\vec{p} - q\vec{A}) + \beta mc^2 + q\phi \right] \psi, \quad (138)$$

or

$$i\hbar\partial_t\psi = H\psi, \quad (139)$$

where the relativistic Hamilton operator is defined as

$$H = c\vec{\alpha}(\vec{p} - q\vec{A}) + \beta mc^2 + q\phi, \quad (140)$$

which is Hermitian by construction:

$$(\alpha^i)^\dagger = (\gamma^0\gamma^i)^\dagger = \gamma^0\gamma^i = \alpha^i. \quad (141)$$

Interestingly, in absence of the vector potential the first two terms in  $H$  can be considered the square root of  $c^2\vec{p}^2 + (mc^2)^2$  entering the relativistic classical Hamiltonian:

$$\begin{aligned} [c\vec{\alpha}\vec{p} + \beta mc^2]^2 &= c^2\alpha^i\alpha^j p_i p_j + \beta^2 (mc^2)^2 + \{\alpha^i, \beta\} p_i mc^3 \\ \alpha^i\alpha^j p_i p_j &= \gamma^0\gamma^i\gamma^0\gamma^j p_i p_j = -\gamma^i\gamma^j p_i p_j = -\frac{1}{2} \{\gamma^i, \gamma^j\} p_i p_j = \vec{p}^2 I_4 \\ \beta^2 (mc^2)^2 &= (mc^2)^2 I_4 \\ \{\alpha^i, \beta\} &= \alpha^i\beta + \beta\alpha^i = \gamma^0\gamma^i\gamma^0 + \gamma^0\gamma^0\gamma^i = -\gamma^i + \gamma^i = 0 \end{aligned}$$

$$\Downarrow$$

$$\left[ c\vec{\alpha}\vec{p} + \beta mc^2 \right]^2 = \left[ c^2\vec{p}^2 + (mc^2)^2 \right] I_4 \quad (142)$$

thus, formally we can write

$$\sqrt{\left[ c^2\vec{p}^2 + (mc^2)^2 \right]} I_4 = c\vec{\alpha}\vec{p} + \beta mc^2. \quad (143)$$

In the case of static vector and scalar potentials the wave function can be written as  $\psi(\vec{r}, t) = \psi(\vec{r}) e^{-\frac{i}{\hbar}Et}$  and we arrive at the eigenvalue problem of the Hamilton operator, called as the *time-independent Dirac equation*,

$$\left[ c\vec{\alpha} \left( \vec{p} - q\vec{A}(\vec{r}) \right) + q\phi(\vec{r}) + \beta mc^2 \right] \psi(\vec{r}) = E \psi(\vec{r}). \quad (144)$$

### 3.2.3 The spectrum of free particles and the positron

In case of zero vector and scalar potentials, the solution of Eq. (144),

$$\begin{pmatrix} mc^2 I_2 & c\vec{\sigma}\vec{p} \\ c\vec{\sigma}\vec{p} & -mc^2 I_2 \end{pmatrix} \psi(\vec{r}) = E \psi(\vec{r}), \quad (145)$$

can be looked for in a plane wave form,

$$\psi(\vec{r}) = U e^{i\vec{k}\vec{r}} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} e^{i\vec{k}\vec{r}}. \quad (146)$$

Substituting  $\psi(\vec{r})$  into the Dirac equation (145) we arrive at the eigenvalue equation,

$$H(\vec{k})U = EU, \quad (147)$$

of the  $4 \times 4$  Hamilton matrix,

$$H(\vec{k}) = \begin{pmatrix} mc^2 I_2 & c\hbar\vec{\sigma}\vec{k} \\ c\hbar\vec{\sigma}\vec{k} & -mc^2 I_2 \end{pmatrix}. \quad (148)$$

Introducing the two-component amplitudes,

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \quad \text{and} \quad v = \begin{pmatrix} u_3 \\ u_4 \end{pmatrix}, \quad (149)$$

the eigenvalue equation of  $H(\vec{k})$  can be transformed to

$$\begin{pmatrix} (E - mc^2) I_2 & -c\hbar\vec{\sigma}\vec{k} \\ -c\hbar\vec{\sigma}\vec{k} & (E + mc^2) I_2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} (E - mc^2) u - c\hbar\vec{\sigma}\vec{k} v \\ -c\hbar\vec{\sigma}\vec{k} u + (E + mc^2) v \end{pmatrix} = 0. \quad (150)$$

Solving the above equations leads to the solutions,

$$v = \frac{c\hbar\vec{\sigma}\vec{k}}{E + mc^2} u \quad (151)$$

and

$$\left( (E - mc^2) (E + mc^2) I_2 - (c\hbar\vec{\sigma}\vec{k})^2 \right) u = 0. \quad (152)$$

Using the identity  $(\vec{\sigma}\vec{k})^2 = k^2 I_2$ , the above equation leads to the solution for  $E$ ,

$$(E - mc^2) (E + mc^2) - c^2 \hbar^2 k^2 = 0 \quad (153)$$

$\Downarrow$

$$E^2 = m^2 c^4 + c^2 \hbar^2 k^2, \quad (154)$$

admitting for the energy of the free particle,

$$E = \pm \sqrt{m^2 c^4 + c^2 \hbar^2 k^2}. \quad (155)$$

The solutions are discussed in details at the practical course. Here we only note that both the positive and the negative energy solutions can be chosen as eigenfunctions of the helicity matrix  $\frac{1}{k} \vec{\sigma}\vec{k}$ , since the Hamiltonian  $H(\vec{k})$  commutes with  $\vec{\sigma}\vec{k}$ ,

$$\left[ \begin{pmatrix} mc^2 I_2 & c\hbar\vec{\sigma}\vec{k} \\ c\hbar\vec{\sigma}\vec{k} & -mc^2 I_2 \end{pmatrix}, \begin{pmatrix} \vec{\sigma}\vec{k} & 0 \\ 0 & \vec{\sigma}\vec{k} \end{pmatrix} \right] = 0 \quad (i = x, y, z). \quad (156)$$

Equation (155) demonstrates that, similar to the Klein-Gordon equation, the energy of the relativistic particles can be either positive or negative,  $E > mc^2$  or  $E < -mc^2$ , respectively. In particular, the spectrum of the energy is not bounded from below. This implies that the Rayleigh-Ritz variational method can not be used to estimate the ground state energy of the system.

A more fundamental problem arises that a system with such an energy spectrum would be unstable: particles can loose their energy by emitting photons while getting endless down along the negative energy levels. A way to avoid this problem was proposed by Dirac (1927) based on the fact that the electrons are fermions. He supposed that in the ground state all the negative energy states are occupied with electrons which prevents the transition of electrons with positive energy down to the negative energy states. This ground state is called the *Dirac sea*.

Based on the assumption of the Dirac sea, an electron with  $E < -mc^2$  could be excited into an unoccupied state with  $E > mc^2$  by absorbing energy from electromagnetic radiation larger than  $2mc^2$ . During this process, an electron with  $E > mc^2$  and a hole in the Dirac sea will be created. Such a hole is detected as an electron missing in the Dirac sea: the missing energy  $E < -mc^2$  is associated with a positive energy of the hole,  $E > mc^2$ , and also the missing charge  $-e$  corresponds to a positive charge of  $e$ . The particle, related to a hole in the Dirac sea, is called the *positron* (antiparticle of the electron), which has the same (positive) mass as the electron, but a positive unit charge. The creation of an electron-positron pair (termed as pair production) was confirmed by Carl Anderson in 1932. Interestingly, the creation of an electron-positron pair can not happen just by absorbing a single photon, since within such a process the conservation of the momentum can not be fulfilled (see practical course). Usually it is the nucleus of an atom that absorbs the excess momentum of the photon. Obviously, the reverse process can also occur: an electron-positron pair is annihilated by emitting a photon.

### 3.2.4 Continuity equation

Let us start from the Dirac equation (138),

$$i\hbar\partial_t\psi = \left[ c\vec{\alpha}(\vec{p} - q\vec{A}) + q\phi + \beta mc^2 \right] \psi, \quad (157)$$

or written for the components of the wave function,

$$i\hbar\partial_t\psi_r = c\vec{\alpha}_{rs}(\vec{p}\psi_s - q\vec{A}\psi_s) + q\phi\psi_r + mc^2\beta_{rs}\psi_s \quad (r = 1, 2, 3, 4). \quad (158)$$

The complex conjugate of the above equations reads as

$$-i\hbar\partial_t\psi_r^* = c\vec{\alpha}_{rs}^*(-\vec{p}\psi_s^* - q\vec{A}\psi_s^*) + q\phi\psi_r^* + mc^2\beta_{rs}^*\psi_s^* \quad (159)$$

$$= -c(\vec{p}\psi_s^*)\vec{\alpha}_{sr} - q\vec{A}\psi_s^*\vec{\alpha}_{sr} + q\phi\psi_r^* + mc^2\psi_s^*\beta_{sr}, \quad (160)$$

where we used that  $(\vec{p}\psi_r)^* = (\frac{\hbar}{i}\vec{\nabla}\psi_r)^* = -\frac{\hbar}{i}\vec{\nabla}\psi_r^* = -\vec{p}\psi_r^*$ , and that  $\vec{\alpha}$  and  $\beta$  are Hermitian matrices. The above equations can be summarized as,

$$-i\hbar\partial_t\psi^\dagger = c(-\vec{p} - q\vec{A})\psi^\dagger\vec{\alpha} + q\phi\psi^\dagger + mc^2\psi^\dagger\beta, \quad (161)$$

with

$$\psi^\dagger = (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*) \quad (162)$$

the adjoint of the Dirac spinor. If we multiply Eq. (157) with  $\psi^+$  from the left, and Eq. (161) with  $\psi$  from the right and subtract the two equations from each other, we get

$$i\hbar[\psi^+\partial_t\psi + (\partial_t\psi^+)\psi] = c[\psi^+\vec{\alpha}(\vec{p}\psi) + (\vec{p}\psi^+)\vec{\alpha}\psi] \quad (163)$$

$\Downarrow$

$$\partial_t(\psi^+\psi) + \vec{\nabla}(c\psi^+\vec{\alpha}\psi) = 0. \quad (164)$$

Defining the finding probability density and the current density as,

$$\rho = \psi^+\psi \quad \text{and} \quad \vec{j} = c\psi^+\vec{\alpha}\psi, \quad (165)$$

we arrive at the continuity equation,

$$\partial_t\rho + \vec{\nabla}\vec{j} = 0. \quad (166)$$

The probability density is obviously positive definite, so the probabilistic interpretation of quantum mechanics applies to the particle described by the Dirac equation.

Later on we will prove that the four-current

$$j^\mu = (c\rho, \vec{j}) = c\psi^+\gamma^0\gamma^\mu\psi \quad (167)$$

is a contravariant four-vector, leading to the covariant form of the continuity equation,

$$\partial_\mu j^\mu = 0 \quad (168)$$

which applies in any inertial system. An important consequence of the continuity equation is the conservation of the finding probability of the particle,

$$\int \rho(\vec{r}, t) d^3r = \int \psi^+(\vec{r}, t) \psi(\vec{r}, t) d^3r = 1. \quad (169)$$

### 3.2.5 Conjugate spinor and the conjugate Dirac equation

The continuity equation can also be obtained from the covariant form of the Dirac equation, see Eq. (126)

$$\gamma^\mu \partial_\mu \psi + \frac{iq}{\hbar} \gamma^\mu A_\mu \psi + i\kappa \psi = 0. \quad (170)$$

Writing it out by components,

$$\gamma_{rs}^\mu \partial_\mu \psi_s + \frac{iq}{\hbar} \gamma_{rs}^\mu A_\mu \psi_s + i\kappa \psi_r = 0, \quad (171)$$

then taking the complex conjugate,

$$(\gamma_{rs}^\mu)^* \partial_\mu \psi_s^* - \frac{iq}{\hbar} (\gamma_{rs}^\mu)^* A_\mu \psi_s^* - i\kappa \psi_r^* = 0, \quad (172)$$

$$\Downarrow \quad (173)$$

$$(\partial_\mu \psi_s^*) (\gamma^\mu)_{sr}^\dagger - \frac{iq}{\hbar} A_\mu \psi_s^* (\gamma^\mu)_{sr}^\dagger - i\kappa \psi_r^* = 0, \quad (174)$$

leads to

$$(\partial_\mu \psi^\dagger) (\gamma^\mu)^\dagger - \frac{iq}{\hbar} A_\mu \psi^\dagger (\gamma^\mu)^\dagger - i\kappa \psi^\dagger = 0. \quad (175)$$

We make use of the identity for the adjoint of the  $\gamma^\mu$  matrices,

$$(\partial_\mu \psi^\dagger) \gamma^0 \gamma^\mu \gamma^0 - \frac{iq}{\hbar} A_\mu \psi^\dagger \gamma^0 \gamma^\mu \gamma^0 - i\kappa \psi^\dagger = 0, \quad (176)$$

and introduce the *conjugate spinor* wavefunction.

$$\bar{\psi} = \psi^\dagger \gamma^0, \quad (177)$$

to arrive at the *conjugate Dirac equation*,

$$(\partial_\mu \bar{\psi}) \gamma^\mu - \frac{iq}{\hbar} A_\mu \bar{\psi} \gamma^\mu - i\kappa \bar{\psi} = 0. \quad (178)$$

Multiplying Eq. (170) with  $\bar{\psi}$  from the left and Eq. (178) with  $\psi$  from the right, then adding the two equations leads to

$$\bar{\psi} \gamma^\mu \partial_\mu \psi + (\partial_\mu \bar{\psi}) \gamma^\mu \psi = 0, \quad (179)$$

or

$$\partial_\mu (\bar{\psi} \gamma^\mu \psi) = 0. \quad (180)$$

By defining the four-current density as

$$j^\mu = c \bar{\psi} \gamma^\mu \psi = c \psi^\dagger \gamma^0 \gamma^\mu \psi \quad (181)$$

we obtain the Lorentz invariant continuity equation

$$\partial_\mu j^\mu = 0. \quad (182)$$

### 3.3 Lorentz invariance of the Dirac equation

Let us start with the Dirac equation in coordinate representation,

$$(\gamma^\mu (i\hbar\partial_\mu - qA_\mu(x)) - mc)\psi(x) = 0, \quad (183)$$

and recall the Lorentz transformation of the covariant vectors appearing in the Dirac equation,

$$x'^\mu = \Lambda^\mu_{\ \nu} x^\nu, \quad (184)$$

$$\partial_\mu = \frac{\partial x'^\nu}{\partial x^\mu} \partial'_\nu = \partial'_\nu \Lambda^\nu_{\ \mu} = (\Lambda^T)^\nu_{\ \mu} \partial'_\nu \rightarrow \partial'_\mu = \bar{\Lambda}_\mu^{\ \nu} \partial_\nu \quad (185)$$

$$A'_\mu(x') = \bar{\Lambda}_\mu^{\ \nu} A_\nu(x). \quad (186)$$

Under Lorentz transformation the Dirac equation should be written as

$$(\gamma^\mu (i\hbar\partial'_\mu - qA'_\mu(x')) - mc)\psi'(x') = 0 \quad (187)$$

$\Downarrow$

$$\left[ \gamma^\mu (\bar{\Lambda})^\nu_{\ \mu} (i\hbar\partial_\nu - qA_\nu(x)) - mc \right] \psi'(x') = 0, \quad (188)$$

where  $\psi'(x')$  is the transformed wavefunction. Introducing the Lorentz transformation of the Dirac spinor,  $\mathcal{S}(\Lambda)$ , being a  $4 \times 4$  matrix (similarly to  $\gamma^\mu$ ),

$$\psi'(x') = \mathcal{S}(\Lambda) \psi(x), \quad (189)$$

equation (188) transforms to

$$\gamma^\mu \bar{\Lambda}_\mu^{\ \nu} \mathcal{S}(\Lambda) (i\hbar\partial_\nu - qA_\nu(x)) \psi(x) - mc \mathcal{S}(\Lambda) \psi(x) = 0. \quad (190)$$

Applying  $\mathcal{S}(\Lambda)$  to equation (183) we obtain

$$\mathcal{S}(\Lambda) \gamma^\mu (i\hbar\partial_\mu - qA_\mu(x)) \psi(x) - mc \mathcal{S}(\Lambda) \psi(x) = 0. \quad (191)$$

By taking the difference of the above two equations, the  $mc \mathcal{S}(\Lambda) \psi(x)$  term vanishes,

$$\left( \gamma^\mu \bar{\Lambda}_\mu^{\ \nu} \mathcal{S}(\Lambda) - \mathcal{S}(\Lambda) \gamma^\nu \right) (i\hbar\partial_\nu - qA_\nu(x)) \psi(x) = 0. \quad (192)$$

The above equality applies to arbitrary wavefunctions, implying the identity

$$\gamma^\mu \bar{\Lambda}_\mu^{\ \nu} \mathcal{S}(\Lambda) - \mathcal{S}(\Lambda) \gamma^\nu = 0, \quad (193)$$

or by using  $\bar{\Lambda}^T = \Lambda^{-1}$ ,

$$(\bar{\Lambda}^T)^\nu_{\ \mu} \gamma^\mu \mathcal{S}(\Lambda) - \mathcal{S}(\Lambda) \gamma^\nu = (\bar{\Lambda}^{-1})^\nu_{\ \mu} \gamma^\mu \mathcal{S}(\Lambda) - \mathcal{S}(\Lambda) \gamma^\nu = 0, \quad (194)$$

$\Downarrow$

$$\gamma^\mu \mathcal{S}(\Lambda) - \Lambda^\mu_{\ \nu} \mathcal{S}(\Lambda) \gamma^\nu = 0 \quad (195)$$

$\Downarrow$

$$\gamma^\mu = \Lambda^\mu_{\ \nu} \mathcal{S}(\Lambda) \gamma^\nu \mathcal{S}(\Lambda)^{-1}. \quad (196)$$

The above equation manifests the invariance of the  $\gamma^\mu$  matrices under Lorentz transformations, which lies at the heart of the Lorentz invariance of the Dirac equation. The above expression can be rewritten as,

$$\mathcal{S}(\Lambda)^{-1} \gamma^\mu \mathcal{S}(\Lambda) = \Lambda^\mu_{\nu} \gamma^\nu, \quad (197)$$

which we will use next for the construction of  $\mathcal{S}(\Lambda)$ .

*Statement:* Given the infinitesimal Lorentz transformation  $\omega$ , the solution of equation (197) for  $\Lambda = e^\omega$  is

$$\mathcal{S}(\Lambda) = e^{\frac{1}{4}\omega_{\mu\nu}\gamma^\mu\gamma^\nu}. \quad (198)$$

*Proof:* Let us introduce the infinitesimal spinor transformation  $\tau$  as,

$$\mathcal{S}(\Lambda) = e^{-\tau}. \quad (199)$$

On the one hand, using the Hausdorff expansion yields

$$\begin{aligned} \mathcal{S}(\Lambda)^{-1} \gamma^\mu \mathcal{S}(\Lambda) &= e^\tau \gamma^\mu e^{-\tau} \\ &= \gamma^\mu + [\tau, \gamma^\mu] + \frac{1}{2!} [\tau, [\tau, \gamma^\mu]] + \frac{1}{3!} [\tau, [\tau, [\tau, \gamma^\mu]]] + \dots \end{aligned} \quad (200)$$

On the other hand, the right-hand side of equation (197) can be expanded as

$$\Lambda^\mu_{\nu} \gamma^\nu = (e^\omega)^\mu_{\nu} \gamma^\nu = \gamma^\mu + \omega^\mu_{\nu} \gamma^\nu + \frac{1}{2!} (\omega^2)^\mu_{\nu} \gamma^\nu + \frac{1}{3!} (\omega^3)^\mu_{\nu} \gamma^\nu + \dots \quad (201)$$

Obviously, the first terms of (200) and (201) are identical. Let's prove the identity of the second terms,

$$[\tau, \gamma^\mu] = \omega^\mu_{\nu} \gamma^\nu, \quad (202)$$

provided  $\tau = -\frac{1}{4}\omega_{\mu\nu}\gamma^\mu\gamma^\nu$ . By using the anticommutation relations of the  $\gamma^\mu$  matrices,

$$\begin{aligned} \tau \gamma^\mu &= -\frac{1}{4}\omega_{\lambda\delta}\gamma^\lambda\gamma^\delta\gamma^\mu = \frac{1}{4}\omega_{\lambda\delta}\gamma^\lambda\gamma^\mu\gamma^\delta - \frac{1}{2}\omega_{\lambda\delta}\gamma^\lambda g^{\mu\delta} \\ &= -\frac{1}{4}\omega_{\lambda\delta}\gamma^\mu\gamma^\lambda\gamma^\delta + \frac{1}{2}\omega_{\lambda\delta}\gamma^\delta g^{\lambda\mu} - \frac{1}{2}\omega_{\lambda\delta}\gamma^\lambda g^{\mu\delta} \\ &= \gamma^\mu\tau + \frac{1}{2}\omega_{\lambda\delta}\gamma^\delta g^{\lambda\mu} + \frac{1}{2}g^{\mu\delta}\omega_{\delta\lambda}\gamma^\lambda = \gamma^\mu\tau + \omega^\mu_{\lambda}\gamma^\lambda, \end{aligned} \quad (203)$$

where we used that  $\omega_{\lambda\delta} = -\omega_{\delta\lambda}$  and  $g^{\lambda\mu} = g^{\mu\lambda}$ . Next we show by complete induction that

$$\underbrace{[\tau, [\tau, \dots [\tau, \gamma^\mu]]]}_n = (\omega^n)^\mu_{\nu} \gamma^\nu, \quad (204)$$

where  $n$  is the number of the embedded commutators (or the number of  $\tau$  matrices) on the left-hand side,  $n \geq 2$ .

$$\begin{aligned} [\tau, \underbrace{[\tau, [\tau, \dots [\tau, \gamma^\mu]]}_{n-1}] &= [\tau, (\omega^{n-1})^\mu_{\nu} \gamma^\nu] = (\omega^{n-1})^\mu_{\nu} [\tau, \gamma^\nu] \\ &= (\omega^{n-1})^\mu_{\nu} \omega^\nu_{\lambda} \gamma^\lambda = (\omega^n)^\mu_{\lambda} \gamma^\lambda. \end{aligned} \quad (205)$$

Thus, the proof is complete.

We note that the spinor transformation  $\mathcal{S}(\Lambda)$  is often written in the form,

$$\mathcal{S}(\Lambda) = e^{-\frac{i}{\hbar}\omega_{\mu\nu}\sigma^{\mu\nu}}, \quad (206)$$

where

$$\sigma^{\mu\nu} = \frac{i\hbar}{4}\gamma^\mu\gamma^\nu = \frac{i\hbar}{8}[\gamma^\mu, \gamma^\nu] \quad (\mu \neq \nu), \quad (207)$$

is the infinitesimal generator of the spinor representation of the Lorentz transformations.

### 3.3.1 The four-current density

Let's consider an operator  $O$  in the space of Dirac spinors,  $O = f(\gamma^\mu)$ , where  $f$  is an analytic function. The Minkowski-space density of  $O$  is defined as

$$O(x) = \psi(x)^\dagger O \psi(x) , \quad (208)$$

or using the conjugate Dirac spinor  $\bar{\psi}(x) = \psi(x)^\dagger \gamma^0$  and the operator

$$\bar{O} = \gamma^0 O , \quad (209)$$

as

$$O(x) = \bar{\psi}(x) \bar{O} \psi(x) . \quad (210)$$

An example is the finding probability density and current density,

$$\varrho(x) = \bar{\psi}(x) \gamma^0 \psi(x) , \quad j^k(x) = c \bar{\psi}(x) \gamma^k \psi(x) \quad (k = 1, 2, 3) , \quad (211)$$

or the four-current density,

$$j^\mu(x) = (c\varrho(x), \vec{j}(x)) = \bar{\psi}(x) (c\gamma^\mu) \psi(x) . \quad (212)$$

Since the  $O$  operator is Lorentz invariant (in any inertial system we use the same  $\gamma^\mu$  matrices), the Lorentz transform of the density  $O(x)$  can be calculated as,

$$\begin{aligned} O'(x') &= \psi'(x')^\dagger O \psi'(x') \\ &= [\mathcal{S}(\Lambda) \psi(x)]^\dagger O \mathcal{S}(\Lambda) \psi(x) \\ &= \psi(x)^\dagger \mathcal{S}(\Lambda)^\dagger O \mathcal{S}(\Lambda) \psi(x) \\ &= \bar{\psi}(x) \gamma^0 \mathcal{S}(\Lambda)^\dagger \gamma^0 \gamma^0 O \mathcal{S}(\Lambda) \psi(x) \\ &= \bar{\psi}(x) \gamma^0 \mathcal{S}(\Lambda)^\dagger \gamma^0 \bar{O} \mathcal{S}(\Lambda) \psi(x) . \end{aligned} \quad (213)$$

*Statement:* For a Lorentz transformation  $\Lambda = e^\omega$ ,

$$\mathcal{S}(\Lambda)^{-1} = \gamma^0 \mathcal{S}(\Lambda)^\dagger \gamma^0 . \quad (214)$$

*Proof:* Substituting  $\mathcal{S}(\Lambda) = e^{-\tau}$ , the right-hand side of the above equality can be written as

$$\gamma^0 \mathcal{S}(\Lambda)^\dagger \gamma^0 = \gamma^0 (e^{-\tau})^\dagger \gamma^0 = \sum_{n=0}^{\infty} \frac{1}{n!} \gamma^0 (-\tau^\dagger)^n \gamma^0 . \quad (215)$$

Using  $\tau = -\frac{1}{4} \omega_{\mu\nu} \gamma^\mu \gamma^\nu$ ,  $\gamma^\mu = \gamma^0 (\gamma^\mu)^\dagger \gamma^0$ , and the fact that  $\omega_{\mu\nu}$  is an antisymmetric matrix,

$$\gamma^0 (-\tau^\dagger) \gamma^0 = \frac{1}{4} \omega_{\mu\nu} \gamma^0 (\gamma^\nu)^\dagger (\gamma^\mu)^\dagger \gamma^0 = \frac{1}{4} \omega_{\mu\nu} \gamma^0 (\gamma^\nu)^\dagger \gamma^0 \gamma^0 (\gamma^\mu)^\dagger \gamma^0 \quad (216)$$

$$= \frac{1}{4} \omega_{\mu\nu} \gamma^\nu \gamma^\mu = -\frac{1}{4} \omega_{\mu\nu} \gamma^\mu \gamma^\nu = \tau \quad (217)$$

$\Downarrow$

$$\gamma^0 (-\tau^\dagger)^2 \gamma^0 = (\gamma^0 (-\tau^\dagger) \gamma^0)^2 = \tau^2 \dots \gamma^0 (-\tau^\dagger)^n \gamma^0 = \tau^n \quad (218)$$

$\Downarrow$

$$\gamma^0 \mathcal{S}(\Lambda)^\dagger \gamma^0 = \sum_{n=0}^{\infty} \frac{\tau^n}{n!} = e^\tau = \mathcal{S}(\Lambda)^{-1} . \quad (219)$$

The Lorentz transform of the density  $O(x)$ , Eq. (213), can then be expressed as

$$O'(x') = \bar{\psi}(x) (\mathcal{S}(\Lambda)^{-1} \overline{O} \mathcal{S}(\Lambda)) \psi(x) . \quad (220)$$

Based on equation (197) and (220),

$$\begin{aligned} j'^\mu(x') &= c \bar{\psi}(x) (\mathcal{S}(\Lambda)^{-1} \gamma^\mu \mathcal{S}(\Lambda)) \psi(x) \\ &= c \bar{\psi}(x) (\Lambda^\mu_{\ \nu} \gamma^\nu) \psi(x) = \Lambda^\mu_{\ \nu} [c \bar{\psi}(x) \gamma^\nu \psi(x)] \\ &= \Lambda^\mu_{\ \nu} j^\nu(x) , \end{aligned} \quad (221)$$

proving that the four-current density transforms as a contravariant vector. From this it immediately follows that the finding probability density,  $\varrho(x) = \frac{1}{c} j^0(x)$ , is not invariant under the Lorentz transformation.

One can prove that  $\bar{\psi}\psi$  and  $\bar{\psi}\gamma^5\psi$  transform as a scalar,  $\bar{\psi}\gamma^5\gamma^\mu\psi$  as a vector, and  $\bar{\psi}\gamma^\mu\gamma^\nu\psi$  as a tensor of rank 2. (Note that  $\gamma^5$  commutes with  $\mathcal{S}(\Lambda)$ !)

### 3.3.2 Rotations and the spin

Let's consider an anti-clockwise infinitesimal spatial rotation by angle  $\varphi$  around the axis  $\vec{n}$ ,

$$\vec{r}' = \vec{r} + \varphi \vec{n} \times \vec{r} , \quad (222)$$

or for  $i = 1, 2, 3$

$$x'^i = x^i + \varepsilon_{ikj} n_k x^j \varphi = \omega^i_{\cdot j} x^j .$$

The matrix of the infinitesimal Lorentz transformation is then given by

$$\omega^i_{\cdot j} = \varepsilon_{ikj} n_k \varphi \quad (i, j = 1, 2, 3) , \quad (223)$$

$$\omega^{\mu}_{\cdot 0} = \omega^0_{\cdot \mu} = 0 \quad (\mu = 0, 1, 2, 3) , \quad (224)$$

since  $x'^0 = x^0$ . The corresponding infinitesimal transformation in spinor space then reads as

$$\frac{1}{4} \omega_{\mu\nu} \gamma^\mu \gamma^\nu = \frac{1}{4} \omega_{ij} \gamma^i \gamma^j = -\frac{1}{4} \varepsilon_{ikj} \gamma^i \gamma^j n_k \varphi = \frac{1}{4} \varepsilon_{kij} \gamma^i \gamma^j n_k \varphi = -\frac{i}{\hbar} \vec{n} \cdot \vec{S} \varphi ,$$

where

$$S^k = \frac{i\hbar}{4} \varepsilon_{kij} \gamma^i \gamma^j \rightarrow \vec{S} = \frac{i\hbar}{4} \vec{\gamma} \times \vec{\gamma} . \quad (225)$$

The components are then given by

$$S^1 = \frac{\hbar}{2} i \gamma^2 \gamma^3 , \quad S^2 = \frac{\hbar}{2} i \gamma^3 \gamma^1 , \quad S^3 = \frac{\hbar}{2} i \gamma^1 \gamma^2 . \quad (226)$$

Let us calculate the commutation relations of the  $S^i$  matrices,

$$[S^1, S^2] = \left( \frac{i\hbar}{2} \right)^2 [\gamma^2 \gamma^3, \gamma^3 \gamma^1] = \left( \frac{i\hbar}{2} \right)^2 (\gamma^2 \gamma^3 \gamma^3 \gamma^1 - \gamma^3 \gamma^1 \gamma^2 \gamma^3) \quad (227)$$

$$= \left( \frac{i\hbar}{2} \right)^2 (-\gamma^2 \gamma^1 + \gamma^1 \gamma^2) = i\hbar \frac{i\hbar}{2} \gamma^1 \gamma^2 = i\hbar S^3 , \quad (228)$$

and similarly,

$$[S^2, S^3] = i\hbar S^1, \quad [S^3, S^1] = i\hbar S^2, \quad (229)$$

that can be compactly written as

$$[S^i, S^j] = i\hbar \varepsilon_{ijk} S^k. \quad (230)$$

Thus the infinitesimal generators  $S^i$  of the spatial rotations satisfy the Lie algebra. Furthermore,

$$(S^1)^2 = -\frac{\hbar^2}{4} \gamma^2 \gamma^3 \gamma^2 \gamma^3 = \frac{\hbar^2}{4} I \quad (231)$$

and, similarly,

$$(S^2)^2 = (S^3)^2 = \frac{\hbar^2}{4} I, \quad (232)$$

so

$$\vec{S}^2 = \frac{3\hbar^2}{4} I = \hbar^2 s(s+1) I \implies s = \frac{1}{2}. \quad (233)$$

The spatial rotations of the spinor wavefunctions in the Dirac equation are generated by  $s = \frac{1}{2}$  angular momentum operators that can be interpreted such that the particles described by the Dirac equation (electrons) exhibit a  $\frac{1}{2}$ -spin. It is now clear that the Klein-Gordon equation, where the wavefunction was scalar, can at best describe a spinless particle (field).

Using the standard representation of the  $\gamma^\mu$  matrices,

$$\vec{S} = \frac{i\hbar}{4} \vec{\gamma} \times \vec{\gamma} = \frac{i\hbar}{4} \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix} \times \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix} \quad (234)$$

$$= -\frac{i\hbar}{4} \begin{pmatrix} \vec{\sigma} \times \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \times \vec{\sigma} \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} = \frac{\hbar}{2} \vec{\Sigma}, \quad (235)$$

where

$$\vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}. \quad (236)$$

Reassuringly, this form of the spin operators is a natural extension of the spin operators  $\frac{\hbar}{2} \vec{\sigma}$  introduced in the Pauli-Schrödinger equation.

### 3.3.3 The total angular momentum

As we have seen, the transformation of the wavefunction under a spatial rotation  $R(\varphi, \vec{n})$  is given by

$$\psi'(R(\varphi, \vec{n}) \vec{r}, t) = \exp\left(-\frac{i}{\hbar} \vec{n} \vec{S} \varphi\right) \psi(\vec{r}, t) \quad (237)$$

or

$$\psi'(\vec{r}, t) = \exp\left(-\frac{i}{\hbar} \vec{n} \vec{S} \varphi\right) \psi(R(-\varphi, \vec{n}) \vec{r}, t). \quad (238)$$

Let us express  $\psi(R(-\varphi, \vec{n}) \vec{r}, t)$  for an infinitesimal rotation,

$$\begin{aligned} \psi(R(-\varphi, \vec{n}) \vec{r}, t) &\simeq \psi(\vec{r}, t) - \varphi (\vec{n} \times \vec{r}) \cdot \vec{\nabla} \psi(\vec{r}, t) \\ &= \psi(\vec{r}, t) - \varphi \vec{n} \cdot (\vec{r} \times \vec{\nabla}) \psi(\vec{r}, t) \\ &= \psi(\vec{r}, t) - \frac{i}{\hbar} \varphi \vec{n} \cdot (\vec{r} \times \vec{p}) \psi(\vec{r}, t) \\ &= \left[ I - \frac{i}{\hbar} \varphi \vec{n} \cdot \vec{L} \right] \psi(\vec{r}, t), \end{aligned} \quad (239)$$

consequently, for finite rotations,

$$\psi(R(-\varphi, \vec{n}) \vec{r}, t) = \lim_{N \rightarrow \infty} (I - \frac{i}{\hbar} \vec{n} \vec{L} \frac{\varphi}{N})^N = \exp\left(-\frac{i}{\hbar} \vec{n} \vec{L} \varphi\right) \psi(\vec{r}, t) . \quad (240)$$

The complete transformation of the four-component wavefunction under spatial rotations can then be written as,

$$\psi'(\vec{r}, t) = \exp\left(-\frac{i}{\hbar} \vec{n} \vec{J} \varphi\right) \psi(\vec{r}, t) , \quad (241)$$

thus in relativistic quantum theory the infinitesimal generator of the spatial rotations is the *total angular momentum* operator,

$$\vec{J} = \vec{L} + \vec{S} .$$

**Time derivative of the total angular momentum operator** Based on the abstract formalism of quantum mechanics (see Quantum Mechanics 1) and the Schrödinger form of the Dirac equation,

$$i\hbar \partial_t \psi = H \psi \quad (242)$$

the quantum mechanical time derivative of an operator  $A$  can be defined,

$$\frac{dA}{dt} = \frac{i}{\hbar} [H, A] + \partial_t A . \quad (243)$$

This makes possible to establish whether the given operator is a conserved quantity of motion. In the non-relativistic theory, for spherical symmetric potentials we have seen that the angular momentum

$$\vec{L} = \vec{r} \times \vec{p} \quad (244)$$

is conserved. Let us check whether this applies in the relativistic case! In the absence of vector potential the Hamiltonian takes the form,

$$H = c\vec{\alpha}\vec{p} + \beta mc^2 + q\phi . \quad (245)$$

The required commutator can be written as

$$[H, L_i] = [c\vec{\alpha}\vec{p} + q\phi, \varepsilon_{ijk} x_j p_k] , \quad (246)$$

since  $L_i$  commutes with  $\beta mc^2$ . (The  $L_i$  operator acts as an identity in the spinor space, so correctly we had to write  $L_i \otimes I_4$ .) Separating the commutator into two parts,

$$\begin{aligned} [c\alpha_l p_l, \varepsilon_{ijk} x_j p_k] &= c\varepsilon_{ijk} \alpha_l [p_l, x_j p_k] = c\varepsilon_{ijk} \alpha_l (\underbrace{[p_l, x_j] p_k}_{\frac{\hbar}{i} \delta_{lj}} + x_j \underbrace{[p_l, p_k]}_0) \\ &= \frac{\hbar c}{i} (\vec{\alpha} \times \vec{p})_i \end{aligned} \quad (247)$$

and

$$\begin{aligned} [q\phi, \varepsilon_{ijk} x_j p_k] &= q\varepsilon_{ijk} [\phi, x_j p_k] = q\varepsilon_{ijk} (\underbrace{[\phi, x_j] p_k}_0 + x_j \underbrace{[\phi, p_k]}_{-\frac{\hbar}{i} \partial_k \phi}) \\ &= -\frac{\hbar q}{i} (\vec{r} \times \vec{\nabla} \phi)_i , \end{aligned} \quad (248)$$

yields

$$[H, \vec{L}] = \frac{\hbar}{i} \left[ c(\vec{\alpha} \times \vec{p}) - q \left( \vec{r} \times \vec{\nabla} \phi \right) \right] , \quad (249)$$

where for a central potential only the second term vanishes. Thus the angular momentum is not conserved for central potentials in the relativistic case!

We prove that, for central potentials the *total angular momentum*

$$\vec{J} = \vec{L} + \vec{S} \quad (250)$$

is conserved, where

$$\vec{S} = \frac{\hbar}{2} \vec{\Sigma} = \frac{\hbar}{2} \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} , \quad (251)$$

is the spin operator. The commutator of the Hamiltonian and spin operator is indeed

$$\begin{aligned} [H, S_i] &= \frac{\hbar c}{2} [\alpha_j p_j, \Sigma_i] = \frac{\hbar c}{2} p_l [\alpha_j, \Sigma_i] = \\ &= \frac{\hbar c}{2} p_j \left[ \begin{pmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{pmatrix} \begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} - \begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} \begin{pmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{pmatrix} \right] \\ &= \frac{\hbar c}{2} p_j \begin{pmatrix} 0 & [\sigma_j, \sigma_i] \\ [\sigma_j, \sigma_i] & 0 \end{pmatrix} = i\hbar c \varepsilon_{jik} p_j \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix} \\ &= -\frac{\hbar c}{i} (\underline{\alpha} \times \vec{p})_i , \end{aligned} \quad (252)$$

so the quantum mechanical time derivative of  $\vec{J}$  is,

$$\frac{d\vec{J}}{dt} = -q \left( \vec{r} \times \vec{\nabla} \phi \right) = \vec{r} \times \vec{F} , \quad (253)$$

where  $\vec{F} = -\vec{\nabla} (q\phi)$  is the force. The above expression indicates that the time derivative of the total angular momentum equals the torque which is zero for central potentials.

### 3.3.4 Spatial reflection (parity)

The reflection of the spacelike coordinates of the Minkowski space is defined as

$$x^{0'} = x^0 \quad x^{i'} = -x^i \quad (i = 1, 2, 3) , \quad (254)$$

yielding the following simple matrix of transformation,

$$\Lambda = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} = \bar{\Lambda} . \quad (255)$$

Multiplying the transformed Dirac equation,

$$(\gamma^\mu K'_\mu(x') - mc) \psi'(x') = (\gamma^0 K_0(x) - \gamma^i K_i(x) - mc) \psi'(x') = 0 , \quad (256)$$

by  $\gamma^0$  from the left we obtain

$$\begin{aligned} (\gamma^0 \gamma^0 K_0 - \gamma^0 \gamma^i K_i - mc \gamma^0) \psi' &= (\gamma^0 \gamma^0 K_0 + \gamma^i \gamma^0 K_i - mc \gamma^0) \psi' \\ &= (\gamma^0 K_0 + \gamma^i K_i - mc) \gamma^0 \psi' = 0, \end{aligned} \quad (257)$$

where we made use that  $\gamma^0$  and  $K_\mu$  commute, while  $\gamma^0 \gamma^i = -\gamma^i \gamma^0$ . From the last equation the transformation of the wavefunction can be read off:

$$\gamma^0 \psi'(x') = \psi(x) \rightarrow \psi'(x') = \gamma^0 \psi(x). \quad (258)$$

The transformed conjugate Dirac spinor is then given by

$$\bar{\psi}' = (\gamma^0 \psi)^\dagger \gamma^0 = \psi^\dagger (\gamma^0)^\dagger \gamma^0 = \psi^\dagger = \bar{\psi} \gamma^0. \quad (259)$$

Let us investigate how the previously introduced quantities transform under reflection,

$$(\psi')^\dagger \psi' = (\gamma^0 \psi)^\dagger \gamma^0 \psi = \psi^\dagger \gamma^0 \gamma^0 \psi = \psi^\dagger \psi, \quad (260)$$

and

$$\bar{\psi}' \psi' = \bar{\psi} \gamma^0 \gamma^0 \psi = \bar{\psi} \psi. \quad (261)$$

Furthermore,

$$\bar{\psi}' \gamma^5 \psi' = \bar{\psi} \gamma^0 \gamma^5 \gamma^0 \psi = -\bar{\psi} \gamma^5 \psi, \quad (262)$$

i.e. it changes sign under reflection. Such scalars are termed as *pseudoscalars*. Similarly,

$$\bar{\psi}' \gamma^i \psi' = \bar{\psi} \gamma^0 \gamma^i \gamma^0 \psi = -\bar{\psi} \gamma^i \psi \quad \text{and} \quad \bar{\psi}' \gamma^0 \psi' = \bar{\psi} \gamma^0 \psi, \quad (263)$$

so  $j^\mu = c \bar{\psi} \gamma^\mu \psi$  transforms as  $x^\mu$  under reflection. We denote such vectors as *polar* or *normal vectors*. On the contrary,

$$\bar{\psi}' \gamma^i \gamma^5 \psi' = \bar{\psi} \gamma^0 \gamma^i \gamma^5 \gamma^0 \psi = \bar{\psi} \gamma^i \gamma^5 \psi \quad \text{and} \quad \bar{\psi}' \gamma^0 \gamma^5 \psi' = \bar{\psi} \gamma^5 \gamma^0 \psi = -\bar{\psi} \gamma^0 \gamma^5 \psi, \quad (264)$$

i.e. the spatial components of  $\bar{\psi} \gamma^\mu \gamma^5 \psi$  remain unchanged, while the timelike component gets a minus sign. Such vectors are called *pseudo (or axial) vectors*.

### 3.4 Non-relativistic approximation and relativistic corrections

In this section we investigate the non-relativistic limit of the Dirac equation. First we split the solution of the stationary Dirac equation,

$$(c \vec{\alpha} \vec{K} + \beta mc^2 + q\phi) \psi = E \psi \quad (\vec{K} = \vec{p} - q\vec{A}), \quad (265)$$

into two parts (two components for each),

$$\psi = \begin{pmatrix} \chi \\ \varphi \end{pmatrix}. \quad (266)$$

Let's write the equation (265) in details,

$$\begin{pmatrix} E - mc^2 - q\phi & -c\vec{\sigma} \vec{K} \\ -c\vec{\sigma} \vec{K} & E + mc^2 - q\phi \end{pmatrix} \begin{pmatrix} \chi \\ \varphi \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$\Downarrow$

$$(E - mc^2 - q\phi) \chi - c\vec{\sigma} \vec{K} \varphi = 0, \quad (267)$$

$$(E + mc^2 - q\phi) \varphi - c\vec{\sigma} \vec{K} \chi = 0. \quad (268)$$

The component  $\varphi$  can be expressed from equation (268):

$$\varphi = \frac{1}{E + mc^2 - q\phi} c\vec{\sigma}\vec{K}\chi. \quad (269)$$

Restricting ourselves to the *positive energy solutions*, we apply the approximation  $E + mc^2 - q\phi \simeq 2mc^2$ ,

$$\varphi = \frac{1}{2mc} \vec{\sigma}\vec{K}\chi. \quad (270)$$

From the probability density,

$$\rho = \psi^\dagger\psi = \chi^\dagger\chi + \varphi^\dagger\varphi = \chi^\dagger\chi + \frac{1}{4m^2c^2}\chi^\dagger(\vec{\sigma}\vec{K})(\vec{\sigma}\vec{K})\chi, \quad (271)$$

we see that the component  $\varphi$  contributes in order of  $\frac{1}{c^2}$ , therefore, it is called the small component, while  $\chi$  is called the large component. Substituting (270) into equation (267) and introducing the notation  $E' = E - mc^2$  we arrive at

$$\left(E' - q\phi - \frac{1}{2m}(\vec{\sigma}\vec{K})(\vec{\sigma}\vec{K})\right)\chi = 0, \quad (272)$$

which can be transformed into the stationary Pauli-Schrödinger equation,

$$H_P\chi = E'\chi, \quad (273)$$

with the Hamilton operator including the paramagnetic term for the spin,

$$H_P = \frac{1}{2m}(\vec{\sigma}\vec{K})(\vec{\sigma}\vec{K}) + q\phi = \frac{\vec{K}^2}{2m} + q\phi - \frac{\hbar q}{2m}\vec{B}\vec{\sigma}. \quad (274)$$

Here we used the identity  $(\vec{\sigma}\vec{K})(\vec{\sigma}\vec{K}) = \vec{K}^2 + i(\vec{K} \times \vec{K})\vec{\sigma}$  and  $\vec{K} \times \vec{K} = -\frac{\hbar q}{i}\vec{B}$ , thus,  $(\vec{\sigma}\vec{K})(\vec{\sigma}\vec{K}) = \vec{K}^2 - \hbar q\vec{B}\vec{\sigma}$ .

### 3.4.1 Relativistic kinetic energy correction

Let's go beyond the approximation (270),

$$\begin{aligned} \varphi &= \frac{1}{E + mc^2 - q\phi} c\vec{\sigma}\vec{K}\chi = \frac{1}{2mc^2 + E' - q\phi} c\vec{\sigma}\vec{K}\chi \\ &= \frac{1}{2mc^2} \frac{1}{1 + \frac{E' - q\phi}{2mc^2}} c\vec{\sigma}\vec{K}\chi \\ &\simeq \frac{1}{2mc} \left(1 - \frac{E' - q\phi}{2mc^2}\right) (\vec{\sigma}\vec{K})\chi, \end{aligned} \quad (275)$$

which we substitute again into (267),

$$(E' - q\phi)\chi = \frac{1}{2m}(\vec{\sigma}\vec{K})\left(1 - \frac{E' - q\phi}{2mc^2}\right)(\vec{\sigma}\vec{K})\chi. \quad (276)$$

By reordering the above equation and using the definition of  $H_P$ , we obtain

$$\begin{aligned} E'\chi &= \left(\frac{1}{2m}(\vec{\sigma}\vec{K})(\vec{\sigma}\vec{K}) + q\phi\right)\chi - \frac{1}{4m^2c^2}(\vec{\sigma}\vec{K})(E' - q\phi)(\vec{\sigma}\vec{K})\chi \\ &= H_P\chi - \frac{1}{4m^2c^2}(\vec{\sigma}\vec{K})(E' - q\phi)(\vec{\sigma}\vec{K})\chi \end{aligned} \quad (277)$$

$$\begin{aligned} &= H_P\chi - \frac{1}{4m^2c^2}(\vec{\sigma}\vec{K})(\vec{\sigma}\vec{K})(E' - q\phi)\chi - \frac{1}{4m^2c^2}(\vec{\sigma}\vec{K})\left[E' - q\phi, \vec{\sigma}\vec{K}\right]\chi \\ &= H_P\chi - \frac{1}{2mc^2}(H_P - q\phi)(E' - q\phi)\chi - \frac{1}{4m^2c^2}(\vec{\sigma}\vec{K})\left[\vec{\sigma}\vec{K}, q\phi\right]. \end{aligned} \quad (278)$$

The second term on the right-hand side of (278) can be written in first order in  $1/c^2$  as

$$H_M \chi \equiv -\frac{1}{2mc^2} (H_P - q\phi)^2 \chi \simeq -\frac{1}{8m^3c^2} [(\vec{\sigma}\vec{p})(\vec{\sigma}\vec{p})]^2 \chi = -\frac{p^4}{8m^3c^2} \chi, \quad (279)$$

where we neglected the effects of the magnetic field. The term in the Hamilton operator,  $H_M$ , can be associated with the *relativistic correction to the kinetic energy (mass enhancement)*, since using the expansion

$$\sqrt{1+x} \simeq 1 + \frac{x}{2} - \frac{x^2}{8},$$

we find

$$\begin{aligned} E - q\phi &= \sqrt{m^2c^4 + p^2c^2} = mc^2 \sqrt{1 + \frac{p^2}{m^2c^2}} \\ &\simeq mc^2 \left( 1 + \frac{1}{2} \frac{p^2}{m^2c^2} - \frac{1}{8} \left( \frac{p^2}{m^2c^2} \right)^2 \right) \\ &= mc^2 + \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \dots \end{aligned} \quad (280)$$

Let us calculate this correction to the energy levels of the hydrogen atom in first order of time independent perturbation theory:

$$H = H_0 + W,$$

where

$$H_0 = \frac{p^2}{2m} - \frac{ke^2}{r}, \quad E_n^{(0)} = -\frac{ke^2}{2a_0} \frac{1}{n^2}, \quad (281)$$

and

$$\begin{aligned} W = H_M &= -\frac{p^4}{8m^3c^2} = -\frac{1}{2mc^2} \left( \frac{p^2}{2m} \right)^2 = -\frac{1}{2mc^2} \left( H_0 + \frac{ke^2}{r} \right)^2 \\ &= -\frac{1}{2mc^2} \left( H_0^2 + 2H_0 \frac{ke^2}{r} + \frac{(ke^2)^2}{r^2} \right). \end{aligned} \quad (282)$$

Since  $W$  commutes with  $H_0$ , the first order energy corrections are given by the diagonal matrix elements of  $W$ ,

$$\begin{aligned} \delta E_{n\ell m}^{(1)} &= \langle n\ell m | W | n\ell m \rangle \\ &= -\frac{1}{2mc^2} \left( (E_n^{(0)})^2 + 2E_n^{(0)} \underbrace{\left\langle \frac{ke^2}{r} \right\rangle_{n\ell m}}_{-2E_n^{(0)}} + k^2e^4 \underbrace{\left\langle \frac{1}{r^2} \right\rangle_{n\ell m}}_{1/a_0^2 n^3 (\ell+1/2)} \right) \\ &= -\frac{1}{2mc^2} \left( -3(E_n^{(0)})^2 + \frac{k^2e^4}{a_0^2} \frac{1}{n^4} \frac{n}{\ell+1/2} \right) = \\ &= -\frac{2(E_n^{(0)})^2}{mc^2} \left( \frac{n}{\ell+1/2} - \frac{3}{4} \right). \end{aligned} \quad (283)$$

This energy correction splits the energy levels of the hydrogen atom according to the angular quantum number  $\ell$ . This result agrees with the eigenenergies provided by the Klein-Gordon equation when expanded to first order in  $1/c^2$ .

### 3.4.2 Spin-orbit interaction

Again by neglecting the effect of the magnetic field, consider the third term in equation (278),

$$-\frac{1}{4m^2c^2} (\vec{\sigma}\vec{p}) (\vec{\sigma} [\vec{p}, q\phi]) = -\frac{1}{4m^2c^2} \vec{p} [\vec{p}, q\phi] - \frac{i}{4m^2c^2} \vec{\sigma} (\vec{p} \times [\vec{p}, q\phi]) . \quad (284)$$

The first term on the right hand side gives a contribution to the so called *Darwin term* to be discussed later. The second term of equation (284) can further be arranged as

$$\begin{aligned} (\vec{p} \times [\vec{p}, q\phi])_i &= \varepsilon_{ijk} p_j [p_k, q\phi] = \underbrace{[\varepsilon_{ijk} p_j p_k, q\phi]}_{=0} - \varepsilon_{ijk} [p_j, q\phi] p_k \\ &= -([\vec{p}, q\phi] \times \vec{p})_i \\ &\Downarrow \\ -\frac{i}{4m^2c^2} \vec{\sigma} (\vec{p} \times [\vec{p}, q\phi]) &= \frac{\hbar}{4m^2c^2} \vec{\sigma} (\vec{\nabla}(q\phi) \times \vec{p}) . \end{aligned} \quad (285)$$

Since we deal with time-independent potentials, this correction can be written as

$$H_{\text{SO}} = -\frac{\hbar}{4m^2c^2} \vec{\sigma} (q\vec{\mathcal{E}} \times \vec{p}) , \quad (286)$$

which we identify with the *spin-orbit interaction*. Namely, for central potentials,

$$\vec{\nabla}\phi(r) = \frac{d\phi(r)}{dr} \frac{1}{r} \vec{r} , \quad (287)$$

the Hamiltonian (285) transforms to

$$H_{\text{SO}} = \frac{\hbar q}{4m^2c^2} \frac{1}{r} \frac{d\phi(r)}{dr} \vec{\sigma} (\vec{r} \times \vec{p}) = \frac{1}{2m^2c^2} \frac{1}{r} \frac{d(q\phi(r))}{dr} \vec{L} \vec{S} . \quad (288)$$

Using the orbital magnetic moment  $\vec{M}_L = \frac{q}{2m} \vec{L}$  and the spin magnetic moment  $\vec{M}_S = \frac{q}{m} \vec{S}$ , we arrive at

$$H_{\text{SO}} = \frac{1}{c^2 q^2} \frac{1}{r} \frac{d(q\phi(r))}{dr} \vec{M}_L \vec{M}_S . \quad (289)$$

describing the interaction between the orbital- and spin magnetic moments.

Adopting that the electron has a  $\frac{1}{2}$ -spin and a corresponding spin magnetic moment, the spin-orbit interaction has a classical interpretation. Suppose the electron moves in an electric field  $\vec{\mathcal{E}}$ . In its rest frame the electron experiences a magnetic field,

$$\vec{B} = \frac{1}{c^2} (\vec{\mathcal{E}} \times \vec{v}) = \frac{1}{mc^2} (\vec{\mathcal{E}} \times \vec{p}) , \quad (290)$$

where  $\vec{v}$  and  $\vec{p}$  are the velocity and the momentum of the electron in the laboratory frame, respectively, and we considered the non-relativistic limit,  $v \ll c$ . The Larmor energy of the spin magnetic moment in this magnetic field is then given by

$$H_L = -\vec{M}_S \vec{B} = -\frac{1}{m^2c^2} \vec{S} (q\vec{\mathcal{E}} \times \vec{p}) = -\frac{\hbar}{2m^2c^2} \vec{\sigma} (q\vec{\mathcal{E}} \times \vec{p}) , \quad (291)$$

which is twice as large in magnitude as the spin-orbit interaction in (286). The correct result can be obtained if one takes into account that the rest frame is not inertial. The acceleration of the

charged particle,  $\vec{a} = q\vec{\mathcal{E}}/m$ , leads to the Thomas precession of the spin with an angular velocity within the non-relativistic limit,

$$\vec{\omega}_T = \frac{1}{2c^2}(\vec{a} \times \vec{v}) = \frac{1}{2m^2c^2}(q\vec{\mathcal{E}} \times \vec{p}), \quad (292)$$

and a corresponding interaction energy,

$$H_T = \vec{\omega}_T \vec{S} = \frac{1}{2m^2c^2}\vec{S}(q\vec{\mathcal{E}} \times \vec{p}) = \frac{\hbar}{4m^2c^2}\vec{\sigma}(q\vec{\mathcal{E}} \times \vec{p}). \quad (293)$$

The sum of  $H_L$  and  $H_T$  is then clearly identical with  $H_{SO}$  in (286).

Let us calculate the correction to the energy levels of the hydrogen atom due to spin-orbit interaction. First we rewrite the perturbation operator,

$$W = \frac{ke^2}{2m^2c^2} \frac{1}{r^3} \vec{L} \vec{S} \quad (294)$$

as

$$W = \frac{ke^2}{4m^2c^2} \frac{1}{r^3} (J^2 - L^2 - S^2), \quad (295)$$

and we consider the unperturbed wavefunctions in the common eigenbasis of the operators  $J^2$ ,  $J_z$ ,  $L^2$  and  $S^2$ ,

$$|n, \ell, j, m_j\rangle = \sum_{m_s = \pm \frac{1}{2}} C(\ell m_\ell, \frac{1}{2} m_s; j, m_j) |n \ell m_\ell\rangle |\frac{1}{2}, m_s\rangle, \quad (296)$$

where  $C(\ell m_\ell, \frac{1}{2} m_s; j, m_j)$  are the Clebsch-Gordan coefficients and  $j = \ell \pm \frac{1}{2}$ . Clearly, just like  $|n \ell m_\ell\rangle |\frac{1}{2}, m_s\rangle$ , the wavefunctions  $|n, \ell, j, m_j\rangle$  form a basis in the eigenspace corresponding to  $E_n^{(0)}$ . The first order energy corrections can then be evaluated as,

$$\delta E_{n, \ell, j, m_j}^{(1)} = \frac{\hbar^2 ke^2}{4m^2c^2} \left\langle \frac{1}{r^3} \right\rangle_{n\ell} \left( j(j+1) - \ell(\ell+1) - \frac{3}{4} \right). \quad (297)$$

Using the expectation value of  $1/r^3$  for the eigenstates of the H atom,

$$\left\langle \frac{1}{r^3} \right\rangle_{n\ell} = \frac{1}{n^3 a_0^3} \frac{1}{(\ell + \frac{1}{2}) \ell (\ell + 1)}, \quad (298)$$

we get,

$$\delta E_{n, \ell, j, m_j}^{(1)} = \frac{\hbar^2 ke^2}{4m^2c^2 a_0^3} \frac{1}{n^4} n \frac{j(j+1) - \ell(\ell+1) - \frac{3}{4}}{(\ell + \frac{1}{2}) \ell (\ell + 1)} \quad (299)$$

$$= \frac{2(E_n^{(0)})^2}{mc^2} n \frac{j(j+1) - \ell(\ell+1) - \frac{3}{4}}{(2\ell + 1) \ell (\ell + 1)}, \quad (300)$$

where we used

$$a_0 = \frac{\hbar^2}{mke^2} \rightarrow \frac{\hbar^2 ke^2}{4m^2c^2 a_0^3} \frac{1}{n^4} = \frac{1}{mc^2} \frac{(ke^2)^2}{4a_0^2 n^4} = \frac{(E_n^{(0)})^2}{mc^2}. \quad (301)$$

After some algebra,

$$\begin{aligned} \frac{j(j+1) - \ell(\ell+1) - \frac{3}{4}}{(2\ell + 1) \ell (\ell + 1)} & \underset{j=\ell+\frac{1}{2}}{=} \frac{(\ell + \frac{1}{2})(\ell + \frac{3}{2}) - \ell(\ell+1) - \frac{3}{4}}{(2\ell + 1) \ell (\ell + 1)} = \frac{1}{(2\ell + 1) (\ell + 1)} \\ & = \frac{1}{\ell + \frac{1}{2}} - \frac{1}{\ell + 1} = \frac{1}{\ell + \frac{1}{2}} - \frac{1}{j + \frac{1}{2}}, \end{aligned} \quad (302)$$

$$\begin{aligned} \frac{j(j+1) - \ell(\ell+1) - \frac{3}{4}}{(2\ell+1)\ell(\ell+1)} & \underset{j=\ell-\frac{1}{2}}{=} \frac{(\ell-\frac{1}{2})(\ell+\frac{1}{2}) - \ell(\ell+1) - \frac{3}{4}}{(2\ell+1)\ell(\ell+1)} = -\frac{1}{(2\ell+1)\ell} \\ & = \frac{1}{\ell+\frac{1}{2}} - \frac{1}{\ell} = \frac{1}{\ell+\frac{1}{2}} - \frac{1}{j+\frac{1}{2}}, \end{aligned} \quad (303)$$

we end up with

$$\delta E_{n,\ell,j,m_j}^{(1)} = \frac{2(E_n^{(0)})^2}{mc^2} \left( \frac{n}{\ell+\frac{1}{2}} - \frac{n}{j+\frac{1}{2}} \right). \quad (304)$$

Together with the relativistic kinetic energy correction of the energy levels, this gives

$$\begin{aligned} \delta E_{n,\ell,j,m_j}^{(1)} &= -\frac{2(E_n^{(0)})^2}{mc^2} \left( \frac{n}{\ell+1/2} - \frac{3}{4} \right) + \frac{2(E_n^{(0)})^2}{mc^2} \left( \frac{n}{\ell+\frac{1}{2}} - \frac{n}{j+\frac{1}{2}} \right) \\ &= -\frac{2(E_n^{(0)})^2}{mc^2} \left( \frac{n}{j+1/2} - \frac{3}{4} \right), \end{aligned} \quad (305)$$

which agrees with the first order term of the  $1/c^2$  expansion of the eigenenergies calculated directly from the Dirac equation. However, the above consideration does not apply for  $\ell = 0$  because in that case  $\langle \frac{1}{r^3} \rangle_{n\ell} \sim \frac{1}{\ell}$  is infinite. Assuming an atomic nucleus of finite size,  $\langle \frac{1}{r^3} \rangle_{n\ell}$  becomes finite, but the matrixelement of the  $\vec{L}\vec{S}$  operator for  $\ell = 0$  is zero, hence the spin-orbit coupling doesn't influence the energy of the  $s$ -orbitals in first-order perturbation theory. This means that taking into account the first-order perturbation correction due to the relativistic kinetic energy term, for  $\ell = 0$  we get

$$\delta E_{n,0,j=\frac{1}{2},m_j}^{(1)} = -\frac{2(E_n^{(0)})^2}{mc^2} \left( 2n - \frac{3}{4} \right), \quad (306)$$

whereas, from (305) we deduce the correct result,

$$\delta E_{n,0,j=\frac{1}{2},m_j}^{(1)} = -\frac{2(E_n^{(0)})^2}{mc^2} \left( n - \frac{3}{4} \right). \quad (307)$$

It is tempting that there should exist a third type of  $\frac{1}{c^2}$  correction due to relativistic effects acting only on the  $\ell = 0$  states of the hydrogen atom.

### 3.4.3 Darwin term

So far, we have not paid attention to the condition that the two-component wavefunction describing the state of the electron within the non-relativistic approach extended by  $1/c^2$  corrections should preserve the norm of the four-component wavefunction  $\psi$ ,

$$\int d^3r \psi^\dagger \psi = \int d^3r (\chi^\dagger \chi + \varphi^\dagger \varphi). \quad (308)$$

Neglecting the effect of magnetic field as before, the norm of the small component can be expressed in first order of  $\frac{1}{c^2}$  as

$$\begin{aligned} \int d^3r \varphi^\dagger \varphi &\simeq \frac{1}{4m^2c^2} \int d^3r \chi^\dagger (\vec{\sigma}\vec{p}) (\vec{\sigma}\vec{p}) \chi \\ &\simeq \frac{1}{4m^2c^2} \int d^3r \chi^\dagger p^2 \chi, \end{aligned} \quad (309)$$

thus the norm of the wavefunction  $\psi$  can be approximated as

$$\int d^3r \psi^\dagger \psi \simeq \int d^3r \chi^\dagger \left(1 + \frac{p^2}{4m^2c^2}\right) \chi. \quad (310)$$

Let us introduce the modified large component  $\chi_S$  the norm of which equals that of  $\psi$  within the above approximation,

$$\int d^3r \chi_S^\dagger \chi_S = \int d^3r \chi^\dagger \left(1 + \frac{p^2}{4m^2c^2}\right) \chi. \quad (311)$$

This leads to the relationship between  $\chi_S$  and  $\chi$ ,

$$\chi_S = \left(1 + \frac{p^2}{8m^2c^2}\right) \chi \quad \Longrightarrow \quad \chi = \left(1 - \frac{p^2}{8m^2c^2}\right) \chi_S, \quad (312)$$

Thus working with the modified large component  $\chi_S$ , the norm of the eliminated small component  $\varphi$  is taken into account in first order of  $1/c^2$ . Note that the above procedure is equivalent with a *unitary transformation* between the four-component wavefunctions  $\psi$  and  $\begin{pmatrix} \chi_S \\ 0 \end{pmatrix}$ .

Let us summarize the Hamiltonian derived in the previous sections by  $H'$ ,

$$H' = H_P + H_M + H_{SO} - \frac{1}{4m^2c^2} \vec{p} [\vec{p}, q\phi], \quad (313)$$

where we included the first term on the right hand side of equation (284). Employing (312) to the eigenvalue equation,

$$E' \chi = H' \chi, \quad (314)$$

we obtain

$$E' \chi_S = \left(1 + \frac{p^2}{8m^2c^2}\right) H' \left(1 - \frac{p^2}{8m^2c^2}\right) \chi_S. \quad (315)$$

In first order of  $1/c^2$  the above equation can be transformed to

$$E' \chi_S \simeq \left(H' + \left[\frac{p^2}{8m^2c^2}, H_P\right]\right) \chi_S \simeq \left(H' + \frac{1}{8m^2c^2} [p^2, q\phi]\right) \chi_S, \quad (316)$$

where we again neglected the contributions from the magnetic field. Treating the new term together with the last term on the right hand side of equation (313), we get the third  $1/c^2$  correction to the non-relativistic Hamilton operator, called the *Darwin term*,

$$\begin{aligned} H_D &= -\frac{1}{4m^2c^2} \vec{p} [\vec{p}, q\phi] + \frac{1}{8m^2c^2} [p^2, q\phi] \\ &= \frac{1}{8m^2c^2} ([\vec{p}, q\phi] \vec{p} - \vec{p} [\vec{p}, q\phi]) \\ &= -\frac{1}{8m^2c^2} [\vec{p}, [\vec{p}, q\phi]] = \frac{\hbar^2}{8m^2c^2} [\vec{\nabla}, [\vec{\nabla}, q\phi]] \\ &= \frac{\hbar^2}{8m^2c^2} \Delta(q\phi). \end{aligned} \quad (317)$$

The Darwin term has no classical interpretation. The origin of this term can be traced back to the phenomenon that the electron cannot be localized to a point-like particle but it rapidly oscillates on a length scale of the reduced Compton wavelength,  $\lambda_C = \frac{1}{\kappa} = \frac{\hbar}{mc}$  (Zitterbewegung). The Zitterbewegung follows from the fact that the wavepacket state of a localized electron is the

superposition of positive and negative energy components, and the rapid oscillations both in space and time arise due to the interference of these contributions. Such a phenomenon is indeed out of the scope of a classical theory. Thus, the electron experiences the average of the potential on a length scale of  $\lambda_C$ ,

$$\begin{aligned}\langle V(\vec{r}) \rangle &\simeq V(\vec{r}_0) + \langle \delta x_i \rangle \partial_i V(\vec{r}_0) + \frac{1}{2} \langle \delta x_i \delta x_j \rangle \partial_i \partial_j V(\vec{r}_0) \\ &\simeq V(\vec{r}_0) + \frac{1}{2} \frac{\lambda_C^2}{3} \Delta V(\vec{r}_0) = V(\vec{r}_0) + \frac{\hbar^2}{6m^2 c^2} \Delta V(\vec{r}_0),\end{aligned}$$

where we assumed  $\langle \delta x_i \rangle = 0$  and  $\langle \delta x_i \delta x_j \rangle \simeq \delta_{ij} \langle \delta r^2 \rangle / 3 \simeq \delta_{ij} \lambda_C^2 / 3$ . This simple estimation gives a very good qualitative agreement with the Darwin term derived in (317).

Since for the Coulomb potential,  $q\phi = -ke^2/r \rightarrow \Delta(q\phi) = 4\pi\delta(\vec{r})$ ,

$$H_D(\vec{r}) = \frac{ke^2 \hbar^2 \pi}{2m^2 c^2} \delta(\vec{r}), \quad (318)$$

the first order correction of the energy levels of the H atom is given by

$$\delta E_{n\ell m}^{(1)} = \int d^3r \psi_{n\ell m}^*(\vec{r}) H_D(\vec{r}) \psi_{n\ell m}(\vec{r}) = \frac{ke^2 \hbar^2 \pi}{2m^2 c^2} |\psi_{n\ell m}(0)|^2. \quad (319)$$

This gives a non-zero contribution only for the  $s$ -orbitals,

$$\psi_{n00}(0) = \frac{1}{\sqrt{\pi}} \frac{1}{(na_0)^{3/2}} \rightarrow |\psi_{n00}(0)|^2 = \frac{1}{\pi} \frac{1}{n^3 a_0^3},$$

and from that

$$\delta E_{n00}^{(1)} = \frac{ke^2 \hbar^2}{2m^2 c^2} \frac{1}{n^3 a_0^3} = \frac{2}{mc^2} \left( \frac{ke^2}{2a_0 n^2} \right)^2 \underbrace{\frac{\hbar^2}{mke^2 a_0}}_{=1} n = \frac{2(E_n^{(0)})^2}{mc^2} n. \quad (320)$$

Adding this correction to (306), we indeed get the correct result (307)!

In summary, the  $1/c^2$  expansion of the Dirac equation results in to the eigenvalue equation,

$$(H_P + H_M + H_{SO} + H_D) \chi_S = (E - mc^2) \chi_S, \quad (321)$$

where  $\chi_S$  is the normalized two-component wavefunction,

$$H_P = \frac{1}{2m} \vec{K}^2 - \frac{\hbar q}{2m} \vec{B} \vec{\sigma} + q\phi, \quad (322)$$

is the Pauli-Schrödinger Hamiltonian,

$$H_M = -\frac{1}{8m^3 c^2} p^4 \quad (323)$$

is the relativistic kinetic energy correction,

$$H_{SO} = -\frac{\hbar q}{4m^2 c^2} \vec{\sigma} \left( \vec{\mathcal{E}} \times \vec{p} \right) \quad (324)$$

is the spin-orbit interaction and

$$H_D = \frac{\hbar^2}{8m^2 c^2} \Delta(q\phi) \quad (325)$$

is the Darwin term.

### 3.4.4 Spin magnetization probability current density

Finally, we derive the non-relativistic limit of the probability current density. We will consider the leading term only, therefore, it is sufficient to use the non-relativistic approximation for the small component,

$$\varphi \simeq \frac{\vec{\sigma} \vec{K} \chi}{2mc}, \quad (326)$$

and we also neglect the normalization of the large component. Consequently, the current density can be expressed as

$$\begin{aligned} \vec{j} &= c\psi^\dagger \vec{\alpha} \psi = c(\chi^\dagger, \varphi^\dagger) \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \begin{pmatrix} \chi \\ \varphi \end{pmatrix} = c(\chi^\dagger \vec{\sigma} \varphi + \varphi^\dagger \vec{\sigma} \chi) \\ &= \frac{1}{2m} \left[ \chi^\dagger \vec{\sigma} (\vec{\sigma} \vec{K} \chi) + ((\vec{K} \chi)^\dagger \vec{\sigma}) \vec{\sigma} \chi \right]. \end{aligned} \quad (327)$$

By using the identities,

$$\begin{aligned} \sigma_i (\vec{\sigma} \vec{K} \chi) &= \sigma_i \sigma_j K_j \chi = (\delta_{ij} + i\varepsilon_{ijk} \sigma_k) K_j \chi = K_i \chi + i(\vec{K} \times \vec{\sigma} \chi)_i \\ ((\vec{K} \chi)^\dagger \vec{\sigma}) \sigma_i &= (K_j \chi)^\dagger \sigma_j \sigma_i = (K_j \chi)^\dagger (\delta_{ij} - i\varepsilon_{ijk} \sigma_k) = (K_i \chi)^\dagger - i((\vec{K} \chi)^\dagger \times \vec{\sigma})_i \end{aligned}$$

we can proceed as

$$\vec{j} = \frac{1}{2m} \left( \chi^\dagger (\vec{K} \chi) + (\vec{K} \chi)^\dagger \chi \right) + \frac{i}{2m} \left( \chi^\dagger (\vec{K} \times \vec{\sigma} \chi) - (\vec{K} \chi)^\dagger \times \vec{\sigma} \chi \right). \quad (328)$$

The first term gives the well-known non-relativistic expression of the current density:

$$\vec{j}_{\text{nr}} \equiv \frac{1}{m} \text{Re} \left( \chi^\dagger (\vec{K} \chi) \right) = \frac{\hbar}{2im} \left[ \chi^\dagger (\vec{\nabla} \chi) - (\vec{\nabla} \chi^\dagger) \chi \right] - \frac{q}{m} \vec{A} \chi^\dagger \chi. \quad (329)$$

The second term can be evaluated by using

$$\begin{aligned} \chi^\dagger (\vec{K} \times \vec{\sigma} \chi) &= \frac{\hbar}{i} \chi^\dagger (\vec{\nabla} \times \vec{\sigma} \chi) - \chi^\dagger (q \vec{A} \times \vec{\sigma} \chi) \\ (\vec{K} \chi)^\dagger \times \vec{\sigma} \chi &= -\frac{\hbar}{i} (\vec{\nabla} \chi^\dagger) \times \vec{\sigma} \chi - \chi^\dagger (q \vec{A} \times \vec{\sigma} \chi), \end{aligned}$$

and it yields a new contribution:

$$\begin{aligned} \vec{j}_{\text{spin}} &\equiv \frac{i}{2m} \left( \chi^\dagger (\vec{K} \times \vec{\sigma} \chi) - (\vec{K} \chi)^\dagger \times \vec{\sigma} \chi \right) \\ &= \frac{\hbar}{2m} \left( \chi^\dagger (\vec{\nabla} \times \vec{\sigma} \chi) + (\vec{\nabla} \chi^\dagger) \times \vec{\sigma} \chi \right) \\ &= \frac{\hbar}{2m} \vec{\nabla} \times (\chi^\dagger \vec{\sigma} \chi) = \frac{1}{m} \text{curl} (\chi^\dagger \vec{S} \chi) \end{aligned} \quad (330)$$

that means it is proportional to the curl of the spin-density. It is obvious that the divergence of  $\vec{j}_{\text{spin}}$  is zero, this is why the continuity equation we derived within the non-relativistic theory was valid though  $\vec{j}_{\text{spin}}$  has not appeared in the probability current density. With the operator of the spin magnetic moment  $\vec{M}_S = \frac{q}{m} \vec{S}$  and the spin-magnetization density,

$$\vec{M}_S(\vec{r}, t) = \chi^\dagger(\vec{r}, t) \vec{M}_S \chi(\vec{r}, t), \quad (331)$$

the corresponding charge current reads as

$$q \vec{j}_{\text{spin}}(\vec{r}, t) = \text{curl} \vec{M}_S(\vec{r}, t). \quad (332)$$