Relativistic Quantum Mechanics
Lecture notes

László Szunyogh and Bendegúz Nyári

5th December 2021

Contents

1 Special relativity in a nutshell

1.1 Minkowski vectors and Lorentz transformations

The *contravariant* coordinates of the vectors of the four-dimensional Minkowski space are denoted by upper indices,

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

The squared distance in four-space is given by

$$ds^2 = c^2t^2 - d\vec{x}^2 = dx^\mu g_{\mu\nu}\, dx^\nu \ , \quad (2)$$

where we introduced the metric tensor,

$$g_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \ . \quad (3)$$

Since ds^2 in Eq. (2) 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 *ligthlike* vectors with zero norm represented with the light cone (world-lines).

Let's introduce the inverse of the metric tensor $g^{\mu\nu}$ as

$$g^{\mu\nu} g_{\nu\lambda} = \delta^{\mu}_{\cdot\lambda} \ , \quad (4)$$
$$g_{\mu\nu} g^{\nu\lambda} = \delta_{\mu}^{\cdot\lambda} \ , \quad (5)$$

where

$$\delta^{\mu}_{\cdot\lambda} = \delta_{\mu}^{\cdot\lambda} = \begin{cases} 1 & \text{if} \quad \mu = \lambda \\ 0 & \text{if} \quad \mu \neq \lambda \end{cases} \ . \quad (6)$$

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

$$g^{\mu\nu} = g_{\mu\nu} \ . \quad (7)$$

It is worth to define the *covariant* components of the Minkowski vectors,

$$x_\mu \equiv g_{\mu\nu}\, x^\nu = (ct, -\vec{x}) \ , \quad (8)$$

yielding the inverse relation,

$$x^\mu = \delta^{\mu}_{\cdot\nu}\, x^\nu = g^{\mu\lambda} g_{\lambda\nu} x^\nu = g^{\mu\lambda}\, x_\lambda \ . \quad (9)$$

The squared distance can then be simply expressed as

$$ds^2 = dx_\mu dx^\mu = dx^\mu dx_\mu \ . \quad (10)$$

We demand that ds^2 should be invariant from the choice of inertial system. Different inertial systems are related by real valued linear transformations of the Minkowski space,

$$x'^\mu = \Lambda^{\mu}_{\cdot\nu}\, x^\nu + a^\mu \quad (11)$$

or in matrix-vector notation,

$$\Lambda = \{\Lambda^{\mu}_{\cdot\,\nu}\} \Longrightarrow x' = \Lambda\, x + a\,, \tag{12}$$

called Poincaré transformations. Invariance means,

$$dx'^{\mu} g_{\mu\tau} dx'^{\tau} = \Lambda^{\mu}_{\cdot\,\nu}\, dx^{\nu} g_{\mu\tau} \Lambda^{\tau}_{\cdot\,\lambda}\, dx^{\lambda} = dx^{\nu} \Lambda^{\mu}_{\cdot\,\nu}\, g_{\mu\tau} \Lambda^{\tau}_{\cdot\,\lambda}\, dx^{\lambda} = dx^{\nu} g_{\nu\lambda} dx^{\lambda}\,, \tag{13}$$

implying the identity

$$\Lambda^{\mu}_{\cdot\,\nu}\, g_{\mu\tau}\, \Lambda^{\tau}_{\cdot\,\lambda} = g_{\nu\lambda}\ . \tag{14}$$

Introducing the trasponse of Λ,

$$\left(\Lambda^{T}\right)_{\nu}^{\cdot\,\mu} \equiv \Lambda^{\mu}_{\cdot\,\nu}\,, \tag{15}$$

the previous identity can be written as

$$\left(\Lambda^{T}\right)_{\nu}^{\cdot\,\mu}\ g_{\mu\tau}\, \Lambda^{\tau}_{\cdot\,\lambda} = g_{\nu\lambda} \tag{16}$$

or in brief notation,

$$\Lambda^{T} g\, \Lambda = g\,. \tag{17}$$

The Lorentz transformation have the following properties:

1. $\Lambda^{\mu}_{\cdot\,\nu} = \delta^{\mu}_{\cdot\,\nu}$ is a solution: $\delta_{\mu}^{\cdot\,\tau} g_{\tau\sigma} \delta^{\sigma}_{\cdot\,\nu} = g_{\mu\nu}$
2. $\det \Lambda^{T} g \Lambda = (\det \Lambda)^{2} \det g = \det g \Longrightarrow \det \Lambda = \pm 1$
3. If Λ satisfies Eq. (17), then Λ^{-1} does it, too: $g = \left(\Lambda^{T}\right)^{-1} g \Lambda^{-1} = (\Lambda^{-1})^{T} g \Lambda^{-1}$
4. If Λ_1 and Λ_2 satisfy Eq. (17), then $\Lambda_1\Lambda_2$ also satisfies Eq. (17):
$(\Lambda_1\Lambda_2)^{T} g \Lambda_1\Lambda_2 = \Lambda_2^{T}\Lambda_1^{T} g \Lambda_1\Lambda_2 = \Lambda_2^{T} g \Lambda_2 = g$

From the above properties it follows that the real valued matrices Λ satisfying Eq. (17) form a group (Lorentz group). As what follows we deal with the Lorentz tranformations for which $a^{\mu} = 0$.

The infinitesimal Lorentz tansformation defined as

$$\Lambda^{\mu}_{\cdot\,\nu} = \delta^{\mu}_{\cdot\,\nu} + \omega^{\mu}_{\cdot\,\nu} \longleftrightarrow \Lambda = I + \omega\,. \tag{18}$$

Eq. (17) then implies for the matrix ω,

$$\left(I + \omega^{T}\right) g\, (I + \omega) = g \underset{\text{in 1st order}}{\Longrightarrow} \omega^{T} g = -g\omega\,. \tag{19}$$

Since

$$\left(\omega^{T}\right)_{\mu}^{\cdot\,\tau} g_{\tau\nu} = \omega^{\tau}_{\cdot\,\mu}\, g_{\tau\nu} = g_{\nu\tau}\, \omega^{\tau}_{\cdot\,\mu} = \omega_{\nu\mu} \quad \text{and} \quad g_{\mu\tau}\omega^{\tau}_{\cdot\,\nu} = \omega_{\mu\nu} \tag{20}$$

$$\Downarrow$$

$$\omega_{\mu\nu} = -\omega_{\nu\mu}\,, \tag{21}$$

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

The Lorentz transform of the covariant vectors can be obtained as

$$x'_{\mu} = g_{\mu\nu}\, x'^{\nu} = g_{\mu\nu} \Lambda^{\nu}_{\cdot\,\tau}\, x^{\tau} = g_{\mu\nu} \Lambda^{\nu}_{\cdot\,\tau} g^{\tau\sigma} x_{\sigma} = \overline{\Lambda}_{\mu}^{\cdot\,\sigma} x_{\sigma} \tag{22}$$

where

$$\overline{\Lambda}_{\mu}^{\boldsymbol{\cdot}\sigma} \equiv g_{\mu\nu}\Lambda^{\nu}_{\boldsymbol{\cdot}\tau}g^{\tau\sigma} \tag{23}$$

or

$$\overline{\Lambda} = g\,\Lambda\,g^{-1}\,. \tag{24}$$

The matrix of transformation of the covariant vectors $\overline{\Lambda}$ satisfies the following identity:

$$\left(\Lambda^T\right)_{\mu}^{\boldsymbol{\cdot}\tau}\,\overline{\Lambda}_{\tau}^{\boldsymbol{\cdot}\nu} = \Lambda^{\tau}_{\boldsymbol{\cdot}\mu}\,g_{\tau\sigma}\Lambda^{\sigma}_{\boldsymbol{\cdot}\rho}\,g^{\rho\nu} = g_{\mu\rho}\,g^{\rho\nu} = \delta_{\mu}^{\boldsymbol{\cdot}\nu} \tag{25}$$

or

$$\Lambda^T\,\overline{\Lambda} = \Lambda^T g\Lambda g^{-1} = g\,g^{-1} = I \Longrightarrow \overline{\Lambda} = \left(\Lambda^{-1}\right)^T\,. \tag{26}$$

We can show that the four-derivative,

$$\partial_{\mu} = \frac{\partial}{\partial x^{\mu}} \Longrightarrow \partial_{\mu}x^{\nu} = \delta_{\mu}^{\boldsymbol{\cdot}\nu} \tag{27}$$

$$\partial_{\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, \vec{\nabla}\right) \tag{28}$$

is a covariant vector:

$$\partial'_{\mu} = \frac{\partial}{\partial x'^{\mu}} = \frac{\partial x^{\nu}}{\partial x'^{\mu}}\frac{\partial}{\partial x^{\nu}} = \left(\Lambda^{-1}\right)^{\nu}_{\boldsymbol{\cdot}\mu}\,\partial_{\nu} = \left[\left(\Lambda^{-1}\right)^T\right]_{\mu}^{\boldsymbol{\cdot}\nu}\,\partial_{\nu} = \overline{\Lambda}_{\mu}^{\boldsymbol{\cdot}\nu}\,\partial_{\nu}\,. \tag{29}$$

The contravariant derivative is then given by

$$\partial^{\mu} = \left(\frac{1}{c}\partial_t, -\vec{\nabla}\right)\,, \tag{30}$$

while the d'Alambert operator,

$$\partial_{\mu}\partial^{\mu} = \frac{1}{c^2}\partial_t^2 - \vec{\nabla}^{\,2} \equiv \Box \tag{31}$$

is invariant under Lorentz transformations.

1.2 Relativistic kinematics

The four-velocity is defined as

$$u^{\mu} = \frac{\mathrm{d}x^{\mu}}{\mathrm{d}\tau} = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\left(c, \vec{v}\right)\,, \tag{32}$$

where τ is the proper time,

$$d\tau = \frac{1}{c}ds = dt\sqrt{1-\frac{v^2}{c^2}}\,, \tag{33}$$

with

$$\vec{v} = \frac{\mathrm{d}\vec{x}\,(t)}{\mathrm{d}t}\,. \tag{34}$$

It is simple to prove that the norm of the four-velocity is Lorentz invariant,

$$u^{\mu}g_{\mu\nu}u^{\nu} = u^{\mu}u_{\mu} = \frac{c^2-v^2}{1-\frac{v^2}{c^2}} = c^2\,. \tag{35}$$

The energy and the momentum also form a four-vector, called as four-momentum,

$$p^\mu = mu^\mu = \left(\frac{E}{c}, \vec{p}\right) , \tag{36}$$

where m is the rest mass of the particle. The momentum and the energy are defined as

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

and

$$E = \frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}} , \tag{37}$$

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

$$p^\mu p_\mu = m^2c^2 , \tag{38}$$

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

$$E^2 - c^2\vec{p}^2 = \left(mc^2\right)^2 . \tag{39}$$

1.3 Electrodynamics

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

$$A^\mu = \left(\frac{\phi}{c}, \vec{A}\right) , \tag{40}$$

from which we can define the field strength tensor,

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{41}$$

$$F^{i0} = \frac{\mathcal{E}_i}{c} , \qquad F^{ij} = -\epsilon_{ijk}B_k . \tag{42}$$

The four-current is defined as

$$j^\mu = \left(c\rho, \vec{j}\right) , \tag{43}$$

satisfying the continuity equation,

$$\partial_\mu j^\mu = 0 , \tag{44}$$

or in the conventional form,

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\vec{j} = 0 . \tag{45}$$

The Maxwell equations can then be summarized as

$$\partial_\mu F^{\mu\nu} = \mu_0 j^\nu , \tag{46}$$

$$\epsilon_{\lambda\tau\mu\nu}\partial^\tau F^{\mu\nu} = 0 , \tag{47}$$

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 we use the *minimal substitution*,

$$p^\mu \to K^\mu = p^\mu - qA^\mu$$

$$\Downarrow$$

$$K^{\mu} = \left(\frac{E - q\phi}{c}, \vec{p} - q\vec{A}\right) \tag{48}$$

$$K_{\mu} = \left(\frac{E - q\phi}{c}, -(\vec{p} - q\vec{A})\right) \tag{49}$$

the length of which is Lorentz invariant,

$$K^{\mu}K_{\mu} = m^2 u^{\mu} u_{\mu} = m^2 c^2 \tag{50}$$

from which the energy-momentum dispersion,

$$\frac{(E - q\phi)^2}{c^2} - \left(\vec{p} - q\vec{A}\right)^2 = m^2 c^2 \tag{51}$$

$$\Downarrow$$

$$E = \sqrt{m^2 c^4 + c^2 \left(\vec{p} - q\vec{A}\right)^2} + q\phi\,. \tag{52}$$

can be obtained.

2 First steps to Relativistic Quantum Mechanics

2.1 Operators in coordinate representation

Relativistic quantum mechanics should provide with a Lorentz invariant wave equation reflecting the dispersion relation (39). 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

$$x_i \to x_i \cdot \quad p_i \to \frac{\hbar}{i}\partial_i = -i\hbar\partial_i . \tag{53}$$

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

$$x^\mu \to x^\mu \cdot \quad p^\mu \to i\hbar\partial^\mu = (\frac{i\hbar\partial_t}{c}, -i\hbar\vec{\bigtriangledown}) = \left(p^0, \vec{p}\right) . \tag{54}$$

The corresponding covariant four-vector operators are

$$x_\mu \to x_\mu \cdot \quad p_\mu \to i\hbar\partial_\mu = (\frac{i\hbar\partial_t}{c}, i\hbar\vec{\bigtriangledown}) = (p_0, -\vec{p}) . \tag{55}$$

Obviously, the following commutation relations apply,

$$[p_\mu, x^\nu] = i\hbar\ [\partial_\mu, x^\nu] = i\hbar[\frac{\partial}{\partial x^\mu}, x^\nu] = i\hbar\, \delta_\mu^{\cdot\nu} . \tag{56}$$

The covariant (Lorentz invariant) forms of these commutator relations are

$$[p^\mu, x^\nu] = i\hbar\, g^{\mu\nu} , \tag{57}$$

or

$$[p_\mu, x_\nu] = i\hbar\, g_{\mu\nu} , \tag{58}$$

which are proper extensions of the commutators postulated in non-relativistic quantum mechanics,

$$\left[p^i, x^j\right] = i\hbar\, g^{ij} = \left\{ \begin{array}{lll} \frac{\hbar}{i} & \text{if} & i = j \\ 0 & \text{if} & i \neq j \end{array} \right. . \tag{59}$$

As an excersize, let's prove that the commutation relations (57) are Lorentz ivariant:

$$C^{\mu\nu} \equiv [p^\mu, x^\nu] \to \mathsf{C} = \imath\hbar g^{-1} \tag{60}$$

$$C'^{\mu\nu} = [p'^\mu, x'^\nu] = \Lambda^\mu_{\cdot\tau}\Lambda^\nu_{\cdot\lambda}\left[p^\tau, x^\lambda\right] = \Lambda^\mu_{\cdot\tau}\left[p^\tau, x^\lambda\right]\left(\Lambda^T\right)_\lambda^{\cdot\nu} = \Lambda^\mu_{\cdot\tau}C^{\tau\lambda}\left(\Lambda^T\right)_\lambda^{\cdot\nu} \tag{61}$$

or, in brief,

$$C' = \Lambda C \Lambda^T = \imath\hbar\Lambda g^{-1}\Lambda^T = \imath\hbar g^{-1} g\Lambda g^{-1}\Lambda^T = \imath\hbar g^{-1}\overline{\Lambda}\Lambda^T = \imath\hbar g^{-1} . \tag{62}$$

We easily can verify the commutator of the timelike coodinates,

$$\left[p^0, x^0\right] = i\hbar\,[\partial_t, t] = i\hbar . \tag{63}$$

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

$$E = \imath\hbar \frac{\partial}{\partial t} . \tag{64}$$

The kinetic momentum operators are then given by

$$K^\mu = p^\mu - qA^\mu = \imath\hbar\partial^\mu - qA^\mu = \left(\frac{1}{c} \left(\imath\hbar\partial_t - q\phi \right), -\imath\hbar\vec{\nabla} - q\vec{A} \right) , \tag{65}$$

$$K_\mu = p_\mu - qA_\mu = \imath\hbar\partial_\mu - qA_\mu = \left(\frac{1}{c} \left(\imath\hbar\partial_t - q\phi \right), \imath\hbar\vec{\nabla} + q\vec{A} \right) . \tag{66}$$

2.2 The Klein–Gordon equation

If we substitute the operators (65) and (66) into the Lorentz invariant expression (50) and let this operator act on a wavefunction we obtain with the *Klein–Gordon equation*,

$$\left[K^\mu K_\mu - m^2 c^2 \right] \psi\left(\vec{r}, t\right) = 0 \tag{67}$$

$$\Downarrow$$

$$\left[\frac{1}{c^2} \left(\imath\hbar\partial_t - q\phi \right)^2 - \left(\imath\hbar\vec{\nabla} + q\vec{A} \right)^2 \right] \psi\left(\vec{r}, t\right) = m^2 c^2 \psi\left(\vec{r}, t\right) . \tag{68}$$

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

$$\left[\Box + \kappa^2 \right] \psi\left(\vec{r}, t\right) = 0 \tag{69}$$

where $\kappa = \frac{mc}{\hbar}$ is the Compton wavelength and $\Box = \frac{1}{c^2}\partial_t^2 - \vec{\nabla}^{\,2}$ is the d'Alambert operator as introduced already.

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

$$\psi\left(\vec{r}, t\right) = \psi\left(\vec{r}\right) e^{-\frac{\imath}{\hbar}Et} \tag{70}$$

leading to the *stationary Klein–Gordon equation*,

$$\left[\left(i\hbar\vec{\nabla} + q\vec{A} \right)^2 - \frac{\left(E - q\phi \right)^2}{c^2} + m^2 c^2 \right] \psi\left(\vec{r}\right) = 0 \quad . \tag{71}$$

For a free partical, the solution of the equation

$$\left[-\hbar^2 \Delta - \frac{E^2}{c^2} + m^2 c^2 \right] \psi\left(\vec{r}\right) = 0 \tag{72}$$

is a plane wave,

$$\psi\left(\vec{r}\right) = A\, e^{\frac{\imath}{\hbar}\vec{p}\vec{r}} \tag{73}$$

which by substitution yields

$$c^2 p^2 + m^2 c^4 - E^2 = 0 , \tag{74}$$

in agreement with Eq. (39). The energy of the free particle with momentum $\vec{p}$ can then take

$$E = \pm\sqrt{m^2c^4 + c^2p^2} \tag{75}$$

allowing for $E \geq mc^2$ or $E \leq -mc^2$. The appearance of the negative energy solutions is a new feature related 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 $\left(\vec{A}\left(\vec{r}\right) = 0,\ \phi\left(\vec{r}\right) = -Ze^2/r\right)$. Expanding the eigenenergies up to first order in $1/c^2$, we obtain

$$E_{n\ell} \simeq mc^2 - \frac{m\left(Ze^2\right)^2}{2\hbar^2n^2} + \frac{m\left(Ze^2\right)^4}{4\hbar^4n^4c^2}\left(\frac{3}{2} - \frac{4n}{2\ell+1}\right) + \ldots \quad , \tag{76}$$

where n and ℓ are the well known principal and orbital quantum numbers of the non-relativistic H atom. Clearly, the second term on the righthand side is identiacal what we got in the nonrelativistic theory which reproduces the correct coarse 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 occur even more serious problems if we try to derive the continuity equation needed for the probabilistic interpretation. On the one hand we conjugate Eq. (69),

$$\left[\Box + \kappa^2\right]\psi\left(\vec{r},t\right)^* = 0\ , \tag{77}$$

and multiply it from the left with $\psi\left(\vec{r},t\right)$, on the other hand we multiply Eq. (69) from the right with $\psi\left(\vec{r},t\right)^*$. Subtracting the two equations from each other we get,

$$\psi\left(\vec{r},t\right)\Box\psi\left(\vec{r},t\right)^* - \psi\left(\vec{r},t\right)^*\Box\psi\left(\vec{r},t\right) = 0\,, \tag{78}$$

which can be written as

$$\begin{aligned}
&\frac{1}{c^2}\left(\psi\partial_t^2\psi^* - \psi^*\partial_t^2\psi\right) - \left(\psi\Delta\psi^* - \psi^*\Delta\psi\right)\\
&= \frac{1}{c^2}\partial_t\left(\psi\partial_t\psi^* - \psi^*\partial_t\psi\right) - \vec{\nabla}\left(\psi\vec{\nabla}\psi^* - \psi^*\vec{\nabla}\psi\right) = 0\ .
\end{aligned} \tag{79}$$

By multiplying the above equation with $\frac{\imath\hbar}{2m}$ and defining the probability density as,

$$\begin{aligned}
\rho\left(\vec{r},t\right) &= \frac{\imath\hbar}{2mc^2}\left(\psi^*\left(\vec{r},t\right)\partial_t\psi\left(\vec{r},t\right) - \psi\left(\vec{r},t\right)\partial_t\psi^*\left(\vec{r},t\right)\right)\\
&= \mathrm{Re}\left(\psi^*\left(\vec{r},t\right)\frac{\imath\hbar}{mc^2}\partial_t\psi\left(\vec{r},t\right)\right)
\end{aligned} \tag{80}$$

and the probability current density as,

$$\begin{aligned}
\vec{j}\left(\vec{r},t\right) &= \frac{\hbar}{2\imath m}\left(\psi^*\left(\vec{r},t\right)\vec{\nabla}\psi\left(\vec{r},t\right) - \psi\left(\vec{r},t\right)\vec{\nabla}\psi^*\left(\vec{r},t\right)\right)\\
&= \mathrm{Re}\left(\psi^*\left(\vec{r},t\right)\frac{\vec{p}}{m}\psi\left(\vec{r},t\right)\right)\ ,
\end{aligned} \tag{81}$$

we obtain the continuity equation,

$$\partial_t\rho\left(\vec{r},t\right) + \vec{\nabla}\vec{j}\left(\vec{r},t\right) = 0\ . \tag{82}$$

While the definition of the probability current density is the same what we obtained from the Schrödinger equation, the probability density is different. The problem is that $\rho(\vec{r}, t)$ is not positive definite. Based on the formula (80) the probability density of a free particle equals

$$\rho(\vec{r}, t) = \frac{E}{mc^2} |\psi(\vec{r}, t)|^2 \ , \tag{83}$$

which gives a negative probability for the negative energy particles, $E = -\sqrt{m^2c^4 + p^2c^2}$. The fundamental problem arises from the fact that the probability density (80) 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 π mezons.

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 in 1928, the wave equation of the free particles is written as a linear form of the four-momentum p_μ,

$$(\gamma^\mu p_\mu - M)\,\psi = 0\ , \tag{84}$$

where M is a scalar and the γ^μ operators commute with the p_μ operators. Consequently, the ψ 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_μ act and another Hilbert space where the γ^μ act.

We demand that the solutions of Eq. (84) 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 (84) with the operator $\gamma^\nu p_\nu + M$,

$$(\gamma^\nu p_\nu + M)\,(\gamma^\mu p_\mu - M)\,\psi = \left(\gamma^\nu\gamma^\mu p_\nu p_\mu - M^2\right)\psi = 0\ . \tag{85}$$

Making use that the operators p_μ commute with each other,

$$\gamma^\nu\gamma^\mu p_\mu p_\nu = \frac{1}{2}\left(\gamma^\nu\gamma^\mu p_\mu p_\nu + \gamma^\nu\gamma^\mu p_\nu p_\mu\right) \tag{86}$$

$$= \frac{1}{2}\left(\gamma^\nu\gamma^\mu + \gamma^\mu\gamma^\nu\right) p_\mu p_\nu\ . \tag{87}$$

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

$$M = mc$$

and by demanding

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

The latter condition implies the following anticommutator relations for γ^μ,

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}\ , \tag{88}$$

from which

$$\left(\gamma^0\right)^2 = I \quad \text{and} \quad \left(\gamma^i\right)^2 = -I \quad (i = 1,2,3)\ , \tag{89}$$

immediately follow.

If we substitute $p_\mu = \imath\hbar\partial_\mu$, we get the Dirac equation for a free particle,

$$(\imath\hbar\gamma^\mu\partial_\mu - mc)\,\psi = 0 \tag{90}$$

$$\Downarrow$$

$$(\imath\gamma^\mu\partial_\mu - \kappa)\,\psi = 0\ , \tag{91}$$

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 representations of the operators γ^μ, n must be even.

Proof:

(i) First we show that the trace of the operators γ^μ is zero. Namely,

$$\mathrm{Tr}\gamma^0 = -\mathrm{Tr}\gamma^0\left(\gamma^j\right)^2 = -\mathrm{Tr}\gamma^j\gamma^0\gamma^i = \mathrm{Tr}\left(\gamma^j\right)^2\gamma^0 = -\mathrm{Tr}\gamma^0 \ , \tag{92}$$

$$\mathrm{Tr}\gamma^j = \mathrm{Tr}\gamma^j\left(\gamma^0\right)^2 = \mathrm{Tr}\gamma^0\gamma^j\gamma^0 = -\mathrm{Tr}\left(\gamma^0\right)^2\gamma^j = -\mathrm{Tr}\gamma^j \ , \tag{93}$$

where we made use of the anticommutation relations (88) and the invariance of the trace under cyclic permutations of the operators.

(ii) Since $\left(\gamma^0\right)^2 = 1$, the eigenvalues of γ^0 are ± 1, while due to $\left(\gamma^j\right)^2 = -1$ the eigenvalues of γ^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,

$$\mathrm{Tr}\gamma^0 = m - (n-m) = 2m - n = 0 \quad , \tag{94}$$

$$\mathrm{Tr}\gamma^j = [m - (n-m)]\,i = (2m-n)\,i = 0 \quad , \tag{95}$$

from which it follows that n is an even integer.

It is easy to prove that the operators (or matrices) γ^μ ($\mu = 0, 1, 2, 4$) 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. Among them, only the four-dimensional representations can satisfy the relationships (88) in case of four γ^μ operators. Thus, as what follows the operators γ^μ will be represented by 4×4 matrices. In particular, we will use the so-called *standard (Dirac) representation*,

$$\gamma^0 = \begin{pmatrix} 1_2 & 0 \\ 0 & -1_2 \end{pmatrix} \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} , \tag{96}$$

where 1_2 is the 2×2 unit matrix and σ^i ($i = 1, 2, 3$) are the Pauli matrices. We can easily check that the required anticommutator relations apply:

$$\left\{\gamma^i, \gamma^k\right\} = -\begin{pmatrix} \left\{\sigma^i, \sigma^k\right\} & 0 \\ 0 & \left\{\sigma^i, \sigma^k\right\} \end{pmatrix} = -2\delta^{ik}\,1_4 = 2g^{ik}\,1_4 \tag{97}$$

and

$$\left\{\gamma^0, \gamma^i\right\} = \begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix} + \begin{pmatrix} 0 & -\sigma^i \\ -\sigma^i & 0 \end{pmatrix} = 0\,. \tag{98}$$

Later on we will see that in order to define a hermitian Hamilton operator, we demand

$$\left(\gamma^0\right)^\dagger = \gamma^0 \quad \left(\gamma^0\gamma^i\right)^\dagger = \gamma^0\gamma^i\,. \tag{99}$$

From these relationship we deduce

$$\gamma^0\gamma^i = \left(\gamma^i\right)^+\gamma^0 \Longrightarrow \left(\gamma^i\right)^+ = \gamma^0\gamma^i\gamma^0\,. \tag{100}$$

Since $(\gamma^0)^2 = I$, we can generally write

$$\left(\gamma^\mu\right)^+ = \gamma^0\gamma^\mu\gamma^0\,. \tag{101}$$

It is worth noting that the Dirac matrix $\gamma^5 = \gamma^0\gamma^1\gamma^2\gamma^3$,

$$\gamma^5 = \underbrace{\begin{pmatrix} 1_2 & 0 \\ 0 & -1_2 \end{pmatrix} \begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix}}_{\begin{pmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{pmatrix}} \underbrace{\begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{pmatrix}}_{\begin{pmatrix} -i\sigma_1 & 0 \\ 0 & -i\sigma_1 \end{pmatrix}} = -i \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix} . \tag{102}$$

anticommutes with the γ^μ matrices ($\mu = 0, 1, 2, 3$). According to Eq. (88), e.g.

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

while

$$\left(\gamma^5\right)^2 = -I_4 \, .$$

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

$$\psi\left(\vec{r}, t\right) = \begin{pmatrix} \psi_1\left(\vec{r}, t\right) \\ \psi_2\left(\vec{r}, t\right) \\ \psi_3\left(\vec{r}, t\right) \\ \psi_4\left(\vec{r}, t\right) \end{pmatrix} , \tag{103}$$

which we call *Dirac spinors*.

3.2 The Dirac equation in the presence of electromagnetic field

In the presence of electromagnetic field the canonical momentum p_μ is substituted by the kinetic momentum K_μ in the Dirac equation,

$$\left(\gamma^\mu K_\mu - mc\right)\psi = 0 \tag{104}$$

where

$$K_\mu = p_\mu - qA_\mu = i\hbar\partial_\mu - qA_\mu \tag{105}$$

$$\Downarrow$$

$$\left(\gamma^\mu\left(i\hbar\partial_\mu - qA_\mu\right) - mc\right)\psi = 0 \tag{106}$$

or

$$\left(\gamma^\mu\left(\partial_\mu + \frac{iq}{\hbar}A_\mu\right) + i\kappa\right)\psi = 0 \, . \tag{107}$$

3.2.1 Gauge transformation

Similarly to the non-relativistic theory, we consider the *gauge transformation*,

$$\vec{A}' = \vec{A} + \vec{\nabla}\Lambda, \; \phi' = \phi - \partial_t\Lambda \Rightarrow A'_\mu = \left(\frac{\phi - \partial_t\Lambda}{c}, -\vec{A} - \vec{\nabla}\Lambda\right) = A_\mu - \partial_\mu\Lambda \, . \tag{108}$$

The Dirac equation then transforms to

$$\left(\gamma^\mu\left(\partial_\mu + \frac{iq}{\hbar}A'_\mu\right) + i\kappa\right)\psi' = 0 \tag{109}$$

$$\Downarrow$$

$$\left[\gamma_\mu \left(\partial_\mu + \frac{iq}{\hbar}A_\mu - \frac{iq}{\hbar}\partial_\mu\Lambda\right) + i\kappa\right]\psi' = 0\,. \tag{110}$$

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

$$\psi' = \psi e^{\frac{iq}{\hbar}\Lambda}\,. \tag{111}$$

This relies on the obvious identity,

$$\partial_\mu(\psi e^{\frac{iq}{\hbar}\Lambda}) = e^{\frac{iq}{\hbar}\Lambda}\partial_\mu\psi + \frac{iq}{\hbar}\left(\partial_\mu\Lambda\right)\psi e^{\frac{iq}{\hbar}\Lambda}, \tag{112}$$

which should be substituted into (106),

$$\left[\gamma_\mu \left(\partial_\mu + \frac{iq}{\hbar}A_\mu\right) + i\kappa\right]\psi e^{\frac{iq}{\hbar}\Lambda} = 0\,. \tag{113}$$

After dividing by the factor $e^{\frac{iq}{\hbar}\Lambda}$ we obtain equation (107).

3.2.2 The Dirac Hamiltonian

Substituting the time and space like components of K_μ

$$K_\mu = \left(\frac{1}{c}\left(i\hbar\partial_t - q\phi\right), -\left(\vec{p} - q\vec{A}\right)\right)\,, \tag{114}$$

into Eq. (104) we get

$$\left[\frac{1}{c}\gamma^0\left(i\hbar\partial_t - q\phi\right) - \vec{\gamma}\left(\vec{p} - q\vec{A}\right) - mc\right]\psi = 0\,. \tag{115}$$

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

$$\left[\left(-i\hbar\partial_t + q\phi\right) + c\gamma^0\vec{\gamma}\left(\vec{p} - q\vec{A}\right) + \gamma^0 mc^2\right]\psi = 0\,, \tag{116}$$

we arrive at the following equation,

$$i\hbar\partial_t\psi = \left[c\gamma^0\vec{\gamma}\left(\vec{p} - q\vec{A}\right) + q\phi + \gamma^0 mc^2\right]\psi\,. \tag{117}$$

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

$$\alpha_i = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}\begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}\,, \tag{118}$$

while the matrix β is

$$\beta = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}\,. \tag{119}$$

With the previously defined matricies, the Dirac equation can be written as

$$i\hbar\partial_t\psi = \left[c\vec{\alpha}\left(\vec{p} - q\vec{A}\right) + q\phi + \beta mc^2\right]\psi\,, \tag{120}$$

or

$$i\hbar\partial_t\psi = H\psi\,, \tag{121}$$

where the relativistic Hamilton operator is defined as

$$H = c\vec{\alpha}\left(\vec{p} - q\vec{A}\right) + q\phi + \beta mc^2\,, \tag{122}$$

which is hermitian by construction: either we . 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*,

$$\left[c\vec{\alpha}\left(\vec{p} - q\vec{A}(\vec{r})\right) + q\phi(\vec{r}) + \beta mc^2\right]\psi(\vec{r}) = E\,\psi(\vec{r})\,. \tag{123}$$

3.2.3 The spectrum of free particles and the positron

In case of zero vector and scalar potentials, the solution of Eq. (123),

$$\begin{pmatrix} mc^2 I_2 & c\vec{\sigma}\vec{p} \\ c\vec{\sigma}\vec{p} & -mc^2 I_2 \end{pmatrix}\psi(\vec{r}) = E\psi(\vec{r})\,, \tag{124}$$

can be searched in the form of a planewave,

$$\psi(\vec{r}) = Ue^{i\vec{k}\vec{r}} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} e^{i\vec{k}\vec{r}}\,, \tag{125}$$

After substituting into the Dirac equation (124) we obtain the eigenvalue equation,

$$H(\vec{k})U = E\,U \tag{126}$$

of the 4×4 matrix of the Hamilton operator,

$$H(\vec{k}) = \begin{pmatrix} mc^2 I_2 & c\hbar\vec{\sigma}\vec{k} \\ c\hbar\vec{\sigma}\vec{k} & -mc^2 I_2 \end{pmatrix}\,. \tag{127}$$

Non-trivital ($U \neq 0$) solution only exists if the secular determinant, $\det(E - H(\vec{k}))$, is zero. Using the identity $(\vec{\sigma}\vec{k})^2 = k^2 I_2$, this gives

$$\left(E - mc^2\right)\left(E + mc^2\right) - c^2\hbar^2k^2 = 0 \tag{128}$$

$$\Downarrow$$

$$E^2 = m^2c^4 + c^2\hbar^2k^2\,, \tag{129}$$

admitting for the energy of the free particle,

$$E = \pm\sqrt{m^2c^4 + c^2\hbar^2k^2}\,. \tag{130}$$

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 operator (matrix) $\frac{1}{k}\vec{\sigma}\vec{k}$, since the Hamiltonian commutes with $\vec{\sigma}\vec{k}$,

$$\left[\begin{pmatrix} mc^2 I_2 & c\hbar\vec{\sigma}\vec{k} \\ c\hbar\vec{\sigma}\vec{k} & -mc^2 I_2 \end{pmatrix}, \begin{pmatrix} \vec{\sigma}\vec{k} & 0 \\ 0 & \vec{\sigma}\vec{k} \end{pmatrix}\right] = 0 \quad (i = x, y, z)\,. \tag{131}$$

Equation (130) demonstrates that 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 into an endless downward spiral 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*.

Following from the assumption of the Dirac sea, an electron with $E < -mc^2$ could be transferred 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. Such a particle, related to a hole in the Dirac sea, is called the *positron* (anti-particle of the electron), which as indicated above 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 Anderson in 1931. 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). Obviously, the reverse process can also occur: an electron-positron pair is annihilated by emitting a pair of photons.

3.2.4 Continuity equation

Let us start from the following form of the Dirac equation,

$$i\hbar\partial_t\psi = \left[c\vec{\alpha}\left(\vec{p} - q\vec{A}\right) + q\phi + \beta mc^2\right]\psi \ , \tag{132}$$

or for the components of the wave function,

$$i\hbar\partial_t\psi_r = c\vec{\alpha}_{rs}(\vec{p}\psi_s - q\vec{A}\psi_s) + q\phi\psi_r + mc^2\beta_{rs}\psi_s \ . \tag{133}$$

The complex conjugate of the above equation reads

$$-i\hbar\partial_t\psi_r^* = -c\vec{\alpha}_{rs}^*(\vec{p}\psi_s^* - q\vec{A}\psi_s^*) + q\phi\psi_r^* + mc^2\beta_{rs}^*\psi_s^* \tag{134}$$

$$= -c\left(\vec{p}\psi_s^*\right)\vec{\alpha}_{sr} - q\vec{A}\psi_s^*\vec{\alpha}_{sr} + q\phi\psi_r^* + mc^2\psi_s^*\beta_{sr} \ , \tag{135}$$

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

$$-i\hbar\partial_t\psi^+ = c\left(-\vec{p} - q\vec{A}\right)\psi^+\vec{\alpha} + q\phi\psi^+ + mc^2\psi^+\beta \ , \tag{136}$$

where

$$\psi^+ = (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*) \tag{137}$$

is the adjoint of the wave function. If we multiply Eq. (123) with ψ^+ from the left, and Eq. (136) with ψ from the right and substract the two equations from each other, we get

$$i\hbar\left[\psi^+\partial_t\psi + \left(\partial_t\psi^+\right)\psi\right] = c\left[\psi^+\vec{\alpha}\vec{p}\psi + \left(\vec{p}\psi^+\right)\vec{\alpha}\psi\right] \tag{138}$$

$$\Downarrow$$

$$\partial_t\left(\psi^+\psi\right) + \vec{\nabla}\left(c\psi^+\vec{\alpha}\psi\right) = 0\,. \tag{139}$$

Defining the finding probability density and the current density as,

$$\rho = \psi^+\psi \quad \text{and} \quad \vec{j} = c\psi^+\vec{\alpha}\psi\,, \tag{140}$$

we arrive at the continuity equation,

$$\partial_t\rho + \vec{\nabla}\vec{j} = 0\,. \tag{141}$$

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

$$j^\mu = (c\rho, \vec{j}) = c\psi^+\gamma^0\gamma^\mu\psi \tag{142}$$

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

$$\partial_\mu j^\mu = 0 \tag{143}$$

which applies in any intertial system. An important consequence of the continuity equation is the conservation of the finding probability of the particle normalized to unity,

$$\int \rho\left(\vec{r},t\right)d^3r = \int \psi^+\left(\vec{r},t\right)\psi\left(\vec{r},t\right)\,d^3r = 1\,. \tag{144}$$

3.2.5 Conjugate spinor and the conjugate Dirac equation

The continuity equation can be obtained from the covariant form of the Dirac equation, see Eq. (107)

$$\gamma^\mu\partial_\mu\psi + \frac{iq}{\hbar}\gamma^\mu A_\mu\psi + i\kappa\psi = 0\,. \tag{145}$$

Writing it out by components,

$$\gamma^\mu_{rs}\partial_\mu\psi_s + \frac{iq}{\hbar}\gamma^\mu_{rs}A_\mu\psi_s + i\kappa\psi_r = 0\,. \tag{146}$$

then by taking the complex conjugate,

$$\left(\gamma^\mu_{rs}\right)^*\partial_\mu\psi^*_s - \frac{iq}{\hbar}\left(\gamma^\mu_{rs}\right)^*A_\mu\psi^*_s - i\kappa\psi^*_r = 0 \tag{147}$$

$$\Downarrow \tag{148}$$

$$\left(\partial_\mu\psi^*_s\right)\left(\gamma^\mu\right)^+_{sr} - \frac{iq}{\hbar}A_\mu\psi^*_s\left(\gamma^\mu\right)^+_{sr} - i\kappa\psi^*_r = 0\,. \tag{149}$$

$$\Downarrow \tag{150}$$

$$\left(\partial_\mu\psi^+\right)\left(\gamma^\mu\right)^+ - \frac{iq}{\hbar}A_\mu\psi^+\left(\gamma^\mu\right)^+ - i\kappa\psi^+ = 0\,. \tag{151}$$

Using the identity for the adjoint of the γ^μ matrices,

$$\left(\partial_\mu \psi^\dagger\right) \gamma^0 \gamma^\mu \gamma^0 - \frac{iq}{\hbar} A_\mu \psi^+ \gamma^0 \gamma^\mu \gamma^0 - i\kappa \psi^+ = 0\,, \tag{152}$$

it is worth to introduce the *conjugate spinor* wavefunction.

$$\overline{\psi} = \psi^+ \gamma^0\,, \tag{153}$$

and we arrive at the *conjugate Dirac equation*,

$$\left(\partial_\mu \overline{\psi}\right) \gamma^\mu - \frac{iq}{\hbar} A_\mu \overline{\psi} \gamma^\mu - i\kappa \overline{\psi} = 0\,. \tag{154}$$

Multiplying Eq. (145) with $\overline{\psi}$ from the left and Eq. (110) with ψ from the right and adding the two equations leads to

$$\overline{\psi} \gamma^\mu \partial_\mu \psi + \left(\partial_\mu \overline{\psi}\right) \gamma^\mu \psi = 0\,, \tag{155}$$

or

$$\partial_\mu \left(\overline{\psi} \gamma^\mu \psi\right) = 0\,. \tag{156}$$

By defining the four-current density as

$$j^\mu = c\overline{\psi} \gamma^\mu \psi = c\psi^\dagger \gamma^0 \gamma^\mu \psi \tag{157}$$

we immediately obtain the covariant continuity equation

$$\partial_\mu j^\mu = 0\,. \tag{158}$$

3.3 Lorentz invariance of the Dirac equation

Let us write the Dirac equation in the form,

$$\left(\gamma^\mu \left(i\hbar \partial_\mu - qA_\mu(x)\right) - mc\right) \psi(x) = 0\,, \tag{159}$$

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

$$x'^\mu = \Lambda^\mu_{\cdot\,\nu}\, x^\nu, \tag{160}$$

$$\partial_\mu = \frac{\partial x'^\nu}{\partial x^\mu} \partial'_\nu = \partial'_\nu \Lambda^\nu_{\cdot\,\mu} = \left(\Lambda^T\right)_\mu^{\cdot\,\nu} \partial'_\nu \;\rightarrow\; \partial'_\mu = \overline{\Lambda}_\mu^{\cdot\,\nu} \partial_\nu \tag{161}$$

$$A'_\mu(x') = \overline{\Lambda}_\mu^{\cdot\,\nu} A_\nu(x)\;. \tag{162}$$

Under Lorentz transformation the Dirac equation should be written as

$$\left(\gamma^\mu \left(i\hbar \partial'_\mu - qA'_\mu(x')\right) - mc\right) \psi'(x') = 0 \tag{163}$$

$$\Downarrow$$

$$\left[\gamma^\mu \left(\overline{\Lambda}\right)_\mu^{\cdot\,\nu} \left(i\hbar \partial_\nu - qA_\nu(x)\right) - mc\right] \psi'(x') = 0\,. \tag{164}$$

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

$$\psi'(x') = \mathcal{S}(\Lambda)\, \psi(x)\;, \tag{165}$$

equation (164) transforms to

$$\gamma^\mu \overline{\Lambda}_{\mu}^{\cdot\nu} \mathcal{S}(\Lambda)(i\hbar\partial_\nu - qA_\nu(x))\psi(x) - mc\mathcal{S}(\Lambda)\psi(x) = 0 \ . \tag{166}$$

Applying $\mathcal{S}(\Lambda)$ on equation (159) we obtain

$$\mathcal{S}(\Lambda)\gamma^\mu(i\hbar\partial_\mu - qA_\mu(x))\psi(x) - mc\mathcal{S}(\Lambda)\psi(x) = 0 \ . \tag{167}$$

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

$$\left(\gamma^\mu \overline{\Lambda}_{\mu}^{\cdot\nu} \mathcal{S}(\Lambda) - \mathcal{S}(\Lambda)\gamma^\nu\right)(i\hbar\partial_\nu - qA_\nu(x))\psi(x) = 0 \,. \tag{168}$$

The above equality applies to arbitary wavefunctions, which implies

$$\gamma^\mu \overline{\Lambda}_{\mu}^{\cdot\nu} \mathcal{S}(\Lambda) - \mathcal{S}(\Lambda)\gamma^\nu = 0 \ , \tag{169}$$

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

$$\gamma^\mu = \Lambda^{\mu}_{\cdot\nu}\, \mathcal{S}(\Lambda)\gamma^\nu \mathcal{S}(\Lambda)^{-1} \ . \tag{170}$$

The above equation manifests the invariance of the γ^μ matrices under Lorentz tranformations, which lies at the heart of the Lorentz invariance of the Dirac equation. The above expression can be rewritten as,

$$\mathcal{S}(\Lambda)^{-1}\gamma^\mu\,\mathcal{S}(\Lambda) = \Lambda^{\mu}_{\cdot\nu}\gamma^\nu \ , \tag{171}$$

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

Statement: Given the infinitesimal Lorentz transformation ω, $\Lambda = e^\omega$, the solution of equation (171) is

$$\mathcal{S}(\Lambda) = e^{\frac{1}{4}\omega_{\mu\nu}\gamma^\mu\gamma^\nu} \ . \tag{172}$$

Proof: Let us introduce the infinitesimal spinor transformation τ as,

$$\mathcal{S}(\Lambda) = e^{-\tau} \ . \tag{173}$$

On the one hand, using the Hausdorff expansion yields

$$\begin{aligned}\mathcal{S}(\Lambda)^{-1}\gamma^\mu\,\mathcal{S}(\Lambda) &= e^\tau\gamma^\mu e^{-\tau} \\ &= \gamma^\mu + [\tau,\gamma^\mu] + \frac{1}{2!}[\tau,[\tau,\gamma^\mu]] + \frac{1}{3!}[\tau,[\tau,[\tau,\gamma^\mu]]] + \dots \ .\end{aligned} \tag{174}$$

On the other hand, the right-hand side of equation (171) can be expanded as

$$\Lambda^{\mu}_{\cdot\nu}\gamma^\nu = \left(e^\omega\right)^{\mu}_{\cdot\nu}\gamma_\nu = \gamma^\mu + \omega^{\mu}_{\cdot\nu}\gamma^\nu + \frac{1}{2!}\left(\omega^2\right)^{\mu}_{\cdot\nu}\gamma^\nu + \frac{1}{3!}\left(\omega^3\right)^{\mu}_{\cdot\nu}\gamma^\nu + \dots \ . \tag{175}$$

Obviously, the first terms of (174) and (166) are identical. Let's prove the identity of the second terms,

$$[\tau,\gamma^\mu] = \omega^{\mu}_{\cdot\nu}\gamma^\nu \ , \tag{176}$$

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

$$\begin{aligned}\tau\gamma^\mu &= -\frac{1}{4}\omega_{\lambda\delta}\gamma^\lambda\gamma^\delta\gamma^\mu = \frac{1}{4}\omega_{\lambda\delta}\gamma^\lambda\gamma^\mu\gamma^\delta - \frac{1}{2}\omega_{\lambda\delta}\gamma^\lambda g^{\mu\delta} \\ &= -\frac{1}{4}\omega_{\lambda\delta}\gamma^\mu\gamma^\lambda\gamma^\delta + \frac{1}{2}\omega_{\lambda\delta}\gamma^\delta g^{\lambda\mu} - \frac{1}{2}\omega_{\lambda\delta}\gamma^\lambda g^{\mu\delta} \\ &= \gamma^\mu\tau + \frac{1}{2}\omega_{\lambda\delta}\gamma^\delta g^{\lambda\mu} + \frac{1}{2}g^{\mu\delta}\omega_{\delta\lambda}\gamma^\lambda = \gamma^\mu\tau + \omega^{\mu}_{\cdot\lambda}\gamma^\lambda \ ,\end{aligned} \tag{177}$$

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

$$\underbrace{[\tau, [\tau, \dots [\tau, \gamma^\mu]]]}_{n} = (\omega^n)^{\mu}_{\cdot\nu} \gamma^\nu , \tag{178}$$

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

$$\begin{aligned} [\tau, \underbrace{[\tau, [\tau, \dots [\tau, \gamma^\mu]]]]}_{n-1} &= \left[\tau, \left(\omega^{n-1}\right)^{\mu}_{\cdot\nu} \gamma^\nu\right] = \left(\omega^{n-1}\right)^{\mu}_{\cdot\nu} [\tau, \gamma^\nu] \\ &= \left(\omega^{n-1}\right)^{\mu}_{\cdot\nu} \omega^{\nu}_{\cdot\lambda} \gamma^\lambda = (\omega^n)^{\mu}_{\cdot\lambda} \gamma^\lambda , \end{aligned} \tag{179}$$

Thus the proof is complete.

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

$$\mathcal{S}(\Lambda) = e^{-\frac{i}{\hbar}\omega_{\mu\nu}\sigma^{\mu\nu}} , \tag{180}$$

where

$$\sigma^{\mu\nu} = \frac{i\hbar}{4}\gamma^\mu\gamma^\nu = \frac{i\hbar}{8}[\gamma^\mu, \gamma^\nu] \quad (\mu \neq \nu) , \tag{181}$$

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

$$O(x) = \psi(x)^+ O\psi(x) , \tag{182}$$

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

$$\overline{O} = \gamma^0 O , \tag{183}$$

as

$$O(x) = \overline{\psi}(x)\overline{O}\,\psi(x) . \tag{184}$$

An example is the finding probability density and current density,

$$\varrho(x) = \overline{\psi}(x)\gamma^0\psi(x) , \; j^k(x) = c\overline{\psi}(x)\gamma^k\psi(x) \;\; (k = 1, 2, 3) , \tag{185}$$

or the four-current density,

$$j^\mu(x) = \left(c\varrho(x), \vec{j}(x)\right) = \overline{\psi}(x)\, (c\gamma^\mu)\, \psi(x) . \tag{186}$$

Statement: For an arbitrary Lorentz transformation Λ,

$$\mathcal{S}(\Lambda)^{-1} = \gamma^0\mathcal{S}(\Lambda)^+ \gamma^0 . \tag{187}$$

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

$$\gamma^0\mathcal{S}(\Lambda)^+ \gamma^0 = \gamma^0 \left(e^{-\tau}\right)^+ \gamma^0 = \sum_{n=0}^{\infty} \frac{1}{n!}\gamma^0 \left(-\tau^+\right)^n \gamma^0 . \tag{188}$$

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

$$\gamma^0 \left(-\tau^+\right) \gamma^0 = \frac{1}{4}\omega_{\mu\nu}\gamma^0 (\gamma^\nu)^+ (\gamma^\mu)^+ \gamma^0 = \frac{1}{4}\omega_{\mu\nu}\gamma^0 (\gamma^\nu)^+ \gamma^0\gamma^0 (\gamma^\mu)^+ \gamma^0 \tag{189}$$

$$= \frac{1}{4}\omega_{\mu\nu}\gamma^\nu\gamma^\mu = -\frac{1}{4}\omega_{\mu\nu}\gamma^\mu\gamma^\nu = \tau \tag{190}$$

$$\Downarrow$$

$$\gamma^0 \left(-\tau^+\right)^2 \gamma^0 = \left(\gamma^0 \left(-\tau^+\right) \gamma^0\right)^2 = \tau^2 \ \ldots \ \gamma^0 \left(-\tau^+\right)^n \gamma^0 = \tau^n \tag{191}$$

$$\Downarrow$$

$$\gamma^0 \mathcal{S}(\Lambda)^+ \gamma^0 = \sum_{n=0}^{\infty} \frac{\tau^n}{n!} = e^\tau = \mathcal{S}(\Lambda)^{-1} \ . \tag{192}$$

Since the spinor operator O is Lorentz invariant, the Lorentz transform of the density $O(x)$ can be calculated as,

$$\begin{aligned} O'(x') &= \psi'(x')^+ O \,\psi'(x') \\ &= [\mathcal{S}(\Lambda)\,\psi(x)]^+ O\,\mathcal{S}(\Lambda)\ \psi(x) \\ &= \psi(x)^+ \mathcal{S}(\Lambda)^+ O\,\mathcal{S}(\Lambda)\ \psi(x) \\ &= \overline{\psi}(x)\,\gamma^0\mathcal{S}(\Lambda)^+ \gamma^0\gamma^0 O\,\mathcal{S}(\Lambda)\ \psi(x) \\ &= \overline{\psi}(x)\ \left(\mathcal{S}(\Lambda)^{-1}\gamma^0 O\mathcal{S}(\Lambda)\right)\ \psi(x) \\ &= \overline{\psi}(x)\ \left(\mathcal{S}(\Lambda)^{-1}\overline{O}\mathcal{S}(\Lambda)\right)\ \psi(x)\ . \end{aligned} \tag{193}$$

Thus the Lorentz transform of the operator $\overline{O}$,

$$\overline{O}' = \mathcal{S}(\Lambda)^{-1}\overline{O}\mathcal{S}(\Lambda) \tag{194}$$

can be used to express the transformed density,

$$O'(x') = \overline{\psi}(x)\,\overline{O}'\psi(x)\ . \tag{195}$$

Based on equation (171) and (195),

$$\begin{aligned} j'^\mu(x') &= c\overline{\psi}(x)\ \left(\mathcal{S}(\Lambda)^{-1}\gamma^\mu\mathcal{S}(\Lambda)\right)\ \psi(x) \\ &= c\overline{\psi}(x)\ \left(\Lambda^\mu_{\cdot\nu}\gamma^\nu\right)\ \psi(x) = \Lambda^\mu_{\cdot\nu}\left[c\overline{\psi}(x)\ \gamma^\nu\,\psi(x)\right] \\ &= \Lambda^\mu_{\cdot\nu}\, j^\nu(x)\ , \end{aligned} \tag{196}$$

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

One can prove that $\overline{\psi}\psi$ and $\overline{\psi}\gamma^5\psi$ transform as a scalar, $\overline{\psi}\gamma^5\gamma^\mu\psi$ as a vector, and $\overline{\psi}\gamma^\mu\gamma^\nu\psi$ as a two-index tensor. (Note that γ^5 commutes with $\mathcal{S}(\Lambda)$!)

3.3.2 Rotations and the spin

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

$$\vec{r}' = \vec{r} + \varphi\, \vec{n} \times \vec{r}\,, \tag{197}$$

or for $i = 1, 2, 3$

$$x'^i = x^i + \varepsilon_{ikj} n_k x^j \varphi = \omega^{i}_{\bullet j}\, x^j \,.$$

The matrix of the infinitesimal Lorentz transformation is then given by

$$\omega^{i}_{\bullet j} = \varepsilon_{ikj} n_k \varphi \quad (i, j = 1, 2, 3)\ , \tag{198}$$

$$\omega^{\mu}_{\bullet 0} = \omega^{0}_{\bullet \mu} = 0 \quad (\mu = 0, 1, 2, 3)\ , \tag{199}$$

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

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

where

$$S^k = \frac{i\hbar}{4}\varepsilon_{kij}\gamma^i\gamma^j\ , \tag{200}$$

or

$$\vec{S} = \frac{i\hbar}{4}\vec{\gamma} \times \vec{\gamma}\ , \tag{201}$$

or by components,

$$S^1 = \frac{\hbar}{2} i\gamma^2\gamma^3\ ,\ S^2 = \frac{\hbar}{2} i\gamma^3\gamma^1\ ,\ S^3 = \frac{\hbar}{2} i\gamma^1\gamma^2\ . \tag{202}$$

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

$$\left[S^1, S^2\right] = \left(\frac{i\hbar}{2}\right)^2 \left[\gamma^2\gamma^3, \gamma^3\gamma^1\right] = \left(\frac{i\hbar}{2}\right)^2 \left(\gamma^2\gamma^3\gamma^3\gamma^1 - \gamma^3\gamma^1\gamma^2\gamma^3\right) \tag{203}$$

$$= \left(\frac{i\hbar}{2}\right)^2 \left(-\gamma^2\gamma^1 + \gamma^1\gamma^2\right) = i\hbar\frac{i\hbar}{2}\gamma^1\gamma^2 = i\hbar S^3\ , \tag{204}$$

and similarly,

$$\left[S^2, S^3\right] = i\hbar S^1\ ,\ \left[S^3, S^1\right] = i\hbar S^2\,, \tag{205}$$

that can be compactly written as

$$\left[S^i, S^j\right] = i\hbar\varepsilon_{ijk}S^k\ . \tag{206}$$

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

$$\left(S^1\right)^2 = -\frac{\hbar^2}{4}\gamma^2\gamma^3\gamma^2\gamma^3 = \frac{\hbar^2}{4}I \tag{207}$$

and, similarly,

$$\left(S^2\right)^2 = \left(S^3\right)^2 = \frac{\hbar^2}{4}I\ , \tag{208}$$

so

$$\vec{S}^2 = \frac{3\hbar^2}{4}I = \hbar^2 s\,(s+1)\,I \implies s = \frac{1}{2}\ . \tag{209}$$

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 as 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 γ^μ matrices,

$$\vec{S} = \frac{i\hbar}{4}\vec{\gamma} \times \vec{\gamma} = \frac{i\hbar}{4}\begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix} \times \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix} \tag{210}$$

$$= -\frac{i\hbar}{4}\begin{pmatrix} \vec{\sigma}\times\vec{\sigma} & 0 \\ 0 & \vec{\sigma}\times\vec{\sigma} \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} = \frac{\hbar}{2}\vec{\Sigma}\,, \tag{211}$$

where

$$\vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} . \tag{212}$$

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

$$\psi'(R(\varphi,\vec{n})\,\vec{r}, t) = \exp\left(-\frac{i}{\hbar}\vec{n}\vec{S}\varphi\right)\psi(\vec{r}, t) \tag{213}$$

or

$$\psi'(\vec{r}, t) = \exp\left(-\frac{i}{\hbar}\vec{n}\vec{S}\varphi\right)\,\psi(R(-\varphi,\vec{n})\,\vec{r}, t)\ . \tag{214}$$

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

$$\begin{aligned}\psi(R(-\varphi,\vec{n})\,\vec{r}, t) &\simeq \psi(\vec{r}, t) - \varphi\,(\vec{n}\times\vec{r})\,\vec{\nabla}\psi(\vec{r}, t) \\ &= \psi(\vec{r}, t) - \varphi\,\vec{n}\left(\vec{r}\times\vec{\nabla}\right)\psi(\vec{r}, t) \\ &= \psi(\vec{r}, t) - \frac{i}{\hbar}\varphi\,\vec{n}\,(\vec{r}\times\vec{p})\,\psi(\vec{r}, t) \\ &= \left[I - \frac{i}{\hbar}\varphi\,\vec{n}\,\vec{L}\right]\psi(\vec{r}, t)\ ,\end{aligned} \tag{215}$$

consequently, for finite rotations,

$$\psi(R(-\varphi,\vec{n})\,\vec{r}, t) = \exp\left(-\frac{i}{\hbar}\vec{n}\vec{L}\varphi\right)\psi(\vec{r}, t)\ . \tag{216}$$

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

$$\psi'(\vec{r}, t) = \exp\left(-\frac{i}{\hbar}\vec{n}\vec{J}\varphi\right)\psi(\vec{r}, t)\ , \tag{217}$$

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, from the Schrödinger form of the Dirac equation,

$$i\hbar\partial_t\psi = H\psi \tag{218}$$

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

$$\frac{dO}{dt} = \frac{i}{\hbar}\left[H, O\right] \ . \tag{219}$$

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

$$\vec{L} = \vec{r} \times \vec{p} \tag{220}$$

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

$$H = c\vec{\alpha}\vec{p} + \beta mc^2 + q\phi \ . \tag{221}$$

The required commutator can be written as

$$\left[H, L_i\right] = \left[c\vec{\alpha}\vec{p} + q\phi, \varepsilon_{ijk}x_j p_k\right] \ , \tag{222}$$

since L_i commutes with βmc^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,

$$\begin{aligned}\left[c\alpha_l p_l, \varepsilon_{ijk}x_j p_k\right] &= c\varepsilon_{ijk}\alpha_l\left[p_l, x_j p_k\right] = c\varepsilon_{ijk}\alpha_l(\underbrace{[p_l, x_j]}_{\frac{\hbar}{i}\delta_{lj}}p_k + x_j\underbrace{[p_l, p_k]}_{0}) \\ &= \frac{\hbar c}{i}\left(\vec{\alpha}\times\vec{p}\right)_i \end{aligned}\tag{223}$$

and

$$\begin{aligned}\left[q\phi, \varepsilon_{ijk}x_j p_k\right] &= q\varepsilon_{ijk}\left[\phi, x_j p_k\right] = q\varepsilon_{ijk}(\underbrace{[\phi, x_j]}_{0}p_k + x_j\underbrace{[\phi, p_k]}_{-\frac{\hbar}{i}\partial_k\phi}) \\ &= -\frac{\hbar q}{i}\left(\vec{r}\times\vec{\nabla}\phi\right)_i \quad , \end{aligned}\tag{224}$$

yields

$$\left[H, \vec{L}\right] = \frac{\hbar}{i}\left[c\left(\vec{\alpha}\times\vec{p}\right) - q\left(\vec{r}\times\vec{\nabla}\phi\right)\right] \ , \tag{225}$$

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*

$$\vec{J} = \vec{L} + \vec{S} \tag{226}$$

is conserved, where

$$\vec{S} = \frac{\hbar}{2}\vec{\Sigma} = \frac{\hbar}{2}\begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma}\end{pmatrix} \ , \tag{227}$$

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

$$
\begin{aligned}
[H, S_i] &= \frac{\hbar c}{2}\left[\alpha_j p_j, \Sigma_i\right] = \frac{\hbar c}{2} p_l \left[\alpha_j, \Sigma_i\right] = \\
&= \frac{\hbar c}{2} p_j \left[\begin{pmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{pmatrix}\begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} - \begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix}\begin{pmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{pmatrix}\right] \\
&= \frac{\hbar c}{2} p_j \begin{pmatrix} 0 & [\sigma_j, \sigma_i] \\ [\sigma_j, \sigma_i] & 0 \end{pmatrix} = i\hbar c \varepsilon_{jik}\, p_j \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix} \\
&= -\frac{\hbar c}{i}\left(\underline{\alpha} \times \vec{p}\right)_i \quad ,
\end{aligned} \tag{228}
$$

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

$$
\frac{d\vec{J}}{dt} = -q\left(\vec{r} \times \vec{\nabla}\phi\right) = \vec{r} \times \vec{F}\,, \tag{229}
$$

where $\vec{F} = -\vec{\nabla}\left(q\phi\right)$ 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

$$
x^{0\prime} = x^0 \quad x^{i\prime} = -x^i \quad (i = 1, 2, 3)\ , \tag{230}
$$

yielding the following simple matrix of transformation,

$$
\Lambda = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} = \overline{\Lambda}\,. \tag{231}
$$

The transformed Dirac equation,

$$
\left(\gamma^\mu K'_\mu\left(x'\right) - mc\right)\psi'\left(x'\right) = \left(\gamma^0 K_0\left(x\right) - \gamma^i K_i\left(x\right) - mc\right)\psi'\left(x'\right) = 0\ , \tag{232}
$$

can be manipulated as

$$
\left(\gamma^0 K_0 \gamma^0 - \gamma^0\gamma^i K_i - mc\gamma^0\right)\psi' = \left(\gamma^0 K_0 + \gamma^i K_i - mc\right)\gamma^0\psi' = 0\,, \tag{233}
$$

from which the transformation of the wavefunction can be read off,

$$
\psi' = \gamma^0\psi\,. \tag{234}
$$

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

$$
\left(\psi'\right)^+\psi' = \left(\gamma^0\psi\right)^+\gamma^0\psi = \psi^+\gamma^0\gamma^0\psi = \psi^+\psi\,, \tag{235}
$$

and

$$\overline{\psi}'\psi' = [(\gamma^0\psi)^+\gamma^0]\gamma^0\psi = \psi^+\gamma^0\psi = \overline{\psi}\psi\,. \tag{236}$$

Furthermore,

$$\overline{\psi}'\gamma^5\psi' = \psi^+\gamma^5\gamma^0\psi = -\psi^+\gamma^0\gamma^5\psi = -\overline{\psi}\gamma^5\psi\,, \tag{237}$$

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

$$\overline{\psi}'\gamma^i\psi' = \psi^+\gamma^i\gamma^0\psi = -\psi^+\gamma^0\gamma^i\psi = -\overline{\psi}\gamma^i\psi \quad \text{and} \quad \overline{\psi}'\gamma^0\psi' = \overline{\psi}\gamma^0\psi\,, \tag{238}$$

so $j^\mu = c\overline{\psi}\gamma^\mu\psi$ transforms as x^μ under reflection. We denote such vectors as *polar or normal vectors*. In addition,

$$\overline{\psi}'\gamma^i\gamma^5\psi' = \psi^+\gamma^i\gamma^5\gamma^0\psi' = \psi^+\gamma^0\gamma^i\gamma^5\psi = \overline{\psi}\gamma^i\gamma^5\psi \quad \text{and} \quad \overline{\psi}'\gamma^0\gamma^5\psi' = -\overline{\psi}\gamma^0\gamma^5\psi\,, \tag{239}$$

i.e. the components of the transform of $\overline{\psi}\gamma^\mu\gamma^5\psi$ change sign with respect to those of polar vectors, therefore, it is called a *pseudo (or axial) vector* .

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,

$$\left[c\vec{\alpha}\left(\vec{p} - q\vec{A}\right) + q\phi + \beta mc^2\right]\psi = E\psi\,, \tag{240}$$

into two parts (two components for each),

$$\psi = \begin{pmatrix} \chi \\ \varphi \end{pmatrix}\,, \tag{241}$$

where χ and φ are the so called large and small components, respectively. Let's write equation (240) in details,

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

$$\Downarrow$$

$$\left(E - mc^2 - q\phi\right)\chi - c\vec{\sigma}\vec{K}\varphi = 0\,, \tag{242}$$

$$\left(E + mc^2 - q\phi\right)\varphi - c\vec{\sigma}\vec{K}\chi = 0\,. \tag{243}$$

The small component φ can be expressed from equation (243),

$$\varphi = \left(E + mc^2 - q\phi\right)^{-1} c\vec{\sigma}\vec{K}\chi\,. \tag{244}$$

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

$$\varphi = \frac{1}{2mc}\vec{\sigma}\vec{K}\chi\,. \tag{245}$$

Substituting this expression into equation (242) and introducing the notation $E' = E - mc^2$, we arrive at

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

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

$$H_P\,\chi = E'\chi, \tag{246}$$

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

$$H_P = \frac{1}{2m}\left(\vec{\sigma}\vec{K}\right)\left(\vec{\sigma}\vec{K}\right) + q\phi = \frac{\vec{K}^2}{2m} + q\phi - \frac{\hbar q}{2m}\vec{B}\vec{\sigma}\;. \tag{247}$$

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 $\vec{K}\times\vec{K} = -\frac{\hbar q}{i}\vec{B}$.

3.4.1 Relativistic kinetic energy correction

Let's go beyond the approximation (245),

$$\begin{aligned}\varphi &= \frac{c\vec{\sigma}\vec{K}\,\chi}{E+mc^2-q\phi} = \frac{c\vec{\sigma}\vec{K}\,\chi}{2mc^2+E'-q\phi}\\ &= \frac{1}{2mc^2}\left(1+\frac{E'-q\phi}{2mc^2}\right)^{-1} c\vec{\sigma}\vec{K}\chi\\ &\simeq \frac{1}{2mc^2}\left(1-\frac{E'-q\phi}{2mc^2}\right) c\vec{\sigma}\vec{K}\chi\;,\end{aligned} \tag{248}$$

which we substitute again into (242),

$$(E'-q\phi)\,\chi = \frac{1}{2m}\vec{\sigma}\vec{K}\left(1-\frac{E'-q\phi}{2mc^2}\right)\vec{\sigma}\vec{K}\chi\;. \tag{249}$$

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

$$\begin{aligned}E'\chi &= H_P\chi - \frac{1}{4m^2c^2}\vec{\sigma}\vec{K}\;(E'-q\phi)\;\vec{\sigma}\vec{K}\;\chi\\ &= H_P\chi - \frac{1}{4m^2c^2}\left(\vec{\sigma}\vec{K}\right)\left(\vec{\sigma}\vec{K}\right)(E'-q\phi)\,\chi - \frac{1}{4m^2c^2}\vec{\sigma}\vec{K}\left[E'-q\phi,\vec{\sigma}\vec{K}\right]\chi\\ &= H_P\chi - \frac{1}{2mc^2}\left(H_P-q\phi\right)(E'-q\phi)\,\chi - \frac{1}{4m^2c^2}\vec{\sigma}\vec{K}\left[\vec{\sigma}\vec{K},q\phi\right]\;.\end{aligned} \tag{250}$$

The second term of on the right-hand side of (250) can be written in first order in $1/c^2$ as

$$H_M\,\chi \equiv -\frac{1}{2mc^2}\left(H_P-q\phi\right)^2\chi \simeq -\frac{1}{8m^3c^2}[(\vec{\sigma}\vec{p})(\vec{\sigma}\vec{p})]^2\chi \simeq -\frac{p^4}{8m^3c^2}\,\chi, \tag{251}$$

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

$$E'-q\phi = \sqrt{m^2c^4+p^2c^2} - mc^2 \simeq \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \ldots \tag{252}$$

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

$$H_0 = H_0 + W\;,$$

where

$$H_0 = \frac{p^2}{2m} - \frac{ke^2}{r} \quad , \quad E_n^{(0)} = -\frac{ke^2}{2a_0}\frac{1}{n^2} \quad , \tag{253}$$

and

$$\begin{aligned} W = H_M &= -\frac{p^4}{8m^3c^2} = -\frac{1}{2mc^2}\left(\frac{p^2}{2m}\right)^2 = -\frac{1}{2mc^2}\left(H_0 + \frac{ke^2}{r}\right)^2 \\ &= -\frac{1}{2mc^2}\left(H_0^2 + 2H_0\frac{ke^2}{r} + \frac{(ke^2)^2}{r^2}\right) . \end{aligned} \tag{254}$$

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

$$\begin{aligned} \delta E_{n\ell m}^{(1)} &= \langle n\ell m\,|W|\,n\ell m\rangle \\ &= -\frac{1}{2mc^2}\left(\left(E_n^{(0)}\right)^2 + 2E_n^{(0)}\underbrace{\left\langle\frac{ke^2}{r}\right\rangle_{n\ell m}}_{-2E_n^{(0)}} + k^2e^4\underbrace{\left\langle\frac{1}{r^2}\right\rangle_{n\ell m}}_{1/a_0^2n^3(\ell+1/2)}\right) \\ &= -\frac{1}{2mc^2}\left(-3(E_n^{(0)})^2 + \frac{k^2e^4}{a_0^2}\frac{1}{n^4}\frac{n}{(\ell+1/2)}\right) = \\ &= -\frac{2(E_n^{(0)})^2}{mc^2}\left(\frac{n}{\ell+1/2} - \frac{3}{4}\right) . \end{aligned} \tag{255}$$

This energy correction splits the energy levels of the hydrogen atom according to the angular quantum number ℓ. 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 (250),

$$-\frac{1}{4m^2c^2}\left(\vec{\sigma}\vec{p}\right)\left(\vec{\sigma}\left[\vec{p},q\phi\right]\right) = -\frac{1}{4m^2c^2}\vec{p}\left[\vec{p},q\phi\right] - \frac{i}{4m^2c^2}\vec{\sigma}\left(\vec{p}\times\left[\vec{p},q\phi\right]\right) . \tag{256}$$

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 (256) can further be arranged as

$$\begin{aligned} \left(\vec{p}\times\left[\vec{p},q\phi\right]\right)_i = \varepsilon_{ijk}p_j\left[p_k,q\phi\right] &= [\underbrace{\varepsilon_{ijk}p_jp_k}_{=0},q\phi] - \varepsilon_{ijk}\left[p_j,q\phi\right]p_k \\ &= -(\left[\vec{p},q\phi\right]\times\vec{p})_i \end{aligned}$$

$$\Downarrow$$

$$\frac{i}{4m^2c^2}\vec{\sigma}\left(\left[\vec{p},q\phi\right]\times\vec{p}\right) = \frac{\hbar}{4m^2c^2}\vec{\sigma}\left(\left[\vec{\nabla}\left(q\phi\right)\right]\times\vec{p}\right) . \tag{257}$$

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

$$H_{\rm SO} = -\frac{\hbar}{4m^2c^2}\vec{\sigma}\left(q\vec{\mathcal{E}}\times\vec{p}\right) , \tag{258}$$

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

$$\vec{\nabla}\phi\,(r) = \frac{d\phi\,(r)}{dr}\frac{1}{r}\vec{r}, \tag{259}$$

the Hamiltonian (257) transforms to

$$H_{\rm SO} = \frac{\hbar q}{4m^2c^2}\frac{1}{r}\frac{d\phi\,(r)}{dr}\,\vec{\sigma}\,(\vec{r}\times\vec{p}) = \frac{1}{2m^2c^2}\frac{1}{r}\frac{d\,(q\phi\,(r))}{dr}\,\vec{L}\vec{S}\,. \tag{260}$$

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

$$H_{\rm SO} = \frac{1}{c^2q^2}\frac{1}{r}\frac{d\,(q\phi\,(r))}{dr}\,\vec{M}_L\vec{M}_S\,. \tag{261}$$

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,

$$\vec{B} = \frac{1}{c^2}(\vec{\mathcal{E}}\times\vec{v}) = \frac{1}{mc^2}(\vec{\mathcal{E}}\times\vec{p})\,, \tag{262}$$

where $\vec{v}$ and $\vec{p}$ are the velocity and momentum of the elecron 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

$$H_L = -\vec{M}_S\,\vec{B} = -\frac{1}{m^2c^2}\vec{S}(q\vec{\mathcal{E}}\times\vec{p}) = -\frac{\hbar}{2m^2c^2}\vec{\sigma}(q\vec{\mathcal{E}}\times\vec{p})\,, \tag{263}$$

which is twice as large in magnitude as the spin-orbit interaction in (258). 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,

$$\vec{\omega}_T = \frac{1}{2c^2}(\vec{a}\times\vec{v}) = \frac{1}{2m^2c^2}(q\vec{\mathcal{E}}\times\vec{p})\,, \tag{264}$$

and a corresponding interaction energy,

$$H_T = \vec{\omega}_T\,\vec{S} = \frac{1}{2m^2c^2}\vec{S}(q\vec{\mathcal{E}}\times\vec{p}) = \frac{\hbar}{4m^2c^2}\vec{\sigma}(q\vec{\mathcal{E}}\times\vec{p})\,. \tag{265}$$

The sum of H_L and H_T is then clearly identical with $H_{\rm SO}$ in (258).

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

$$W = \frac{ke^2}{2m^2c^2}\frac{1}{r^3}\,\vec{L}\vec{S} \tag{266}$$

as

$$W = \frac{ke^2}{4m^2c^2}\frac{1}{r^3}\,\left(J^2 - L^2 - S^2\right)\,, \tag{267}$$

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

$$|n,\ell,j,m_j\rangle = \sum_{m_s=\pm\frac{1}{2}} C(\ell\, m_\ell, \frac{1}{2} m_s; j, m_j)\ |n\,\ell\,m_\ell\rangle\,|\frac{1}{2}, m_s\rangle\ , \tag{268}$$

where $C(\ell\, m_\ell, \frac{1}{2} m_s; j, m_j)$ are 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 subspace corresponding to the eigenvalue $E_n^{(0)}$. The first order energy corrections can then be evaluated as,

$$\delta E^{(1)}_{n,\ell,j,m_j} = \frac{\hbar^2 k e^2}{4m^2c^2}\left\langle\frac{1}{r^3}\right\rangle_{n\ell}\left(j\,(j+1) - \ell\,(\ell+1) - \frac{3}{4}\right)\ . \tag{269}$$

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

$$\left\langle\frac{1}{r^3}\right\rangle_{n\ell} = \frac{1}{n^3 a_0^3}\frac{1}{\left(\ell+\frac{1}{2}\right)\ell\,(\ell+1)}\ , \tag{270}$$

we get,

$$\delta E^{(1)}_{n,\ell,j,m_j} = \frac{\hbar^2 k e^2}{4m^2c^2a_0^3}\frac{1}{n^4}\,n\,\frac{j\,(j+1) - \ell\,(\ell+1) - \frac{3}{4}}{\left(\ell+\frac{1}{2}\right)\ell\,(\ell+1)} \tag{271}$$

$$= \frac{2(E_n^{(0)})^2}{mc^2}\,n\,\frac{j\,(j+1) - \ell\,(\ell+1) - \frac{3}{4}}{(2\ell+1)\,\ell\,(\ell+1)}\ , \tag{272}$$

where we used

$$a_0 = \frac{\hbar^2}{mke^2} \to \frac{\hbar^2 k e^2}{4m^2c^2a_0^3}\frac{1}{n^4} = \frac{1}{mc^2}\frac{(ke^2)^2}{4a_0^2n^4} = \frac{(E_n^{(0)})^2}{mc^2}\ . \tag{273}$$

After some algebra,

$$\frac{j\,(j+1) - \ell\,(\ell+1) - \frac{3}{4}}{(2\ell+1)\,\ell\,(\ell+1)} \underset{j=\ell+\frac{1}{2}}{=} \frac{\left(\ell+\frac{1}{2}\right)\left(\ell+\frac{3}{2}\right) - \ell\,(\ell+1) - \frac{3}{4}}{(2\ell+1)\,\ell\,(\ell+1)} = \frac{1}{(2\ell+1)\,(\ell+1)}$$
$$= \frac{1}{\ell+\frac{1}{2}} - \frac{1}{\ell+1} = \frac{1}{\ell+\frac{1}{2}} - \frac{1}{j+\frac{1}{2}}\ , \tag{274}$$

$$\frac{j\,(j+1) - \ell\,(\ell+1) - \frac{3}{4}}{(2\ell+1)\,\ell\,(\ell+1)} \underset{j=\ell-\frac{1}{2}}{=} \frac{\left(\ell-\frac{1}{2}\right)\left(\ell+\frac{1}{2}\right) - \ell\,(\ell+1) - \frac{3}{4}}{(2\ell+1)\,\ell\,(\ell+1)} = -\frac{1}{(2\ell+1)\,\ell}$$
$$= \frac{1}{\ell+\frac{1}{2}} - \frac{1}{\ell} = \frac{1}{\ell+\frac{1}{2}} - \frac{1}{j+\frac{1}{2}}\ , \tag{275}$$

we end up with

$$\delta E^{(1)}_{n,\ell,j,m_j} = \frac{2(E_n^{(0)})^2}{mc^2}\left(\frac{n}{\ell+\frac{1}{2}} - \frac{n}{j+\frac{1}{2}}\right)\ . \tag{276}$$

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

$$\delta E^{(1)}_{n,\ell,j,m_j} = -\frac{2(E_n^{(0)})^2}{mc^2}\left(\frac{n}{\ell+1/2} - \frac{3}{4}\right) + \frac{2(E_n^{(0)})^2}{mc^2}\left(\frac{n}{\ell+\frac{1}{2}} - \frac{n}{j+\frac{1}{2}}\right)$$
$$= -\frac{2(E_n^{(0)})^2}{mc^2}\left(\frac{n}{j+1/2} - \frac{3}{4}\right)\ , \tag{277}$$

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 $\left\langle \frac{1}{r^3} \right\rangle_{n\ell} \sim \frac{1}{\ell}$ is infinite. Assuming an atomic nucleus of finite size, $\left\langle \frac{1}{r^3} \right\rangle_{n\ell}$ becomes finite, but the matrixelement of the $\vec{L}\vec{S}$ operator for $\ell = 0$ is zero, hence the spin-orbit coupling doesn't ifluence the energy of the s-orbitals in first-order perturbation theory. This means that taking into account the first-order perturbation correction due the relativistic kinetic energy term and the spin-orbit interaction, for $\ell = 0$ we get

$$\delta E^{(1)}_{n,0,j=\frac{1}{2},m_j} = -\frac{2(E^{(0)}_n)^2}{mc^2} \left(2n - \frac{3}{4} \right) , \tag{278}$$

whereas, from (277) we deduce the correct result,

$$\delta E^{(1)}_{n,0,j=\frac{1}{2},m_j} = -\frac{2(E^{(0)}_n)^2}{mc^2} \left(n - \frac{3}{4} \right) . \tag{279}$$

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 ψ,

$$\int d^3r \; \psi^+\psi = \int d^3r \; \left(\chi^+\chi + \varphi^+\varphi \right) . \tag{280}$$

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

$$\begin{aligned} \int d^3r \; \varphi^+\varphi &\simeq \frac{1}{4m^2c^2} \int d^3r \; \chi^+ \left(\vec{\sigma}\vec{p}\right) \left(\vec{\sigma}\vec{p}\right) \chi \\ &\simeq \frac{1}{4m^2c^2} \int d^3r \; \chi^+ p^2 \chi \; , \end{aligned} \tag{281}$$

thus the norm of the wavefunction ψ can be approximated as

$$\int d^3r \; \psi^+\psi \simeq \int d^3r \; \chi^+ \left(1 + \frac{p^2}{4m^2c^2} \right) \chi \, . \tag{282}$$

Let us introduce the large component χ_S the norm of which equals that of ψ within the above approximation,

$$\int d^3r \; \chi_S^+ \chi_S = \int d^3r \; \chi^+ \left(1 + \frac{p^2}{4m^2c^2} \right) \chi \, . \tag{283}$$

This leads to the relationship between χ_S and χ,

$$\chi_S = \left(1 + \frac{p^2}{8m^2c^2} \right) \chi \quad \Longrightarrow \quad \chi = \left(1 - \frac{p^2}{8m^2c^2} \right) \chi_S \; , \tag{284}$$

Thus working with the normalized large component χ_S, the norm of the eliminated small component φ 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 ψ and $\begin{pmatrix} \chi_S \\ 0 \end{pmatrix}$.

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

$$H' = H_{\rm P} + H_{\rm M} + H_{\rm SO} - \frac{1}{4m^2c^2}\vec{p}\,[\vec{p}, q\phi]\,, \tag{285}$$

where we included the first term on the right hand side of equation (256). Employing (284) to the eigenvalue equation,

$$E'\chi = H'\chi\;, \tag{286}$$

we obtain

$$E'\chi_S = \left(1 + \frac{p^2}{8m^2c^2}\right) H' \left(1 - \frac{p^2}{8m^2c^2}\right) \chi_S\,. \tag{287}$$

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

$$E'\chi_S \simeq \left(H' + \left[\frac{p^2}{8m^2c^2}, H_P\right]\right) \chi_S \simeq \left(H' + \frac{1}{8m^2c^2}\left[p^2, q\phi\right]\right) \chi_S\;, \tag{288}$$

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 (285), we get the third $1/c^2$ correction to the non-relativistic Hamilton operator, called the *Darwin term*,

$$\begin{aligned} H_D &= -\frac{1}{4m^2c^2}\vec{p}\,[\vec{p}, q\phi] + \frac{1}{8m^2c^2}\left[p^2, q\phi\right] \\ &= \frac{1}{8m^2c^2}\left([\vec{p}, q\phi]\,\vec{p} - \vec{p}\,[\vec{p}, q\phi]\right) \\ &= -\frac{1}{8m^2c^2}\left[\vec{p}, [\vec{p}, q\phi]\right] = \frac{\hbar^2}{8m^2c^2}\left[\vec{\nabla}, \left[\vec{\nabla}, q\phi\right]\right] \\ &= \frac{\hbar^2}{8m^2c^2}\Delta\left(q\phi\right)\;. \end{aligned} \tag{289}$$

The Darwin term has no classical interpretation. The origin of this term can be traced back to the fact that the electron cannot be localized to a point-like particle but it rapidly oscillates in a length scale of the Compton wavelength, $\lambda_C = \hbar/mc$ (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 is space and time arise due to the interference of these components. 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 λ_C,

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

where we assumed $\langle\delta x_i\delta x_j\rangle \simeq \delta_{ij}\,\langle\delta r^2\rangle/3 \simeq \delta_{ij}\,\lambda_C^2/3$. This simple estimation gives a very good qualitative agreement with the Darwin term derived in (289).

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

$$H_D\left(\vec{r}\right) = \frac{ke^2\hbar^2\pi}{2m^2c^2}\,\delta\left(\vec{r}\right)\;, \tag{290}$$

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

$$\delta E^{(1)}_{n\ell m} = \int d^3r \, \psi_{n\ell m}(\vec{r})^* \, H_D(\vec{r}) \, \psi_{n\ell m}(\vec{r}) = \frac{ke^2\hbar^2\pi}{2m^2c^2} |\psi_{n\ell m}(0)|^2 \, . \tag{291}$$

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

$$\delta E^{(1)}_{n00} = \frac{ke^2\hbar^2}{2m^2c^2} \frac{1}{n^3a_0^3} = \frac{2}{mc^2} \left(\frac{ke^2}{2a_0n^2} \right)^2 \underbrace{\frac{\hbar^2}{mke^2a_0}}_{=1} n = \frac{2(E^{(0)}_n)^2}{mc^2} n \, . \tag{292}$$

Adding this correction to (278), we indeed get the correct result (279)!

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

$$(H_\mathrm{P} + H_\mathrm{M} + H_\mathrm{SO} + H_\mathrm{D}) \, \chi_S = \left(E - mc^2\right) \chi_S \, , \tag{293}$$

where χ_S is the normalized two-component wavefunction,

$$H_\mathrm{P} = \frac{1}{2m} \vec{K}^2 - \frac{\hbar q}{2m} \vec{B}\vec{\sigma} + q\phi \, , \tag{294}$$

is the Pauli-Schrödinger Hamiltonian,

$$H_\mathrm{M} = -\frac{1}{8m^3c^2} p^4 \tag{295}$$

is the relativistic kinetic energy correction,

$$H_\mathrm{SO} = -\frac{\hbar q}{4m^2c^2} \vec{\sigma} \left(\vec{\mathcal{E}} \times \vec{p} \right) \tag{296}$$

is the spin-orbit interaction and

$$H_\mathrm{D} = \frac{\hbar^2}{8m^2c^2} \Delta (q\phi) \tag{297}$$

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,

$$\varphi \simeq \frac{\vec{\sigma}\vec{K}\chi}{2mc} \, , \tag{298}$$

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

$$\begin{aligned} \vec{j} = c\psi^+ \vec{\alpha} \psi &= c \left(\chi^+, \varphi^+ \right) \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \begin{pmatrix} \chi \\ \varphi \end{pmatrix} = c \left(\chi^+ \vec{\sigma} \varphi + \varphi^+ \vec{\sigma} \chi \right) \\ &= \frac{1}{2m} \left[\chi^+ \vec{\sigma} (\vec{\sigma}\vec{K}\chi) + ((\vec{K}\chi)^+ \vec{\sigma}) \vec{\sigma} \chi \right] \, . \end{aligned} \tag{299}$$

By using the identities,

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

we can proceed as

$$\vec{j} = \frac{1}{2m}\left(\chi^+(\vec{K}\chi) + (\vec{K}\chi)^+\chi\right) + \frac{i}{2m}\left(\chi^+(\vec{K}\times\vec{\sigma}\chi) - (\vec{K}\chi)^+\times\vec{\sigma}\chi\right) . \tag{300}$$

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

$$\vec{j}_{\rm nr} \equiv \frac{1}{m}\mathrm{Re}\left(\chi^+(\vec{K}\chi)\right) = \frac{\hbar}{2im}\left[\chi^+(\vec{\nabla}\chi) - (\vec{\nabla}\chi^+)\chi\right] - \frac{q}{m}\vec{A}\,\chi^+\chi . \tag{301}$$

The second term can be evaluated by using

$$\chi^+(\vec{K}\times\vec{\sigma}\chi) \;= \frac{\hbar}{i}\chi^+(\vec{\nabla}\times\vec{\sigma}\chi) - \chi^+(q\vec{A}\times\vec{\sigma}\chi)$$

$$(\vec{K}\chi)^+\times\vec{\sigma}\chi = -\frac{\hbar}{i}(\vec{\nabla}\chi^+)\times\vec{\sigma}\chi - \chi^+(q\vec{A}\times\vec{\sigma}\chi)\,,$$

and it yields a new contribution:

$$\begin{aligned}\vec{j}_{\rm spin} &\equiv \frac{i}{2m}\left(\chi^+(\vec{K}\times\vec{\sigma}\chi) - (\vec{K}\chi)^+\times\vec{\sigma}\chi\right)\\ &= \frac{\hbar}{2m}\left(\chi^+(\vec{\nabla}\times\vec{\sigma}\chi) + (\vec{\nabla}\chi^+)\times\vec{\sigma}\chi\right)\\ &= \frac{\hbar}{2m}\vec{\nabla}\times\left(\chi^+\vec{\sigma}\chi\right) = \frac{1}{m}\,\mathrm{curl}\,(\chi^+\vec{S}\chi)\end{aligned} \tag{302}$$

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

$$\vec{M}_S\,(\vec{r},t) = \chi^+\,(\vec{r},t)\;\vec{M}_S\,\chi\,(\vec{r},t)\;, \tag{303}$$

the corresponding charge current reads as

$$q\,\vec{j}_{\rm spin}\,(\vec{r},t) = \mathrm{curl}\;\vec{M}_S\,(\vec{r},t)\;. \tag{304}$$