7 Dirac equation

The Klein-Gordon equation is a quantum-mechanical wave equation for a spinless scalar particle. For a realistic description of the electron, we need a relativistic quantum mechanical wave equation that takes the spin degree of freedom into account.

Dirac's discovery was rather motivated by the conceptual clash between the appearance of 2nd order time derivatives and general properties of quantum mechanics.

7.1 Free Dirac equation

Let us "derive" the free Dirac equation by demanding that it should be a "wave" equation of first order in time, i.e. linear in energy according to the correspondence rule. In order to preserve relativistic covariance it also has to be linear in space derivatives, i.e. spatial momenta. This justifies the following ansatz:

$$\left[(\gamma, p) - mc\right] \psi = 0 \qquad (7.1)$$

where

$$(\gamma, p) = \gamma^\mu p_\mu =: \not{p} \quad (\text{"p slash"}) . \qquad (7.2)$$

As ψ should describe an electron with spin, we already expect that it is not just a complex number, but rather carries several components (c.f. two-component notation for spin in non-relativistic QM). Hence, we may already expect that γ^μ is not merely a 4-vector, but a set of matrices summarized in 4-vector notation.

Furthermore, we demand that a solution ψ still satisfies the relativistic energy-momentum relation:

$$\left((p, p) - m^2 c^2\right) \psi = 0 \qquad (7.3)$$

Multiplying (7.1) with $\left[(\gamma, p) + mc\right]$ from the right, yields

$$(\not{p} + mc)(\not{p} - mc)\psi$$

$$= (\not{p}\not{p} - m^2 c^2)\psi \underset{!}{\overset{!}{=}} 0 \qquad (7.4)$$

For (7.4) to be equivalent to (7.3), we conclude that

$$\not{p}\not{p} = \gamma^\mu \gamma^\nu p_\mu p_\nu = \frac{1}{2}(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu) p_\mu p_\nu$$

$$\stackrel{!}{=} g^{\mu\nu} p_\mu p_\nu \cdot \mathbb{1} \qquad (7.5)$$

has to hold. Introducing the anti-commutator $\{A, B\} = AB + BA$, the γ-matrices have to satisfy

$$\{\gamma^\mu, \gamma^\nu\} = 2 g^{\mu\nu} \mathbb{1} \qquad (7.6)$$

"<u>Dirac algebra</u>"

To summarize: a solution to (7.1) also solves the Klein-Gordon equation, if γ^μ are matrices that satisfy the Dirac algebra (7.6).

For a representation of the Dirac algebra, we need $D=4$ anti-commuting matrices γ^μ.

From the description of spin in non-relativistic

QM, we are already familiar with anticommuting objects: the Pauli matrices σ_i, satisfying

$$\{\sigma_i, \sigma_j\} = 2\,\delta_{ij} \qquad , i = 1,2,3 \tag{7.7}$$

As the Pauli matrices are 2×2 matrices, we cannot construct a forth 2×2 matrix, satisfying (7.7) with all other Pauli matrices. It turns out that the least dimensional representation of (7.6) is in terms of 4×4 matrices.

This representation (once we have one) is not unique, as with a given set of γ^μ, also

$$\tilde{\gamma}^\mu = U \gamma^\mu U^{-1} \tag{7.8}$$

with any invertible U also satisfies (7.6). Representations of (7.6) which are related by (7.8) are said to be equivalent.

Of course, also larger representations exist. For a given set of γ^μ, also the 8×8 matrices

$$\hat{\gamma}^\mu = \begin{pmatrix} \gamma^\mu & 0 \\ 0 & \gamma^\mu \end{pmatrix} \tag{7.9}$$

satisfy (7.6). Obviously, the 4×4 sub blocks $\sim \gamma^\mu$ do not mix but are mapped onto themselves. Such representations are said to be <u>reducible</u>. Those which do not have invariant subspaces are called irreducible.

From a given irreducible representation, arbitrarily many reducible representations can be constructed (analogously to (7.9)).

In particle physics, a common (though not historically the most favorite) representation is the so-called <u>chiral representation</u>:

$$\gamma^0 = \sigma_1 \otimes \sigma_0 = \begin{pmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{pmatrix}, \qquad (7.10)$$

$$\gamma^k = -i\sigma_2 \otimes \sigma_k = \begin{pmatrix} 0 & -\sigma_k \\ \sigma_k & 0 \end{pmatrix},$$

where as usual $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$,

and we define and use

$$\sigma_0 = \mathbb{1}_{2\times 2} \,. \qquad (7.11)$$

We may combine σ_0 and σ_i into a 4 vector-type object:

$$\sigma_\mu = \tilde{\sigma}^\mu = \{\sigma_0, \sigma_i\}$$
$$\sigma^\mu = \tilde{\sigma}_\mu = \{\sigma_0, -\sigma_i\} \,. \qquad (7.12)$$

(We stress that σ^μ nor $\tilde{\sigma}^\mu$ do not satisfy the Dirac algebra (7.6)).

Here, we have introduced both σ_μ and $\tilde{\sigma}_\mu$ in order to notationally account for the relative minus sign of the ~~off~~ diagonal components in the definition (7.10):

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \tilde{\sigma}^\mu & 0 \end{pmatrix} . \qquad (7.13)$$

Obviously, γ^0 is hermitean and γ_i are anti-hermitean.

Let us consider a few useful identities.

Using $$\sigma_i \sigma_j = \delta_{ij} \sigma_0 + i \varepsilon_{ijk} \sigma_k , \qquad (7.14)$$

it is straightforward to show that $(a, b \mathrel{\hat{=}} a^\mu, b^\mu)$

$$(\sigma, a) \cdot (\sigma, b) = \sigma_0 \left(a^0 b^0 + \vec{a} \cdot \vec{b} \right) + \vec{\sigma} \cdot \left(a^0 \vec{b} + \vec{a} b^0 + i \, \vec{a} \times \vec{b} \right) \qquad (7.15)$$

and

$$(\sigma, a) \cdot (\tilde{\sigma}, a) = \sigma_0 \left(a^0 a^0 - \vec{a} \cdot \vec{a} \right) = \sigma_0 \, (a, a) . \qquad (7.16)$$

Such objects indeed occur frequently; for instance in the Dirac equation, we encounter

$$\not{p} = \gamma^\mu p_\mu = \begin{pmatrix} 0 & \sigma^\mu p_\mu \\ \tilde{\sigma}^\mu p_\mu & 0 \end{pmatrix} = \begin{pmatrix} 0 & (\sigma, p) \\ (\hat{\sigma}, p) & 0 \end{pmatrix}. \tag{7.17}$$

The object (σ, a) reads explicitly

$$(\sigma, a) = \sigma^\mu a_\mu = \begin{pmatrix} a^0 + a^3 & a^1 - ia^2 \\ a^1 + ia^2 & a^0 - a^3 \end{pmatrix} \tag{7.18}$$

and obviously parameterizes all possible hermitean 2×2 matrices. The determinant of (7.18) reduces to a well-known expression:

$$\det(\sigma, a) = \det(\tilde{\sigma}, a) = a^{0\,2} - a^{1\,2} - a^{2\,2} - a^{3\,2} = a^\mu a_\mu = (a, a), \tag{7.19}$$

which is a Lorentz invariant quantity.

In the case that $\det(\sigma, a) = 1$, (7.16) implies that

$$(\sigma, a) = (\tilde{\sigma}, a)^{-1}. \tag{7.20}$$

Let us return to the Dirac equation. As the γ^μ are 4×4 matrices, also the wave function ψ is not just a complex number, but a

4 dimensional "vector" (which however does not transform under the ordinary 4D representation of the Lorentz group).
As this 4D Dirac structure is related to spin degrees of freedom, we call ψ a Spinor.

Having decomposed the Dirac matrices into 2x2 blocks, we can also decompose the 4-component Dirac spinor into 2 2-component spinors:

$$\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix} \tag{7.21}$$

These components in the chiral representation are often called left-handed (ϕ) and right-handed (χ) Weyl spinors. In this decomposition, the Dirac equation reads

$$\gamma^\mu p_\mu \psi = \begin{pmatrix} 0 & (\sigma, p) \\ (\tilde{\sigma}, p) & 0 \end{pmatrix} \begin{pmatrix} \phi \\ \chi \end{pmatrix} = \begin{pmatrix} (\sigma, p)\chi \\ (\tilde{\sigma}, p)\phi \end{pmatrix} = mc \begin{pmatrix} \phi \\ \chi \end{pmatrix} \tag{7.22}$$

or

$$\begin{aligned} (\sigma, p)\chi &= mc\,\phi \\ (\tilde{\sigma}, p)\phi &= mc\,\chi , \end{aligned} \tag{7.23}$$

i.e. a coupled set of two 2x2 equations.

Both fields satisfy the Klein - Gordon equation because of (7.3), as can be verified using (7.16):

$$p^2 \chi = m^2c^2\chi \quad , \quad p^2 \Phi = m^2c^2\Phi \qquad (7.24)$$

7.2 Lorentz covariance of the Dirac equation

In the non-relativistic description of spin $\frac{1}{2}$, we already saw that the classical rotation group SO(3) was replaced by the quantum mechanical rotation group SU(2) which, roughly speaking, covers SO(3) twice. A similar structure exists in relativistic quantum mechanics. The classical Lorentz group SO(3,1) will be replaced by SL(2, $\mathbb{C}$) which also covers SO(3,1) twice.

Whereas 4 vectors transform under SO(3,1), spinors transform under SL(2, $\mathbb{C}$) (as the spinorial representation of the Lorentz group (special linear group of 2x2 matrices with complex components and determinant = 1).

We already found that any hermitean 2x2 matrix can be written as

(σ, ξ) with a real 4 vector ξ^μ. Then, also

$$A\,(\sigma, \xi)\,A^+ \qquad \text{with} \quad A \in SL(2,\mathbb{C}) \tag{7.25}$$

is hermitean, such that another 4 vector η^μ must exist, satisfying

$$A\,(\sigma, \xi)\,A^+ = (\sigma, \eta) \tag{7.26a}$$

$$= (\sigma, \Lambda(A)\,\xi) \tag{7.26b}$$

The mapping (7.26a) defines a linear transformation on the 4 vector space

$$\xi \to \eta = \Lambda(A)\,\xi \tag{7.27}$$

As $A \in SL(2,\mathbb{C})$, $\Lambda(A)$ is a representation of this group with the multiplication property

$$(\sigma, \Lambda(A_1 A_2)\,\xi) = A_1 A_2\,(\sigma, \xi)\,A_2^+ A_1^+$$

$$= A_1\,(\sigma, \Lambda(A_2)\,\xi)\,A_1^+$$

$$= (\sigma, \Lambda(A_1)\,\Lambda(A_2)\,\xi)\,, \tag{7.28}$$

as it should be. Also $\Lambda(\mathbb{1}_2) = \mathbb{1}_4$ obviously holds. The group $SL(2,\mathbb{C})$ can be parametrized by 6 real parameters:

$$8 \, {\scriptstyle (\text{complex } 2\times 2 \text{ components})} - 2 \, {\scriptstyle (\text{constraint: } \det A = 1)} = 6 , \quad (7.29)$$

as also holds for the Lorentz group. In order to verify that $\Lambda(A)$ indeed is a Lorentz transformation the scalar product (ξ,ξ) has to remain invariant. In fact:

$$(\xi,\xi) = \det(\sigma,\xi) \overset{\det A = 1}{=} \det A(\sigma,\xi)A^+$$

$$= \det(\sigma, \Lambda(A)\xi) = (\Lambda(A)\xi, \Lambda(A)\xi) , \quad (7.30)$$

as required for a Lorentz transformation.

(NB: As $SL(2,\mathbb{C})$ is connected, the image of $SL(2,\mathbb{C})$ under Λ reproduces the orthochronous "proper" Lorentz group $L_+^\uparrow$ which is one of the four disconnected components of the Lorentz group; the other three components are related to the proper group by time inversion and parity flips. However $SL(2,\mathbb{C})$ covers the proper Lorentz group twice, as $\pm \mathbb{1}_{2\times 2}$ are both mapped onto $\mathbb{1}_{4\times 4}$. The mathematical statement is

$$L_+^\uparrow = SL(2,\mathbb{C})/\mathbb{Z}_2 \quad . \qquad) \tag{7.31}$$

Lorentz transformation of Spinors

The formulation of Lorentz transformations as $SL(2,\mathbb{C})$ matrices lead naturally to the corresponding transformation rules of spinors. Let us first start with the transformation of 2-component left- and right-handed spinors, i.e.

$$(\varphi, \chi)(x) \to (\varphi', \chi')(x') \ , \tag{7.32}$$

such that $\varphi'(x')$ and $\chi(x')$ satisfy the Dirac equation in x' coordinates if $\varphi(x)$ and $\chi(x)$ do so in the x coordinate frame. We demand (cf. (7.23))

$$mc\,\varphi' \overset{!}{=} (\sigma, p')\chi' = (\sigma, \Lambda(A))\chi' \underset{(7.26)}{=} A\,(\sigma, p)A^+ \chi'$$

$$mc\,\chi' \overset{!}{=} (\tilde{\sigma}, p')\varphi' = (\tilde{\sigma}, \Lambda(A)p)\varphi' = A^{+-1}(\tilde{\sigma}, p)A^{-1}\varphi' \tag{7.33}$$

(NB: where Λ has to be chosen as an element of the orthochronous Lorentz group at this point here)

We observe that (7.33) is fulfilled, if the spinors transform as

$$\phi'(x') = A\,\phi(x) \quad , \quad \chi'(x') = A^{+-1}\,\chi(x) \qquad (7.34)$$

and $\phi(x)$ and $\chi(x)$ satisfy the Dirac equation (7.23).

Also parity transformations can be included in the transformation laws. For coordinates or momenta, parity is defined as

$$x^\mu = (x^0, \vec{x}) \longrightarrow (Px)^\mu = (x^0, -\vec{x}) = x' \qquad (7.35)$$

$$p^\mu = (p^0, \vec{p}) \longrightarrow (Pp)^\mu = (p^0, -\vec{p}) = p'$$

A representation of P is obviously given by

$$(P)^\mu{}_\nu = g_{\mu\nu} . \qquad (7.36)$$

Note that σ^μ and $\tilde{\sigma}^\mu$ in (7.12) behave similarly, such that

$$(\sigma, p') = (\tilde{\sigma}, p) \quad , \quad (\tilde{\sigma}, p') = (\sigma, p) . \qquad (7.37)$$

From this, it is obvious that the

parity transformation rule for spinors

$$P: \quad \phi'(x') = \chi(x), \quad \chi'(x') = \phi(x) \qquad (7.38)$$

leads to parity covariance of the Dirac equation. This is an important difference to the non-relativistic theory, in which spinors transform into spinors under parity. In the relativistic case, parity can only be conserved, if both left- and right-handed spinors are present.
Vice versa, if a theory only has, say, right-handed spinors, it necessarily breaks parity.

For instance, neutrinos in the standard-model of particle physics only interact as left-handed spinors (if it had also a right-handed independent component, we couldn't tell from any experiment). If neutrinos were massless (which nowadays is considered as disproved by neutrino oscillation experiments), they would be described by Weyl's equation

$$(\tilde{\sigma}, p)\,\phi = 0 \qquad (7.39)$$

in which no coupling to the χ components occurs anymore.

Lorentz transformations of Dirac spinors

We can construct the transformation rules of the Dirac spinor from that of the Weyl spinors:

$$\Psi'(x') = \begin{pmatrix} \varphi'(x') \\ \chi'(x') \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & A^{+-1} \end{pmatrix} \begin{pmatrix} \varphi(x) \\ \chi(x) \end{pmatrix}$$

$$=: S(A)\,\Psi(x), \qquad (7.40)$$

where S is used as a general symbol for a representation of a transformation in "Dirac space". For instance, we write parity transformations as

$$\Psi'(x') = \begin{pmatrix} \varphi'(x') \\ \chi'(x') \end{pmatrix} = \begin{pmatrix} \chi(x) \\ \varphi(x) \end{pmatrix} = \begin{pmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{pmatrix} \begin{pmatrix} \varphi(x) \\ \chi(x) \end{pmatrix} = \gamma^0 \Psi(x)$$

$$=: S(P)\,\Psi(x) \qquad \text{with } S(P) \equiv \gamma^0 \qquad (7.41)$$

Coming back to the Lorentz transformation (7.41), we observe that

$$S(A)\,(\gamma, p)\,S^{-1}(A) \overset{(7.17)}{=} \begin{pmatrix} A & 0 \\ 0 & A^{\dagger -1} \end{pmatrix} \begin{pmatrix} 0 & (\sigma, p) \\ (\tilde{\sigma}, p) & 0 \end{pmatrix} \begin{pmatrix} A^{-1} & 0 \\ 0 & A^{\dagger} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & A(\sigma, p)A^{\dagger} \\ A^{\dagger -1}(\tilde{\sigma}, p)A^{-1} & 0 \end{pmatrix} = (\gamma, \Lambda(A)\,p)$$

$$= \gamma^{\mu} \Lambda_{\mu}{}^{\nu} p_{\nu} \qquad (7.42)$$

As this holds for any p_{ν}, we deduce

$$\underline{\underline{S(A)\,\gamma^{\mu}\,S^{-1}(A) = \gamma^{\nu}}}\,\Lambda_{\nu}{}^{\mu} \qquad (7.43)$$

This states that γ^{μ}, in fact, transforms as a 4-vector under Lorentz transformations, and implies that products such as $\gamma^{\mu} p_{\nu}$ are indeed Lorentz invariant.

As a 4-vector, j^μ should also transform accordingly under parity transformations. This can immediately be verified,

$$\underbrace{S^{-1}(P)}_{=\gamma^0}\, \gamma^0\, \underbrace{S(P)}_{=\gamma^0} = \gamma^0 \quad , \quad S^{-1}(P)\gamma^i S(P) = \gamma^0 \gamma^i \gamma^0 = -\gamma^i \,. \tag{7.44}$$

For later purposes, we note that the inverse S^{-1} can be related to the hermitean conjugate,

$$S^+ = \begin{pmatrix} A^+ & 0 \\ 0 & A^{-1} \end{pmatrix} \quad , \quad S^{-1} = \begin{pmatrix} A^{-1} & 0 \\ 0 & A^+ \end{pmatrix} \tag{7.45}$$

$$\Rightarrow \qquad \underline{\gamma^0 S^+ \gamma^0 = S^{-1}} \,. \tag{7.46}$$

For the construction of covariant spinor-bilinears in Dirac space, it is thus useful to introduce not only the hermitean conjugate row-like spinor ψ^+, but also the <u>Dirac conjugate object</u>

$$\overline{\psi} := \psi^{+}\gamma^{0} , \qquad (7.47)$$

which transforms as

$$\overline{\psi}(x') = (S\psi(x))^{+}\gamma^{0} = \psi^{+}_{(x)} S^{+}\gamma^{0} \overset{(7.46)}{=} \psi^{+}\gamma^{0}S^{-1}$$

$$= \overline{\psi}(x)\, S^{-1} \qquad (7.48)$$

(c.f. (7.41): $\psi'(x') = S\psi(x)$) .

The decomposition into left- and right-handed components can also be formalized with the aid of projectors. For this, we observe that there exists another matrix

$$\gamma_5 := i\gamma^0\gamma^1\gamma^2\gamma^3 \qquad (7.49)$$

which anti-commutes with all γ^{μ}'s,

$$\{\gamma_5, \gamma^{\mu}\} = 0 . \qquad (7.50)$$

In the chiral representation, its explizit form is

$$\gamma_5 = \sigma_3 \otimes \sigma_0 = \begin{pmatrix} \sigma_0 & 0 \\ 0 & \sigma_0 \end{pmatrix} = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & -1 & \\ 0 & & & -1 \end{pmatrix} \qquad (7.51)$$

Note also that

$$\gamma_5^{+} = \gamma_5 \quad , \quad (\gamma_5)^2 = \mathbb{1} \quad \text{and}$$

$$S^{-1}(P)\,\gamma_5\, S(P) = -\gamma_5 \, . \qquad (7.52)$$

With these properties, we can verify that the projectors

$$P_L = \frac{1}{2}\left(\mathbb{1} + \gamma_5\right) \quad , \quad P_R = \frac{1}{2}\left(\mathbb{1} - \gamma_5\right) \qquad (7.53)$$

satisfy

$$P_L \psi = \begin{pmatrix} \phi \\ 0 \end{pmatrix} \quad , \quad P_R \psi = \begin{pmatrix} 0 \\ \chi \end{pmatrix} \qquad (7.54)$$

and the projector properties

$$\begin{aligned} P_L + P_R &= \mathbb{1} && \text{(completeness)} \\ P_{L,R}^2 &= P_{L,R} && \text{(idempotency)} \\ P_L P_R &= P_R P_L = 0 && \text{(orthogonality)} \end{aligned} \qquad (7.55)$$

With these prerequisites, we can introduce a number of (physically relevant) bilinear (pseudo) tensor fields.

Let us start with the example

$$j^\mu(x) = \bar{\psi}\gamma^\mu\psi . \tag{7.56}$$

We can directly verify that it transforms as

$$j^{\mu\,\prime}(x') = \bar{\psi}'(x')\gamma^\mu\psi'(x') \overset{(7.48)}{=} \bar{\psi}(x)\underbrace{S^{-1}(\Lambda)\gamma^\mu S(\Lambda)}_{\Lambda^\mu{}_\nu\gamma^\nu}\psi(x)$$

$$\overset{(7.43)}{=} \Lambda^\mu{}_\nu\, j^\nu(x) , \tag{7.57}$$

i.e. as a 4-vector field.

In the same manner (and also checking the parity transformation properties) it is straightforward to show the transformation properties of the following 16 (pseudo)-tensor fields:

$$S(x) = \bar{\Psi}(x)\,\Psi(x) \qquad \text{scalar} \qquad 1$$

$$j^\mu \equiv V^\mu = \bar{\Psi}\gamma^\mu\Psi \qquad \text{vector} \qquad 4$$

$$T^{\mu\nu} = \bar{\Psi}(x)\,[\gamma^\mu,\gamma^\nu]\,\Psi(x) \qquad \text{antisym. tensor} \qquad 6$$

$$A^\mu = \bar{\Psi}\gamma_5\gamma^\mu\Psi \qquad \text{pseudovector} \qquad 4$$

$$P = \bar{\Psi}\gamma_5\Psi \qquad \text{pseudo scalar} \qquad 1$$

In fact, on the bilinear level, this exhausts all possible tensor fields as the Dirac algebra generates 16 elements: $\Big(\underbrace{4\times 4}_{\text{matrices}} \cdot \underset{\substack{\uparrow \\ \text{complex} \\ \text{components}}}{2} - \underset{\substack{\uparrow \\ \text{constraints} \\ \text{from } \{\gamma^\mu,\gamma^\nu\} = 2g^{\mu\nu}\mathbb{1}}}{16} = 16 \Big)$.

(NB: a generalization to other dimensions is less straightforward as it may seem. E.g. γ_5-type matrices in irreducible representations exist only in even dimensions.)

This tensor fields occur, for instance, as currents in the strong or weak interactions, or as condensates in chiral theories.

7.3 Plane - wave solutions

For a plane-wave solution of the Dirac equation, let us make the ansatz

$$\psi_p = \frac{1}{\sqrt{2\omega(\vec{k})}} e^{-ikx} u_p \qquad (7.58)$$

where $p^\mu = \hbar k^\mu$ and

$$\omega(\vec{k}) = +\sqrt{\vec{k}^2 + \mu^2} \qquad (7.59)$$

is the angular frequency corresponding to the relativistic energy $E = \hbar\omega$ in the energy-momentum relation.

ψ_p is a solution of the free Dirac equation, if the spinor u_p satisfies

$$(\not{k} - \mu)u_p \equiv (\gamma^\mu k_\mu - \mu)u_p = 0 \qquad (7.60)$$

Whereas the Dirac equation including the correspondence rule $p_\mu \to i\hbar\partial_\mu$ is a differential equation, (7.60) is just an algebraic equation.

For simplicity, let us take a look at the "particle at rest", $\vec{p} = \vec{k} = 0$. Then (7.60) reduces to

$$\gamma^0 k_0 \, u_p = \mu \, u_p \qquad (7.61)$$

$$\gamma^0 k_0 = \begin{pmatrix} 0 & \begin{matrix} k_0 & 0 \\ 0 & k_0 \end{matrix} \\ \begin{matrix} k_0 & 0 \\ 0 & k_0 \end{matrix} & 0 \end{pmatrix}$$

One eigenvector is, for instance, given by $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, yielding the eigenvalue k_0, and thus a positive energy solution

$$k_0 = \mu$$

$$\text{or} \quad p_0 c = \hbar \omega \quad \text{after a Lorentz boost} \qquad (7.62)$$

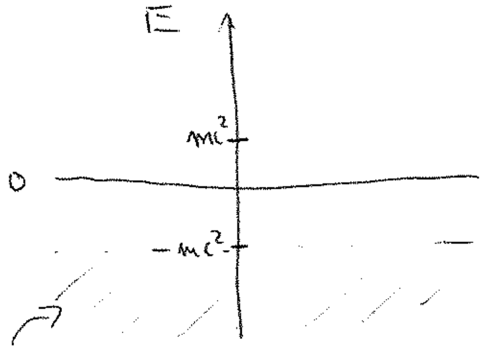
However, another eigenvector is given by $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}$, leading to a negative energy solution

$$k_0 = -\mu$$

$$\text{or} \quad p_0 c = -\hbar \omega \quad \text{after a Lorentz boost} \qquad (7.63)$$

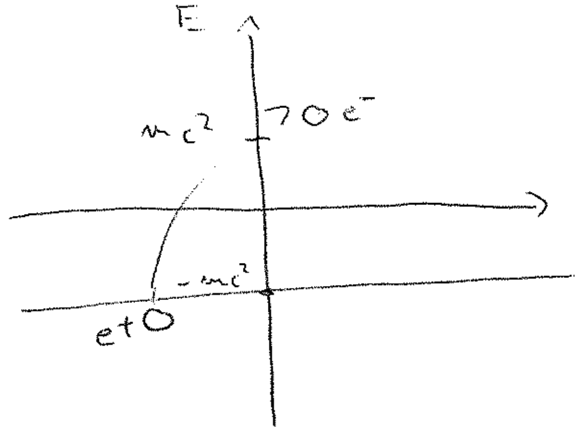
Obviously, the first-order nature of the Dirac equation as a differential equation in time does not solve the negative energy "problem".

After all, it is no longer a problem as it used to be considered historically, as these states can be associated with anti-matter. Dirac himself invoced Pauli's exclusion principle which applies as the Dirac equation describes fermions. He asserted that the vacuum is the state where all negative energy states are occupied. Then we cannot gain energy from positive-energy states dropping down to negative energies due to Pauli blocking.



all states occupied in vacuum
"Dirac sea"

Still, a negative energy state can be lifted to positive energies, implying that a hole remains in the "Dirac sea"



As the hole carries positive charge in the case of electrons, he interpeted and thus predicted the hole as the existence of antimatter, a "positron", a few years before its discovery.

The concept of the Dirac sea has been very important historically for an understanding of negative energy/density states as antimatter. However, there is no exclusion principle and thus no Dirac sea for bosons. Still the reinterpretation of negative energy/density states as antimatter also holds for bosons. Therefore, the Dirac sea should be considered as a useful picture, but not as an essential building block of the formalism and concepts.

Returning to the algebraic equation (7.60),

$$(\not{k} - \mu) u_p = 0 , \tag{7.64}$$

we use the following normalization convention for the spinors u_p:

$$\bar{u}_p u_p = 2\mu \, \mathrm{sign}(p_0) . \tag{7.65}$$

This fixes the solutions of (7.64) in terms of all eigenvalues of $\not{k}$ for a given k^μ, and thus defines the plane-wave solutions of the Dirac equation.

Let us specifically consider a positive-energy solution. Multiplying (7.64) with $\bar{u}_p$ yields

$$\bar{u}_p \not{k} u_p = \mu \, \bar{u}_p u_p \underset{p^0>0}{\overset{(7.65)}{=}} 2\mu^2 = 2k^2$$

$$\Rightarrow \quad \bar{u}_p \gamma^\mu u = 2k^\mu \tag{7.66}$$

The resulting 4-vector current then yields

$$j^\mu = \bar{\psi}_p \gamma^\mu \psi_p = \frac{1}{2\omega} \bar{u}_p \gamma^\mu u_p = \frac{k^\mu}{\omega} \tag{7.67}$$

$$=: (1, \vec{v}) \quad , \quad \vec{v} = \frac{\vec{k}}{\omega} \, ,$$

where $\vec{v}$ can be interpreted as the velocity of a particle. The normalization in the ansatz (7.58), was actually chosen such that the particle density is normalized to unity (one particle per unit volume).

7.4 Coupling to the electromagnetic field

As in the case of the Klein-Gordon equation, we obtain the Dirac equation in an electromagnetic field by the replacement

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + \frac{ie}{\hbar c} A_\mu \,, \tag{7.68}$$

yielding

$$\begin{aligned} \left(i\not{D} - \frac{mc}{\hbar} \right) \psi &= \left(i\not{D} - \mu \right) \psi \,, \\ \not{D} &= \gamma^\mu D_\mu \,. \end{aligned} \tag{7.69}$$

As also the covariant derivative D_μ alone, the Dirac operator $\not{D}$ transforms covariantly under gauge transformations

$$\not{D}_{A'} = e^{\frac{ie}{\hbar c}\lambda} \not{D}_A \, e^{-\frac{ie}{\hbar c}\lambda} \,, \quad A'_\mu = A_\mu - \partial_\mu \lambda \tag{7.70}$$

This implies that the local field transformation

$$(A, \psi) \rightarrow (A', \psi')$$

with

$$\psi'(x) = e^{\frac{ie}{\hbar c}\lambda(x)}\,\psi(x) \qquad (7.71)$$

is a (gauge) symmetry of the Dirac equation. As $e^{\frac{ie}{\hbar c}\lambda(x)} \in U(1)$, we call this a $U(1)$-symmetry. The corresponding Noether current (-> excercize) is given by

$$j^\mu = \bar{\psi}(x)\,\gamma^\mu\,\psi(x) \quad , \partial_\mu j^\mu = 0 \qquad (7.72)$$

and is conserved.

7.5 Hamiltonian formalism

For late purposes, it is useful to rewrite the Dirac equation into a Hamiltonian form. For this, we multiply (7.69) by γ^0,

$$i D_0 \psi = -i\gamma^0\gamma^j D_j \psi + \frac{mc}{\hbar}\gamma^0\psi$$

$$\Rightarrow \quad \underline{-\frac{\hbar}{i}\frac{\partial\psi}{\partial t} = H\psi} \qquad (7.73a)$$

where

$$\underline{H = c(\vec{\alpha},\vec{\pi}) + mc^2\beta + eA_0} \qquad (7.73b)$$

where we have introduced the matrices

$$\alpha_k := \gamma^0 \gamma^k \quad \text{and} \quad \beta := \gamma^0 \tag{7.74}$$

and as usual

$$\vec{\pi} = \vec{p} - \frac{e}{c}\vec{A} \tag{7.75}$$

is the kinetic momentum operator.

Let us verify the Hermiticity properties of H for which we need those of $\vec{\alpha}$ and β. We already know that

$$\beta^\dagger = \gamma^{0\dagger} = \gamma^0 = \beta \tag{7.76}$$

Furthermore,

$$\begin{aligned} \underline{\underline{\alpha_j^\dagger}} &= (\gamma^0 \gamma^j)^\dagger = \gamma^{j\dagger} \gamma^{0\dagger} = (-\gamma^j)\gamma^0 \\ &= -\gamma^0 \underbrace{\gamma^0 \gamma^j}_{=-\gamma^j\gamma^0} \gamma^0 = \gamma^0 \gamma^j (\gamma^0)^2 = \gamma^0 \gamma^j = \underline{\underline{\alpha_j}} \end{aligned} \tag{7.77}$$

from which we conclude that

$$H^\dagger = H \,. \tag{7.78}$$

Of course, the algebra of α_k and β follows directly from the Dirac algebra. Defining

$$\alpha_4 := \beta \,, \qquad (7.79)$$

this algebra can be summarized as

$$\{\alpha_\mu, \alpha_\nu\} = 2\delta_{\mu\nu} \,, \quad \mu,\nu = 1,2,3,4, \qquad (7.80)$$

with $\alpha_\mu^+ = \alpha_\mu$.

Free solutions (in the Dirac representation)

So far, we have used the chiral representation of the γ matrices. Whereas physical observables are independent of the choice of representation (within one equivalence class), some properties become more transparent in certain representations. For instance, the connection between relativistic covariance and parity / chirality discussed above became rather obvious in the chiral representation. In the following we will also use the Dirac and Majorana representation (see exercises) which are convenient for the aspects of positive/negative energy solutions and charge conjugation, respectively.

Let us start with the free Dirac equation in the Hamilton form

$$H\psi = E\psi \, ,$$

$$H = c\,\vec{\alpha}\cdot\vec{p} + mc^2\beta = \alpha_\mu \xi_\mu \qquad (7.81)$$

where

$$\xi_\mu = (cp_1, cp_2, cp_3, mc^2)^T . \qquad (7.82)$$

Using

$$(\alpha_\mu \xi_\mu)^2 = \xi_\mu \xi_\nu \alpha_\mu \alpha_\nu$$

$$= \frac{1}{2}\xi_\mu \xi_\nu \underbrace{\{\alpha_\mu, \alpha_\nu\}}_{=2\delta_{\mu\nu}} = (\xi_\mu)^2 \, , \qquad (7.83)$$

we reproduce the Klein-Gordon equation in the form

$$(H+E)(H-E)\psi = (\xi_\mu \xi_\mu - E^2)\psi$$

$$= (c\vec{p}^{\,2} + m^2c^4 - E^2)\psi = 0 \, . \qquad (7.84)$$

Since $\vec{p}$ and H commute (i.e. are compatible operators), there exists an eigenbasis which diagonalizes both operators. These eigenstates then satisfy

$$E = \pm c \sqrt{m^2c^2 + \vec{p}^2} = \pm \hbar \, \omega(\vec{p}) \qquad (7.85)$$

where $\omega(\vec{p}) > 0$ is the angular frequency corresponding the the momentum $\vec{p}$. We have already seen in (7.61-63) that both positive and negative energy solutions indeed exist. Since ψ is a four-component wave function, both eigenvalues (7.85) are 2-fold degenerate (it is straightforward to see this degeneracy in the chiral representation analogous to (7.61-63)).

As H^2 occuring in (7.84) is diagonal in Dirac space, any plane wave with a constant spinor diagonalizes H^2. However, we are rather interested in the eigenvalues of H. For this, we introduce the projectors

$$P_\pm = \frac{1}{2}\left(\mathbb{1} \pm \frac{H}{\hbar\omega}\right) \qquad (7.86)$$

onto positive and negative energy subspaces. It is straightforward to verify the projector properties

$$P_+ + P_- = \mathbb{1} \quad , \; P_+ P_- = 0 = P_- P_+$$
$$P_\pm^2 = P_\pm \; . \tag{7.87}$$

The decisive property becomes obvious from

$$H \;(P_\pm \psi) = \frac{1}{2} \; H \left(1 \pm \frac{H}{\hbar\omega}\right)\psi = \frac{1}{2}\Big(H \pm \frac{1}{\hbar\omega} \underbrace{H^2}_{=E^2=(\hbar\omega)^2}\Big) \psi$$

$$= \pm \hbar\omega \; \frac{1}{2}\left(1 \pm \frac{H}{\hbar\omega}\right)\psi = \pm \hbar\omega (P_\pm \psi) \, , \tag{7.88}$$

demonstrating that $P_\pm$ projects onto the pos. (neg. energy subspaces.

Let us now illustrate these properties in the Dirac representation of the γ matrices

$$\gamma^0 = \beta = \begin{pmatrix} \sigma_0 & 0 \\ 0 & -\sigma_0 \end{pmatrix} \; , \; \gamma^j = \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix} \tag{7.89}$$

(which is different but unitary equivalent to the chiral representation, cf. exercises).
In this representation, the projectors read

$$P_\pm = \frac{1}{2}\begin{pmatrix} \left(1 \pm \frac{mc^2}{\hbar\omega}\right)\sigma_0 & \pm \frac{c(\vec{\sigma}\cdot\vec{p})}{\hbar\omega} \\ \pm \frac{c}{\hbar\omega}(\vec{\sigma}\cdot\vec{p}) & \left(1 \mp \frac{mc^2}{\hbar\omega}\right)\sigma_0 \end{pmatrix}. \qquad (7.90)$$

For a particle at rest, we have $\vec{p}=0$, $\hbar\omega_0 = mc^2$ and thus simply

$$P_+ = \begin{pmatrix} \sigma_0 & 0 \\ 0 & 0 \end{pmatrix}, \quad P_- = \begin{pmatrix} 0 & 0 \\ 0 & \sigma_0 \end{pmatrix}. \qquad (7.91)$$

Positive energy solutions (at rest) hence occur as the upper components of the Dirac spinor in the Dirac representation and negative energy solutions (at rest) as lower components

(CAVE: upper/lower components in the Dirac representation should not be confused with left-/right-handed spinors in the chiral representation.)

7.6 Charge conjugation

For simplicity, we use the units

$$\hbar = c = 1 \qquad (7.92)$$

in the following. Using dimensional analysis, the correct $\hbar$ and c factors can always be reinstated in the final results <u>unambiguously</u>.

Next, we use a third representation of the Dirac algebra: the Majorana <u>representation</u>. This representation is special as all γ matrices have purely imaginary components:

$$\begin{aligned} \gamma^0 &= \begin{pmatrix} \sigma_2 & 0 \\ 0 & \sigma_2 \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} i\sigma_3 & 0 \\ 0 & i\sigma_3 \end{pmatrix}, \\ \gamma^2 &= \begin{pmatrix} 0 & i\sigma_1 \\ i\sigma_1 & 0 \end{pmatrix}, \quad \gamma^3 = \begin{pmatrix} i\sigma_1 & 0 \\ 0 & -i\sigma_1 \end{pmatrix} \end{aligned} \qquad (7.93)$$

and

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & -i\sigma_1 \\ i\sigma_1 & 0 \end{pmatrix}.$$

Because of this property, charge conjugation works similarly to the Klein-Gordon-case by complex conjugation :

$$\psi \to \psi^* = (\psi^+)^T . \qquad (7.94)$$

To see this, let us complex conjugate the Dirac equation in an electromagnetic field,

$$0 = (i\gamma^\mu(\partial_\mu + ieA_\mu) - m)\psi$$

$$\xrightarrow{c.c.} (i\gamma^\mu(\partial_\mu - ieA_\mu) - m)\psi^* = 0 . \qquad (7.95)$$

So if ψ satisfies the Dirac equation with charge e, ψ^* satisfies the same Dirac equation with charge $(-e)$.

The charge conjugated spinor can also be written in terms of the Dirac conjugated spinor, since

$$\bar\psi = \psi^+\gamma^0 \Rightarrow \psi^+ = \bar\psi\gamma^0$$

$$\Rightarrow \psi^* \overset{(7.94)}{=} (\bar\psi\gamma^0)^T \underset{(7.93)}{=} -\gamma^0\bar\psi^T . \qquad (7.96)$$

It is a standard convention to define the charge conjugated spinor as complex conjugation in the Majorana representation including a phase factor:

$$C: \quad \psi \to \psi_c = -i\,\psi^* = C\,\overline{\psi}^T \qquad (7.97)$$

(valid in any representation) with

$$C = i\gamma^0 \quad \text{in Majorana representation.} \qquad (7.98)$$

Charge conjugation is one of the fundamental discrete symmetries in quantum field theory (that may or may not hold for specific theories). In the Majorana representation, we can directly verify that

$$C\,C^\dagger = \mathbb{1}\,, \quad C^{-1} = C^\dagger\,, \quad C^T = -C^T \qquad (7.99)$$

and

$$C^{-1}\gamma^\mu C = -(\gamma^\mu)^T.$$

Due to the connection with complex conjugation, it also becomes clear that a solution (say for an electron) with positive energy $\psi \sim e^{-iEt}$ turns into a negative-energy solution $\psi_c \sim e^{+iEt}$ under charge conjugation ($\Rightarrow$ a positron-like solution).

Hence, the "positive/negative-energy problem" can be solved by reinterpretation of negative-energy solutions as antimatter states with opposite charge, similarly as we discussed it for the Klein-Gordon equation.

(NB1 explicit forms of the charge conjugation operator in other unitary equivalent representations can immediately be found with the transformation

$$\psi \to \tilde{\psi} = U\psi, \quad \tilde{\gamma}^\mu = U\gamma^\mu U^{-1}. \qquad (7.100)$$

Charge conjugation in this new representation then has the form

$$\tilde{C} = U C U^T \qquad (7.101)$$

which preserves (7.97): $\tilde{\psi}_c = \tilde{C}\bar{\tilde{\psi}}^T$.)

(NB2: Spinor particles which do not carry an electric charge can potentially satisfy

$$\psi = \psi_c \qquad (7.102)$$

If read as a constraint (7.102) effectively removes 2 degrees of freedom from the Dirac spinor. Spinors that satisfy (7.102) are called Majorana spinors. As they form there own antimatter, particle numbers are typically not conserved in processes involving

Majorana particles, as opposed to the analogous theory with Dirac particles. In the particle physics standard model, neutrinos do not carry electric charge. If they are Dirac particles, lepton number would still be a conserved quantity. If they are Majorana fermions, lepton number violating processes are possible. A famous test case is neutrinoless double beta decay which is actively searched for.)

7.7 Nonrelativistic limit

For a study of the non-relativistic limit of the Dirac equation, let us again return to the <u>Dirac</u> representation.

As the relativistic energy contains also the rest Energy $E = mc^2 + \ldots$, it seems advisable to factorize the rapidly oscillating phase factor arising from the rest energy:

$$\psi = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} \cdot e^{-imc^2 t} \qquad (7.103)$$

in an ansatz for the Dirac equation

The Hamilton operator in Dirac representation reads

$$H = \begin{pmatrix} (mc^2 + eA_0)\sigma_0 & c\vec{\sigma}\cdot\vec{\pi} \\ c\vec{\sigma}\cdot\vec{\pi} & (-mc^2 + eA_0)\sigma_0 \end{pmatrix} . \qquad (7.104)$$

The $\psi_\pm$ components thus have to satisfy the coupled system

$$\begin{aligned} (i\partial_t - eA_0)\psi_+ &= c\vec{\sigma}\cdot\vec{\pi}\,\psi_- \\ (i\partial_t - eA_0 + 2mc^2)\psi_- &= c\vec{\sigma}\cdot\vec{\pi}\,\psi_+ \end{aligned} \qquad (7.105)$$

For $\vec{p} = 0$ and $A_\mu = 0$ $(\Rightarrow \vec{\pi} = \vec{p})$, (7.105) is solved by $\psi_- = 0$ and $\psi_+ = $ const., yielding a positive-energy solution upon insertion into (7.103). For small $\vec{p}$ and A_μ, we expect that ψ_-/ψ_+ stays comparatively small, i.e. we expect that $(i\partial_t - eA_0)\psi_-$ can be neglected compared with $2mc^2\psi_-$ in (7.105)

$$\Rightarrow \quad \psi_- \simeq \frac{1}{2mc}\vec{\sigma}\cdot\vec{\pi}\,\psi_+ \quad , \vec{\pi} = \vec{p} - \frac{e}{c}\vec{A} \qquad (7.106)$$

Inserting (7.106) into $(7.105)_{\text{1st line}}$, yields

$$(i\partial_t - eA_0)\Psi_+ = \frac{1}{2m}(\vec{\sigma}\cdot\vec{\pi})^2\,\Psi_+ \quad . \tag{7.107}$$

The right-hand side contains

$$(\vec{\sigma}\cdot\vec{\pi})^2 = \underbrace{\sigma_i\sigma_j}_{=\delta_{ij}+i\varepsilon_{ijk}\sigma_k}\pi_i\pi_j = \vec{\pi}^2 + i\,\varepsilon_{ijk}\,\sigma_k\,\pi_i\pi_j$$

$$= \vec{\pi}^2 + \frac{i}{2}\,\varepsilon_{ijk}\,\sigma_k\underbrace{[\pi_i,\pi_j]}_{=ie\,\varepsilon_{ijk}B_k} \tag{7.108}$$

so that this approximation to the Dirac equation yields

$$i\hbar\,\partial_t\Psi_+ = H_p\,\Psi_+ \tag{7.109}$$

where

$$H_p = \frac{1}{2m}\vec{\pi}^2 + eA_0 - \frac{e\hbar}{2mc}\,\vec{\sigma}\cdot\vec{B} \tag{7.110}$$

denotes the nonrelativistic Pauli-Hamilton-Operator. Introducing $\vec{S} = \frac{1}{2}\hbar\vec{\sigma}$ for the spin operator, we can read off the electron's magnetic moment

$$\mu = \frac{e\hbar}{2mc}\,\vec{\sigma} = \frac{e}{mc}\,\vec{S} \tag{7.111}$$

Writing the magnetic moment as usual as

$$\vec{\mu} = g \, \mu_B \frac{\vec{S}}{\hbar} \, , \quad \mu_B = \frac{e\hbar}{2mc} \tag{7.112}$$

Bohr's magneton

we can read off the gyromagnetic factor of the Dirac electron,

$$\underline{\underline{g = 2}} \, , \tag{7.113}$$

as is in good agreement with experiment. Further small deviations

$$(g-2)_{exp} \sim \frac{\alpha}{\pi} \sim 2 \cdot 10^{-3} \tag{7.114}$$

are very well understood in the framework of quantum electrodynamics (QED).

The non-relativistic limit can be formalized more precisely in terms of the Foldy - Wouthuysen (FW) transformation

For $\vec{p}=0$ and $A_\mu=0$, the Hamiltonian (in Dirac representation) reads

$$H = \begin{pmatrix} mc^2\sigma_0 & 0 \\ 0 & -mc^2\sigma_0 \end{pmatrix} \qquad (7.115)$$

where each component is a 2x2 block. The eigenvectors also decompose into a 2-component structure

same 2-comp. spinor

$$E>0: \quad \psi \sim \begin{pmatrix} \square \\ 0 \end{pmatrix} \quad , \quad E<0: \quad \psi \sim \begin{pmatrix} 0 \\ \square \end{pmatrix} \qquad (7.116)$$

For $\vec{p}\neq 0$ or $A_\mu \neq 0$, the 2-component upper and lower components are coupled. The FW transformation now aims at a decoupling of these components by means of a unitary transformation

$$\psi' \to e^S \psi \quad , \quad S^\dagger = -S \; . \qquad (7.117)$$

S should be constructed such that

The Hamiltonian has the form

$$H' = \begin{pmatrix} H'_+ & 0 \\ 0 & H'_- \end{pmatrix} \qquad (7.118)$$

For the nonrelativistic limit, we construct this transformation order by order in $\frac{\vec{p}}{mc}$. For this, we define two classes of operators with the 2x2 block form

$$\mathcal{E} = \begin{pmatrix} \square & 0 \\ 0 & \square \end{pmatrix} \quad \text{"even"}$$
$$\mathcal{O} = \begin{pmatrix} 0 & \square \\ \square & 0 \end{pmatrix} \quad \text{"odd"} \qquad (7.119)$$

In Dirac representation, where $\beta = \begin{pmatrix} \sigma_0 & 0 \\ 0 & -\sigma_0 \end{pmatrix}$, we observe that

$$\left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \beta\right] = 2\begin{pmatrix} 0 & -b \\ c & 0 \end{pmatrix} \quad \text{and} \quad \left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \beta\right\} = 2\begin{pmatrix} a & 0 \\ 0 & -d \end{pmatrix}, \qquad (7.120)$$

such that even and odd operators can be classified according to

$$[\mathcal{E}, \beta] = 0 \quad \text{and} \quad \{\beta, \mathcal{O}\} = 0 \qquad (7.121)$$

Any operator can be decomposed into a unique sum of even and odd operators. For instance, for the Hamiltonian, we write

$$H = mc^2\beta + \mathcal{E} + \mathcal{O} \quad , \; \mathcal{E} = eA_0 \, , \; \mathcal{O} = c(\vec{\alpha}\cdot\vec{\pi}) \tag{7.122}$$

(Of course, $mc^2\beta$ also belongs to the even part.)

Let us now consider the behavior of the Dirac equation $i\hbar\partial_t\psi = H\psi$ under the transformation (7.117),

$$\begin{aligned} i\hbar\partial_t\psi' &= i\hbar(\partial_t e^S)\psi + i\hbar e^S\partial_t\psi \\ &= i\hbar(\partial_t e^S)\, e^{-S}\psi' + e^S H e^{-S}\psi' \\ &=: H'\psi' \end{aligned} \tag{7.123}$$

with

$$H' = e^S H e^{-S} + i\hbar(\partial_t e^S)\, e^{-S}. \tag{7.124}$$

Since $S=0$ for $\vec{p}=0$ and $A_\mu=0$ (see (7.115)), S, or more precisely $\mathcal{O}$ in (7.112), remains small in the non-relativistic limit. Hence, we study the expansion of (7.124) for small S. The first term yields

$$e^{S} H e^{-S} = H + [S,H] + \frac{1}{2}[S,[S,H]] + \ldots$$
$$=: \sum_{k=0}^{\infty} \frac{1}{k!} (\mathrm{ad}\, S)^k H \,, \qquad (7.125)$$

where ad S is a short form for the linear map

$$(\mathrm{ad}\, S) H = [S,H] \,. \qquad (7.126)$$

For the second term in (7.124), let us first consider

$$\frac{d}{d\lambda}\left((\partial_t e^{\lambda S})\, e^{-\lambda S} \right) = \left(\partial_t\, e^{\lambda S} S \right) e^{-\lambda S} - (\partial_t e^{\lambda S})\, S\, e^{-\lambda S}$$

$$= e^{\lambda S} (\partial_t S)\, e^{-\lambda S}$$

$$\overset{(7.125)}{=} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} (\mathrm{ad}\, S)^k\, \partial_t S \,. \qquad (7.127)$$

Integrating λ from 0 to 1 yields

$$(\partial_t e^{S})\, e^{-S} = \sum_{k=0}^{\infty} \frac{1}{(k+1)!} (\mathrm{ad}\, S)^k\, \partial_t S \,. \qquad (7.128)$$

In summary, we obtain

$$H' = H + \left([S,H] + \frac{1}{2!}[S,[S,H]] + \frac{1}{3!}[S,[S,[S,H]]] + \ldots \right)$$

$$+ i\hbar \left(\partial_t S + \frac{1}{2!} [S, \partial_t S] + \frac{1}{3!} [S,[S,\partial_t S]] + \ldots \right) \qquad (7.129)$$

Let us now try to apply this transformation such that H' becomes diagonal at least to a desired order in c (or $1/c$). For this, we note that the leading term in $\frac{1}{c}$ is $mc^2\beta$. $\mathcal{E}$ and $\mathcal{O}$ are subleading: ($\mathcal{O} \sim c$ and $\mathcal{E} \sim c\left(\frac{e}{c}A_0\right)$).

Obviously, we need to construct an appropriate S. As S is expected to have an expansion in $\frac{1}{c}$ starting with $\left(\frac{1}{c}\right)^1$, as $S=0$ for $c \to \infty$, the leading order correction / transformation is (for $\partial_t S = 0$)

$$H' = mc^2\beta + \underbrace{\mathcal{E} + \mathcal{O}}_{\mathcal{O}(c)} + \underbrace{[S, mc^2\beta]}_{\mathcal{O}(c)} + \mathcal{O}(c^0) \qquad (7.130)$$

Obviously H' becomes even (diagonal), if

$$[S, mc^2\beta] = -\mathcal{O} \qquad (7.131)$$

(7.131) is solved by

$$S = \frac{1}{2mc^2}\,\beta\cdot\mathcal{O}\,, \qquad (7.132)$$

Since

$$\underline{[S, mc^2\beta]} = \frac{1}{2}[\beta\mathcal{O}, \beta] = \frac{1}{2}\beta[\mathcal{O}, \beta]$$

$$= \frac{1}{2}\beta\underbrace{\mathcal{O}\beta}_{=-\beta\mathcal{O}+\underbrace{\{\mathcal{O},\beta\}}_{=0}} - \frac{1}{2}\beta^2\mathcal{O} = -\beta^2\mathcal{O} = \underline{-\mathcal{O}} \qquad (7.133)$$

This transformation brings H' into even form to order c^1. Still, all higher order terms can be computed to arbitrary order using (7.129). The result will be of the form

$$H' = mc^2\beta + \mathcal{E}' + \mathcal{O}' \qquad (7.134)$$

where $\mathcal{E}' \sim \mathcal{O}(c)$ and $\mathcal{O}' \sim \mathcal{O}(c^0)$. We can now iterate the same strategy, i.e. perform a transformation with

$$S' \underset{(7.132)}{=} \frac{1}{2mc^2}\beta\mathcal{O}' \sim \mathcal{O}\left(\frac{1}{c^2}\right) \qquad (7.135)$$

such that

$$H'' = \underbrace{H'}_{\mathcal{O}(c^2)} + \underbrace{[\underbrace{S'}_{\mathcal{O}(\frac{1}{c^2})}, H']}_{\mathcal{O}(c^0)} + \mathcal{O}\left(\frac{1}{c^2}\right) \qquad (7.136)$$

$$= mc^2\beta + \underbrace{\mathcal{E}''}_{\mathcal{O}(c)} + \underbrace{\mathcal{O}''}_{\mathcal{O}(\frac{1}{c})} + \mathcal{O}\left(\frac{1}{c^2}\right) \qquad (7.137)$$

Another iteration decouples the Hamilton operator to order $\frac{1}{c^2}$, yielding (without explicit calculation)

$$H''' = \beta\left(mc^2 + \frac{1}{2m}\vec{\pi}^2 - \frac{1}{8m^3c^2}(\vec{\pi}^2)^2 + \dots\right) + eA_0$$

$$- \frac{e\hbar}{2mc}\,\vec{\sigma}\cdot\vec{B}\,.\,\beta$$

$$- \frac{ie\hbar}{8m^2c^3}\left(\vec{\sigma}\cdot(\vec{\nabla}\times\vec{E})\right) - \frac{e\hbar}{4m^2c^2}\,\vec{\sigma}\cdot(\vec{E}\times\vec{p})$$

$$- \frac{e^2\hbar^2}{8m^2c^2}\,\vec{\nabla}\cdot\vec{E} + \dots \qquad (7.138)$$

The terms in the first parenthesis clearly corresponds to the non-relativistic expansion of the energy-momentum relation

$$c\sqrt{m^2c^2+\vec{\pi}^2} \sim mc^2\left(1+\frac{\vec{\pi}^2}{2m^2c^2}-\frac{1}{8m^4c^4}(\vec{\pi}^2)^2+\ldots\right)$$

already discussed before. (7.139)

The second line in (7.138) is the Pauli term (now in 4-component version) already discovered in our heuristic discussion of the NR limit (cf. (7.110)).

For the 3rd line, we consider a spherical potential $\vec{E} = -\vec{\nabla}\phi(r)$, such that $\vec{\nabla}\times\vec{E}=0$ and

$$\vec{E} = -\vec{\nabla}\phi(r) = \frac{\vec{x}}{r}\partial_r\phi(r) \tag{7.140}$$

yielding

3rd line : $$-\frac{ie\hbar}{8m^2c^3}\vec{\sigma}\cdot\underbrace{(\vec{\nabla}\times\vec{E})}_{=0}-\frac{e\hbar}{4m^2c^2}\vec{\sigma}\cdot(\vec{E}\times\vec{p})$$

$$= \frac{e\hbar}{4m^2c^2}\vec{\sigma}\cdot(\vec{x}\times\vec{p})\frac{\phi'(r)}{r}$$

$\vec{S}=\frac{\hbar}{2}\vec{\sigma}$ $$= \frac{e}{2m^2c^2}\vec{S}\cdot\vec{L}\frac{\phi'(r)}{r} \tag{7.141}$$

This is exactly the "classical" spin-orbit coupling (no $\hbar$ any more) including the correct prefactor from Thomas precession that we had used in (3.97) for determining the fine structure. The last term is the so-called <u>Darwin term</u> which is only nonzero "inside" a source where $\vec{\nabla}\cdot\vec{E} \neq 0$. In the fine structure calculation this term contributes only for the $l=0$ states with non-vanishing wave function at the origin justifying the use of (3.105) also for the $l=0$ states.

7.8 Relativistic Coulomb problem

We will not solve the Dirac equation for the Coulomb problem explicitly here.

The general strategy is rather similar to standard central-force problems: first the angular-momentum eigenstates are constructed and finally the radial eigenfunctions are determined.

For the angular problem a new ingredient is given by the operator (compatible with H, $[K, H] = 0$)

$$K = \beta\left(\frac{\vec{L}}{\hbar}\cdot\vec{\sigma} + \mathbb{1}\right) \qquad (7.142)$$

which measures the parallelism of spin and orbital angular momentum. Its eigenvalues turn out to be

$$k = \pm 1, \pm 2, \pm 3, \ldots, \qquad (7.143)$$

the modulus of which correspond to

$$|k| = j + \frac{1}{2} \qquad (7.144)$$

These eigenvalues together with the principle quantum number n finally show up in the energy spectrum for hydrogen like atoms

$$\frac{E}{mc^2} = \left(1 + \frac{Z^2\alpha^2}{\left(n - |k| + \sqrt{k^2 - Z^2\alpha^2}\right)^2}\right)^{-\frac{1}{2}} \qquad (7.145)$$

$n = 1, 2, 3, \ldots \qquad |k| = 1, 2, 3, \ldots$

but $\quad |k| \leq n$

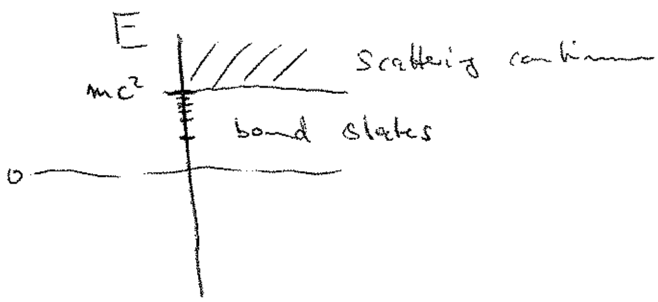
This implies that all states except $n = |k|$ are doubly degenerate.

The expansion of $\frac{E}{mc^2}$ in powers of $Z\alpha$ yields

$$E = mc^2 \left(1 - \frac{(Z\alpha)^2}{2n^2} + \frac{3}{8}\frac{(Z\alpha)^4}{n^4} - \frac{1}{2}\frac{(Z\alpha)^4}{n^3}\frac{1}{j+\frac{1}{2}} + \mathcal{O}\left((Z\alpha)^6\right) \right) \tag{7.146}$$

Corresponding to the rest energy, the Balmer term and the fine structure found in (3.105).

The strength of the potential is parameterized by Z. On the energy axis the bound states lie slightly (for small Z) below the rest mass mc^2 where the scattering continuum starts.



For increasing Z the bound states are dragged more and more into the energy gap between mc^2 and $-mc^2$. The binding energy is

$$E_B = E - mc^2 = mc^2\left(1 + \frac{Z^2\alpha^2}{\left(n - |k| + \sqrt{k^2 - Z^2\alpha^2}\right)^2}\right)^{-1/2}$$

$$- mc^2 \qquad (7.147)$$

For $n = 1$, $|k| = 1$ (lowest state), we get

$$E_B = mc^2\left(1 + \frac{Z^2\alpha^2}{1 - Z^2\alpha^2}\right)^{-1/2} - mc^2$$

$$= mc^2\left(\sqrt{1 - Z^2\alpha^2} - 1\right). \qquad (7.148)$$

So for $Z\alpha = 1$ ($Z \simeq 137$) we have $E = 0$ or $E_B = -mc^2$. From then on, the bound states drop into the negative energy domain and E_B (for large Z) appears to become imaginary.

This result is interpreted as the instability of the vacuum in the vicinity of an extremely strong Coulomb field. Here, one expects that electrons and positrons are generated spontaneously from the vacuum. As this involves particle creation the true processes can only be described within quantum field theory.