Advanced Quantum Physics
Week 11

Michel Ferrero
Thomas Schäfer

Fall 2019

Relativistic quantum mechanics

The aim is of this section is to introduce some notions of relativistic quantum mechanics. We will not enter into fine details as this could be the subject of an entire lecture, but rather discuss some of the implications of relativistic quantum mechanics. Indeed, it is not necessary to embark on a light-speed spaceship to see relativistic physics. A particle is relativistic when its energy becomes large compared to its rest mass mc^2. This may concern electrons such as those used in as high resolution electron microscope or simply photons that are always relativistic! Relativity can then induce new phenomena. A striking example is the production of particle pairs, for example an electron and a positron can be created from a high-energy photon (e.g. γ-rays) interacting with matter. Relativity also naturally explains the existence of the spin, etc. Before discussing the possible forms of a relativistic Hamiltonian, let us (very) briefly recall some notations and conventions from special relativity that we will use in the following.

Special relativity and Lorentz transformations

There are two founding hypotheses behind the construction of special relativity. The first states that the laws that dictate the behavior of physical systems do not change if the systems are described in a given reference frame $\mathcal{R}$ or in any other reference frame that moves in uniform translation with respect to $\mathcal{R}$. The second says that the speed of light is the same in all inertial reference frames.

If an event is described by a time t and coordinates x, y, z in a reference frame $\mathcal{R}$ and a time t' and coordinates x', y', z' in another inertial reference frame $\mathcal{R}'$, then one can show that an alternative way to express both conditions above is simply to require that

$$c^2t^2 - x^2 - y^2 - z^2 = c^2t'^2 - x'^2 - y'^2 - z'^2$$

This condition defines the structure of the group that transforms coordinates from one reference frame to another: the group of *Lorentz transformations*. In the construction of a quantum description of relativistic particles, we will therefore require that the theory is invariant under Lorentz transformations.

Minkowski spacetime

A physical event is described by a time t and coordinates x, y, z in the reference frame. It is useful to group these variables into a spacetime contravariant 4-vector $x = (x^\mu) = (x^0, x^1, x^2, x^3) = (ct, \vec{x})$. Here we use the convention where the first coordinate is ct so that it has the same physical dimension as the other components. Under a Lorentz transformation Λ, 4-vectors change according to

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu$$

Another example of a 4-vector is the energy-momentum tensor $p = (p^\mu) = (E/c, \vec{p})$. The inner product between two spacetime points is

$$x \cdot y = c^2 t_x t_y - \vec{x} \cdot \vec{y} = g_{\mu\nu} x^\mu y^\nu,$$

where we introduced the Minkowski metric through

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad g_\mu{}^\nu g_{\nu\lambda} = \delta_{\mu\lambda}$$

The inner product can be written in a simpler form

$$x \cdot y = x_\mu y^\mu = x^\mu y_\mu,$$

where we introduced the covariant 4-vector x_μ defined by

$$x_\mu = g_{\mu\nu} x^\nu \qquad \Leftrightarrow \qquad x^\mu = g^{\mu\nu} x_\nu$$

Covariant vectors transform in the following way

$$x'_\mu = g_{\mu\beta} \Lambda^\beta{}_\alpha x^\alpha = g_{\mu\beta} \Lambda^\beta{}_\alpha g^{\alpha\nu} x_\nu = \Lambda_\mu{}^\nu x_\nu$$

As discussed above, the Lorentz group consists of linear Lorentz transformations Λ that preserve $x \cdot x$. This yields the defining condition

$$g_{\alpha\beta} x^\alpha y^\beta = g_{\mu\nu} \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta x^\alpha y^\beta \qquad \Leftrightarrow \qquad g_{\alpha\beta} = g_{\mu\nu} \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta$$

Energy-momentum relation

Scalar products yield conserved quantities that are the same in all inertial reference frames. Applying this property to the energy-momentum tensor we find that the following quantity is constant

$$p \cdot p = p_\nu p^\nu = \frac{E^2}{c^2} - \vec{p}^2 = \text{const}$$

In a reference frame where $\vec{p} = 0$, the energy is the rest mass energy $E = mc^2$ and we can therefore write the *energy-momentum relation*

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

This relation will be the basis for the construction of a relativistic equation for quantum mechanics.

The Klein-Gordon equation

The first attempt to construct a relativistic equation for quantum mechanics was the Klein-Gordon equation. The usual Schrödinger equation can be obtained by applying the correspondence principle to the classical Hamiltonian by promoting the canonical momentum $\vec{p}$ to the operator $\hat{\vec{p}} = -i\hbar\vec{\nabla}$ and the energy E to the time derivative $i\hbar\partial_t$. In order to obtain a relativistic version of the Schrödinger equation, a possibility would be to start from the energy-momentum invariant $p^2 = (E/c)^2 - \vec{p}^2 = (mc)^2$ and write

$$E(p) = +(m^2c^4 + \vec{p}^2c^2)^{1/2}$$

Applying the correspondence principle on this equation yields

$$i\hbar\partial_t\Psi = \left[m^2c^4 - \hbar^2c^2\nabla^2\right]^{1/2}\Psi,$$

where m is the rest mass of the particle. There are some issues with such a formulation. It is not easy to make sense of the square root of an operator. One could think of it as a Taylor expansion

$$i\hbar\partial_t\Psi = mc^2\Psi - \frac{\hbar^2\nabla^2}{2m}\Psi - \frac{\hbar^4(\nabla^2)^2}{8m^3c^2}\Psi + \cdots$$

But in this case, one needs an infinite number of boundary conditions to be able to solve the time evolution of Ψ. Another approach that avoids the square root is to directly apply the quantization on

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

and obtain

$$-\hbar^2\partial_t^2\Psi = \left(-\hbar^2c^2\nabla^2 + m^2c^4\right)\Psi$$

This is the *Klein-Gordon equation* that can be written as

$$\left(\partial^2 + k_c^2\right)\Psi = 0,$$

where $\partial^2 = \partial_\mu\partial^\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$ is the d'Alembert operator and $k_c = 2\pi/\lambda_c = mc/\hbar$. By keeping a second-order time derivative, the Klein-Gordon equation still defines a local theory. It is also covariant. The equation has some limitations nevertheless. It is a rotationally-invariant equation and can therefore only describe spin-0 particles. But more importantly, the quantity $|\Psi|^2$ cannot be interpreted as a probability amplitude as in the Schrödinger equation. Indeed, from the latter we can write $\Psi^\dagger(i\hbar\partial_t + \frac{\hbar^2\nabla^2}{2m})\Psi = 0$ and with the complex conjugate of this equation we find

$$\partial_t|\Psi|^2 - i\frac{\hbar}{2m}\nabla\cdot(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*) = 0$$

This continuity equation is expressing the conservation of the probability current $\partial_t\rho + \nabla\cdot\vec{j} = 0$, with $\rho = |\Psi|^2$ and $\vec{j} = -i\frac{\hbar}{2m}(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*)$. This shows that one can consistently interpret $|\Psi|^2$ as a probability density. If we apply the same construction to the Klein-Gordon equation, we obtain

$$\hbar^2\partial_t(\Psi^*\partial_t\Psi - \Psi\partial_t\Psi^*) - \hbar^2c^2\nabla\cdot(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*) = 0.$$

This is again a continuity equation but for

$$\rho = i\frac{\hbar}{2mc^2}(\Psi^*\partial_t\Psi - \Psi\partial_t\Psi^*) \quad \vec{j} = -i\frac{\hbar}{2m}(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*).$$

Introducing the 4-current $j^\mu = (\rho c, \vec{j})$ the continuity equation has the covariant expression $\partial_\mu j^\mu = 0$. We see that the quantity that has to be associated to a probability density in the Klein-Gordon equation is $i\frac{\hbar}{2mc^2}(\Psi^*\partial_t\Psi - \Psi\partial_t\Psi^*)$ and not $|\Psi|^2$. There are therefore several reasons why the Klein-Gordon equation may not be the best candidate for the description of a relativistic quantum theory

- The probability density is not a positive-definite quantity as in the Schrödinger equation. This makes the interpretation of the wavefunction quite puzzling.
- The Klein-Gordon equation is second-order in time, meaning that one needs to know two boundary conditions at $t = 0$, one for Ψ and another one for $\partial_t\Psi$. This is not necessary with the Schrödinger equation.
- The equation that forms the basis of the Klein-Gordon equation $E^2 = m^2c^4 + \vec{p}^2c^2$ has both positive and negative energy solutions. It can actually be shown that the existence of these negative-energy solutions is at the origin of the previous two problems.

One could argue that it is sufficient to ignore the negative energy solutions, but the introduction of interactions makes this difficult, because they typically induce transitions between positive and negative energy states. These difficulties motivated Dirac to look for a different equation to describe relativistic quantum theory.

The Dirac equation

Dirac was convinced that the equation governing relativistic quantum particles had to be first-order in time and space. On the other hand, just like the Klein-Gordon equation, it should yield $p^2 = (mc)^2$ for free particles. His idea was to try to consider the square root of the negative d'Alembert operator. Namely, he was looking for an operator

$$\left(A\partial_x + B\partial_y + C\partial_z + \frac{i}{c}D\partial_t\right)$$

such that its square would be $-\partial^2$, i.e.

$$-\partial^2 = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} = \left(A\partial_x + B\partial_y + C\partial_z + \frac{i}{c}D\partial_t\right)\left(A\partial_x + B\partial_y + C\partial_z + \frac{i}{c}D\partial_t\right)$$

For this equation to be true, all cross terms must vanish and the diagonal terms must be 1. In other words

$$AB + BA = \ldots = 0 \qquad A^2 = B^2 = C^2 = D^2 = 1$$

Dirac, who had been working a lot with the Heisenberg formulation of quantum mechanics, recognized that the operator he was looking for could exist if A, B, C and D could be promoted to *matrices*. Then one could write a first-order matrix equation in time and space

$$\left(A\partial_x + B\partial_y + C\partial_z + \frac{i}{c}D\partial_t\right)\Psi = k_c\Psi$$

Note that the wavefunction Ψ must now be a vector. Applying the operator on both sides again would yield $-\partial^2\Psi = k_c^2\Psi$, meaning that every component of Ψ satisfies the Klein-Gordon equation individually. If we set $A = i\beta\alpha_1$, $B = i\beta\alpha_2$, $C = i\beta\alpha_3$ and $D = \beta$ and recalling that $D^2 = \beta^2 = \mathbb{1}$, the equation above can be written as

$$i\hbar\partial_t\Psi = \left(\beta mc^2 + c\vec{\alpha}\cdot\hat{\vec{p}}\right)\Psi = \hat{\mathcal{H}}\Psi$$

This is a first form of the *Dirac equation* which resembles the Schrödinger equation. It is sometimes referred to as the Dirac-Pauli representation. For $\hat{\mathcal{H}}$ to be Hermitian and for the A, B, C, D matrices to have the right anticommutation relations, we see that

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij}, \qquad \beta^2 = \mathbb{1}, \qquad \{\alpha_i, \beta\} = 0, \qquad \vec{\alpha}^\dagger = \vec{\alpha}, \qquad \beta^\dagger = \beta$$

It can be shown that the matrices α_i and β have to be at least 4×4 to satisfy these requirements. This in turn means that the wavefunction Ψ has 4 components: it is called a *bispinor*.

Covariant form

It is common to see the Dirac equation written in terms of γ matrices, which make its relativistic properties clearer. Introducing γ_i such that $\vec{\alpha} = \gamma^0\vec{\gamma}$ and $\beta = \gamma^0$, we can write the Dirac equation as

$$(i\hbar\gamma^\mu\partial_\mu - mc)\,\Psi = 0$$

The defining properties of the Dirac γ matrices are then

$$\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0, \qquad \{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$$

An explicit standard representation of the γ matrices which easily captures the non-relativistic limit is

$$\gamma^0 = \begin{pmatrix} \mathbb{1}_2 & 0 \\ 0 & -\mathbb{1}_2 \end{pmatrix} \qquad \vec{\gamma} = \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix},$$

where the $\vec{\sigma}$ are the usual Pauli matrices. With these definitions, the $\vec{\alpha}$ and β matrices are

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \qquad \beta = \begin{pmatrix} \mathbb{1}_2 & 0 \\ 0 & -\mathbb{1}_2 \end{pmatrix}$$

Relativistic invariance

Let us check that the Dirac equation is invariant under Lorentz transformations. We need to find how the wavefunction transforms under a transformation Λ or in other words we need to find the matrix $S(\Lambda)$ such that

$$\Psi'(x') = S(\Lambda)\Psi(x),$$

where $\Psi'(x')$ is the solution of the transformed Dirac equation

$$\left(i\hbar\gamma^\mu\partial'_\mu - mc\right)\Psi'(x') = 0$$

The operators ∂_μ transform like a covariant vector $\partial'_\mu = \Lambda_\mu{}^\nu\partial_\nu$ and we can write the Dirac equation above as

$$\left(i\hbar\gamma^\mu\Lambda_\mu{}^\nu\partial_\nu - mc\right)S(\Lambda)\Psi(x) = 0$$

From this equation it is clear that the transformation matrix $S(\Lambda)$ must satisfy

$$\gamma^\mu\Lambda_\mu{}^\nu = S(\Lambda)\gamma^\nu S^{-1}(\Lambda)$$

One can show that for an infinitesimal proper Lorentz transformation $\Lambda^\nu{}_\mu = g^\nu{}_\mu + \omega^\nu{}_\mu$ the matrix $S(\Lambda)$ is

$$S(\Lambda) = \mathbb{1} - \frac{i}{4}\Sigma_{\mu\nu}\omega^{\mu\nu} + \cdots \qquad S^{-1}(\Lambda) = \mathbb{1} + \frac{i}{4}\Sigma_{\mu\nu}\omega^{\mu\nu} + \cdots,$$

where

$$\Sigma_{\alpha\beta} = \frac{i}{2}\left[\gamma_\alpha, \gamma_\beta\right]$$

Probability density and current

In the Dirac equation Ψ is a 4-component complex spinor. So how do we construct the probability density ρ? The Hermitian conjugate of the Dirac equation is

$$\Psi^\dagger\left(-i\hbar\gamma^{\dagger\mu}\overleftarrow{\partial}_\mu - mc\right) = 0,$$

where $\Psi^\dagger \overleftarrow{\partial}_\mu = (\partial_\mu \Psi)^\dagger$. Defining $\bar{\Psi} = \Psi^\dagger \gamma^0$ and using the commutation relations of the γ matrices, we can write

$$\bar{\Psi}\left(i\hbar \overleftarrow{\not{\partial}} + mc\right) = 0,$$

where we used the Feynman slash notation $\not{a} = a_\mu \gamma^\mu$. Subtracting this equation to the usual Dirac equation, we get

$$\bar{\Psi}\left(\overleftarrow{\not{\partial}} + \overrightarrow{\not{\partial}}\right)\Psi = \partial_\mu\left(\bar{\Psi}\gamma^\mu\Psi\right) = 0.$$

This is the continuity equation from which we find the 4-current

$$j^\mu = \bar{\Psi}\gamma^\mu\Psi \qquad \Leftrightarrow \qquad (\rho, \vec{j}) = (\Psi^\dagger\Psi, \Psi^\dagger\vec{\alpha}\Psi)$$

We see that with the Dirac equation, it is possible to consistently define the usual positive-definite probability density $\rho = \Psi^\dagger\Psi$.

Angular momentum and spin

Let us consider the anticlockwise spatial rotation of angle θ about the axis $\hat{n}$. The coordinates transform according to

$$x'_i = [\Lambda x]_i = x_i - \omega_{ij} x_j,$$

where $\omega_{ij} = \epsilon_{ijk} n_k \theta$ for the spatial part of the coordinate and $\Lambda^\mu{}_0 = \Lambda^0{}_\mu = 0$. We conclude that the wavefunction has this expression

$$\Psi(x) = \Psi(\Lambda^{-1}x') = \Psi(x'_0, \vec{x}' + \vec{x}' \times \vec{n}\theta) = (1 - \theta\vec{n}\cdot\vec{x}'\times\vec{\nabla} + \cdots)\Psi(x') = (1 - \frac{i}{\hbar}\theta\,\vec{n}\cdot\hat{\vec{L}} + \cdots)\Psi(x'),$$

where $\hat{\vec{L}} = \hat{\vec{x}} \times \hat{\vec{p}}$ is the usual non-relativistic angular momentum operator. On the other hand, the wavefunction must change according to $\Psi'(x') = S(\Lambda)\Psi(\Lambda^{-1}x')$ under a Lorentz transformation. Above, we have seen that

$$S(\Lambda) = \mathbb{1} - \frac{i}{4}\epsilon_{ijk} n_k \Sigma_{ij}\theta + \cdots$$

Using the explicit representation of the γ matrices and the defining commutator for Σ we have

$$S(\Lambda) = \mathbb{1} - i\vec{n}\cdot\vec{S} + \cdots \qquad \vec{S} = \frac{1}{2}\begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}$$

Combining the equations above, we finally see that the wavefunction has the following transformation property under infinitesimal rotations

$$\Psi'(x') = S(\Lambda)\Psi(\Lambda^{-1}x') = \left(1 - \frac{i}{\hbar}\theta\,\vec{n}\cdot\hat{\vec{J}} + \cdots\right)\Psi(x'),$$

where $\hat{\vec{J}} = \hat{\vec{L}} + \vec{S}$ is the total angular momentum of the particle. It is the sum of an orbital part and an intrinsic component which is the spin. We indeed recognize the property

$$[S_i, S_j] = i\hbar\epsilon_{ijk} S_k \qquad S_i^2 = \frac{\hbar^2}{4}$$

The interpretation of this result is that the Dirac equation describes relativistic particles of spin-1/2. The spin is not a property that has been added artificially as in the non-relativistic Schrödinger equation, but it stems directly from the properties of the Dirac equation.

Free particle solution of the Dirac equation

Let us solve the Dirac equation for free particles. The wavefunction can be written as

$$\Psi(x) = \exp[-ip\cdot x]u(p)$$

This wavefunction is a solution of the Dirac equation with energy $E = cp^0 = \pm\sqrt{\vec{p}^2c^2 + m^2c^4}$ if the spinor $u(p)$ satisfies

$$\left(\not{p} - mc\right)u(p) = 0$$

The Dirac equation also has negative-energy solutions. We will discuss how to give these negative energy states a meaning later. The equation above can be rewritten

$$\left(\not{p} - mc\right)u(p) = \begin{pmatrix} p^0 - mc & -\vec{\sigma}\cdot\vec{p} \\ \vec{\sigma}\cdot\vec{p} & -p^0 - mc \end{pmatrix}\begin{pmatrix} \xi \\ \eta \end{pmatrix} = 0$$

This leads to the following equations

$$(p^0 - mc)\xi = \vec{\sigma}\cdot\vec{p}\eta \qquad \vec{\sigma}\cdot\vec{p}\xi = (p^0 + m)\eta$$

Using $(p^0)^2 = \vec{p}^2c^2 + m^2c^4$ these equations are satisfied if $\eta = \frac{\vec{\sigma}\cdot\vec{p}}{p^0+m}\xi$ and the solutions can be written as

$$u^{(r)}(p) = N(p)\begin{pmatrix} \chi^{(r)} \\ \frac{\vec{\sigma}\cdot\vec{p}}{p^0+mc}\chi^{(r)} \end{pmatrix}$$

The 2-component spinors $\chi^{(r)}$ can be chosen freely as long as they are orthogonal. It is common to choose the $\chi^{(r)}$ to be the two eigenvectors of the helicity operator

$$\frac{\vec{\sigma}}{2}\cdot\frac{\vec{p}}{|\vec{p}|}\chi^{(\pm)} - \pm\frac{1}{2}\chi^{(\pm)}$$

With this choice, the wavefunction can be written

$$\Psi_p^{(\pm)}(x) = N(p)e^{-ip\cdot x}\begin{pmatrix} \chi^{(\pm)} \\ \pm\frac{|\vec{p}|}{p^0+mc}\chi^{(\pm)} \end{pmatrix}$$

For a given choice of p there are therefore 4 solutions to the Dirac equation, two having a positive energy and two having a negative energy.

Low-energy limit of the Dirac equation

To finish our discussion about the Dirac equation, we investigate its low-energy non-relativistic limit in the presence of a magnetic field. As we have discussed earlier, including the magnetic field can be obtained by the minimal substitution

$$p^\mu \to p^\mu - qA^\mu,$$

if we suppose that we are interested in a particle with charge q. In this notation the components of A are the scalar potential is Φ and the vector potential $\vec{A}$, i.e. $A = (\Phi, \vec{A})$. This changes the Dirac equation into

$$(\gamma^\mu(i\hbar\partial_\mu - qA_\mu) - mc)\,\Psi = 0$$

Our strategy to derive the low-energy limit will be based on the following observation. Above, we have found that the wavefunction of a free Dirac particle takes the general form

$$\Psi_p(x) = N(p)e^{-ip\cdot x}\begin{pmatrix}\chi \\ \frac{c\,\vec{\sigma}\cdot\vec{p}}{E+mc^2}\chi\end{pmatrix}$$

At low energies, the second term component of $\Psi_p(x)$ is a factor v/c smaller that the first term. This motivates the idea to construct a perturbation in powers of v/c. Using the Dirac-Pauli representation, the Dirac equation has the form $i\hbar\partial_t\Psi = \hat{\mathcal{H}}\Psi$ with

$$\hat{\mathcal{H}} = \begin{pmatrix} mc^2 + q\Phi & c\vec{\sigma}\cdot(-i\hbar\vec{\nabla} - q\vec{A}) \\ c\vec{\sigma}\cdot(-i\hbar\vec{\nabla} - q\vec{A}) & -mc^2 + q\Phi \end{pmatrix}$$

If we introduce the bispinor $\Psi(x) = (\Psi_a(x), \Psi_b(x))$ the Dirac equation leads to the following set of equations

$$(mc^2 + q\Phi)\Psi_a + c\vec{\sigma}\cdot(\hat{\vec{p}} - q\vec{A})\Psi_b = E\Psi_a$$
$$c\vec{\sigma}\cdot(\hat{\vec{p}} - q\vec{A})\Psi_a - (mc^2 - q\Phi)\Psi_b = E\Psi_b$$

Next, with the parameter $W = E - mc^2$, we can write

$$\Psi_b = \frac{1}{2mc^2 + W - q\Phi}c\vec{\sigma}\cdot(\hat{\vec{p}} - q\vec{A})\Psi_a$$

Order-0: the Pauli equation

At order 0 in v/c, we have

$$\Psi_b \simeq \frac{1}{2mc^2}c\vec{\sigma}\cdot(\hat{\vec{p}} - q\vec{A})\Psi_a$$

Putting this back into the first equation above, we obtain the *Pauli equation*

$$\hat{\mathcal{H}}_{\text{non-rel}}\Psi_a = W\Psi_a \qquad \hat{\mathcal{H}}_{\text{non-rel}} = \frac{1}{2m}\left[\vec{\sigma}\cdot(\hat{\vec{p}} - q\vec{A})\right]^2 + q\Phi$$

Expanding the square and using the Pauli identity $\sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k$ we find the Schrödinger Hamiltonian

$$\hat{\mathcal{H}}_{\text{non-rel}} = \frac{1}{2m}\left(\hat{\vec{p}} - q\vec{A}\right)^2 - \frac{q\hbar}{2m}\vec{\sigma}\cdot\left(\vec{\nabla}\times\vec{A}\right) + q\Phi$$

The second term above is the coupling of the magnetic field with the spin! It is a direct consequence of the relativistic treatment of the particle. The predicted spin magnetic moment is

$$\hat{\vec{\mu}} = \frac{q\hbar}{2m}\vec{\sigma} = \frac{q}{m}\hat{\vec{S}} = \gamma\hat{\vec{S}},$$

where the gyromagnetic ratio is given by $\gamma = q/m = g\,q/2m$ and we see that the corresponding g-factor is $g = 2$. This is very close to the experimentally-measured value $g = 2\times(1.0011596567 \pm 0.0000000035)$.

Order-1: more relativistic corrections

We can pursue our analysis and include the first-order corrections in v/c. For simplicity, we now assume $\vec{A}=0$. We find

$$\Psi_b \simeq \frac{1}{2mc^2}\left(1+\frac{V-W}{2mc^2}\right) c\vec{\sigma}\cdot\hat{\vec{p}}\Psi_a,$$

where we introduced $V=q\Phi$. If we insert this in the first equation we have

$$\left[\frac{1}{2m}(\vec{\sigma}\cdot\hat{\vec{p}})^2+\frac{1}{4m^2c^2}(\vec{\sigma}\cdot\hat{\vec{p}})(V-W)(\vec{\sigma}\cdot\hat{\vec{p}})+V\right]\Psi_a = W\Psi_a$$

Here, we have to be careful about normalization. The original relativistic wavefunction is normalized. We therefore have

$$\begin{aligned}1 &= \int d^3x\Psi^\dagger(\vec{x},t)\Psi(\vec{x},t) = \int d^3x\left(\Psi_a^\dagger(\vec{x},t)\Psi_a(\vec{x},t)+\Psi_b^\dagger(\vec{x},t)\Psi_b(\vec{x},t)\right)\\ &\simeq \int d^3x\Psi_a^\dagger(\vec{x},t)\Psi_a(\vec{x},t)+\frac{1}{(2mc)^2}\int d^3x\Psi_a^\dagger(\vec{x},t)\hat{\vec{p}}^2\Psi_a(\vec{x},t)\end{aligned}$$

The normalized wavefunction is then given by

$$\Psi_s = \left(1+\frac{1}{8m^2c^2}\hat{\vec{p}}^2\right)\Psi_a \qquad \Leftrightarrow \qquad \Psi_a = \left(1-\frac{1}{8m^2c^2}\hat{\vec{p}}^2\right)\Psi_s$$

Previously, we could ignore this modification of the normalization because it does not enter at order 0. Inserting this expression in the equation above, we finally obtain $\hat{\mathcal{H}}_{\mathrm{non-rel}}\Psi_s = W\Psi_s$, with

$$\hat{\mathcal{H}}_{\mathrm{non-rel}} = \frac{\hat{\vec{p}}^2}{2m}-\frac{\hat{\vec{p}}^4}{8m^3c^2}+\frac{\hbar}{4m^2c^2}(\vec{\sigma}\cdot\hat{\vec{p}})V(\vec{\sigma}\cdot\hat{\vec{p}})+V-\frac{1}{8m^2c^2}(V\hat{\vec{p}}^2+\hat{\vec{p}}^2V)$$

We can simplify this expression by using

$$\begin{aligned}[V,\hat{\vec{p}}^2] &= \hbar^2(\nabla^2V)+2i\hbar(\nabla V)\cdot\hat{\vec{p}}\\ (\vec{\sigma}\cdot\hat{\vec{p}})V &= V(\vec{\sigma}\cdot\hat{\vec{p}})+\vec{\sigma}\cdot[v\hat{e}cp,V]\\ (\vec{\sigma}\cdot\hat{\vec{p}})V(\vec{\sigma}\cdot\hat{\vec{p}}) &= V\hat{\vec{p}}^2-i\hbar(\vec{\nabla}V)\cdot\hat{\vec{p}}+\hbar\vec{\sigma}\cdot(\vec{\nabla}V)\times\hat{\vec{p}}\end{aligned}$$

We then find the non-relativistic Schrödinger Hamiltonian

$$\hat{\mathcal{H}}_{\mathrm{non-rel}} = \frac{\hat{\vec{p}}^2}{2m}-\frac{\hat{\vec{p}}^4}{8m^3c^2}+\frac{\hbar}{4m^2c^2}\vec{\sigma}\cdot(\vec{\nabla}V)\times\hat{\vec{p}}+\frac{\hbar^2}{8m^2c^2}(\nabla^2V).$$

In this Hamiltonian, the second term is a relativistic correction to the kinetic energy, the third term is the *spin-orbit interaction* and the fourth term is the *Darwin term*. In a central potential, the spin-orbit term takes the form

$$\hat{\mathcal{H}}_{\mathrm{spin-orbit}} = \frac{\hbar}{4m^2c^2}\vec{\sigma}\cdot\frac{1}{r}(\partial_rV)\vec{r}\times\hat{\vec{p}} = \frac{\hbar}{4m^2c^2}\frac{1}{r}(\partial_rV)\vec{\sigma}\cdot\hat{\vec{L}} = \frac{1}{2m^2c^2}\frac{1}{r}(\partial_rV)\hat{\vec{S}}\cdot\hat{\vec{L}}$$