

Wave mechanics and the beginnings of quantum physics

February 1, 2017

1 The beginnings of quantum physics: an historical overview

In the early years of the twentieth century, a series of experiments showed that the classical notions of particle and wave are both present in matter at very small scales. The insights led to a breakdown of Newtonian mechanics, in favor of the emerging quantum theory.

Here is a brief description of a few of the important results, presented in logical order rather than historical order.

1.1 Planck (1900), blackbody radiation, and $E = \hbar\omega$

The theoretical distribution of frequencies of radiation within a hot cavity based on classical thermodynamics does predicts far too many high frequency states. Planck showed that good agreement could be obtained if one assumed proportionality between the energy and the frequency, $E = hf$ or more conveniently,

$$E = \hbar\omega$$

The requirement of higher energy suppresses the number of high frequency states. This was the first evidence for quantization. The measured value of the constant, called Planck's constant, is $h = 6.62606957 \times 10^{-34} \frac{m^2 kg}{s}$. Recall that the classical prediction is for the energy to vary as the sum of squares of the electric and magnetic fields of the wave, independently of frequency.

The result relates a wave property (frequency) to a particle property (energy).

1.2 Photoelectric effect (1905)

Einstein's Nobel prize winning paper of 1905 argued that if all of the energy, $E = \hbar\omega$, of light of frequency ω were absorbed by an electron in an atom of a metal, and if an energy ϕ were required to free the electron from its atom, then electrons would escape with kinetic energy $KE = \hbar\omega - \phi$. Such a linear relationship with frequency is found. The conclusion, that a *photon* of light must be *entirely* absorbed or not at all, remains one of the more puzzling features of quantum mechanics.

The result points to *quantization*: the idea that in radiation, light moves in discrete all-or-nothing packets called *photons*.

1.3 Electron diffraction through crystals, de Broglie (1924), and $p = \hbar k$

When a beam of electrons of momentum p is directed at a crystal, a circular diffraction pattern appears, just as if the electron traveled as a wave with wavelength $\lambda = \frac{h}{p}$, called the de Broglie wavelength. The pattern is just like that produced by x-rays passing through a crystal. This shows that something thought of as a particle also has wavelike properties.

We may write the relationship using the wave vector,

$$\mathbf{k} = \frac{2\pi}{\lambda} \hat{\mathbf{n}}$$

where $\hat{\mathbf{n}}$ is the direction of propagation of the wave. Then the de Broglie relationship becomes

$$\mathbf{p} = \hbar\mathbf{k}$$

in close parallel to the Planck relation.

2 The Bohr atom (1913) and the Schrödinger equation (1925)

Based in part on these findings, and in an effort to explain the discrete spectra of atoms, Bohr used the wave nature of the electron to develop a simple model for the hydrogen atom.

2.1 The Bohr atom

The Bohr atom assumes the usual electrostatic attraction between an electron and a proton,

$$\mathbf{F} = -\frac{ke^2}{r^2}\hat{\mathbf{r}}$$

Then, for an electron in a circular orbit,

$$\mathbf{a} = -\frac{v^2}{r}\hat{\mathbf{r}}$$

To these classical elements, Bohr added a quantization rule: the angular momentum must be a multiple of Planck's reduced constant,

$$L = mvr = n\hbar$$

Combining the classical elements, we have a relationship between the radius and velocity of circular orbits,

$$\frac{ke^2}{r^2} = \frac{mv^2}{r}$$

Solving for the velocity, we have

$$v = \sqrt{\frac{ke^2}{mr}}$$

Then according to the Bohr quantization rule,

$$\begin{aligned} n\hbar &= mvr \\ &= \sqrt{mrke^2} \end{aligned}$$

or, solving for r ,

$$r_n = \frac{n^2\hbar^2}{mke^2}$$

The total energy of the electron is

$$\begin{aligned} E &= \frac{1}{2}mv^2 - \frac{ke^2}{r} \\ &= -\frac{ke^2}{2r} \\ &= -\frac{mk^2e^4}{2n^2\hbar^2} \\ &= -\frac{13.6eV}{n^2} \end{aligned}$$

This means that the energy of an electron that moves between two orbits will change by

$$\begin{aligned}
 E &= \frac{1}{2}mv^2 - \frac{ke^2}{r} \\
 &= -\frac{ke^2}{2r} \\
 &= -\frac{mk^2e^4}{2n^2\hbar^2} \\
 \Delta E &= -13.6 \left(\frac{1}{n^2} - \frac{1}{m^2} \right) eV
 \end{aligned}$$

If this energy is given off in the form of a photon satisfying the Planck relation, then the frequency of the emitted light will be

$$\omega = \frac{\Delta E}{\hbar}$$

A formula of this form had already been determined experimentally, and was now explained by the Bohr model.

2.2 The Klein-Gordon equation

The Bohr model restricts the electron to circular motion in a plane, and gives incorrect values of total angular momentum for the electrons. A fuller picture was required, and is provided by writing a 3-dimensional wave equation for the electron.

We may use the deBroglie wavelength and the Planck relation, together with the relativistic relationship between energy and momentum, to derive a suitable equation. We have:

$$\begin{aligned}
 E &= \hbar\omega \\
 \mathbf{p} &= \hbar\mathbf{k}
 \end{aligned}$$

The 4-momentum of a particle is given by

$$\begin{aligned}
 p^\alpha &= mu^\alpha \\
 &= m\gamma(c, \mathbf{v}) \\
 &= \left(\frac{E}{c}, \mathbf{p} \right)
 \end{aligned}$$

and the norm of this equation is

$$\begin{aligned}
 \eta_{\alpha\beta}p^\alpha p^\beta &= p^\alpha p_\alpha \\
 &= -(p^0)^2 + \mathbf{p}^2 \\
 &= -\frac{E^2}{c^2} + \mathbf{p}^2
 \end{aligned}$$

On the other hand, we have

$$\begin{aligned}
 \eta_{\alpha\beta}p^\alpha p^\beta &= m^2 u^\alpha u_\alpha \\
 &= -m^2 c^2
 \end{aligned}$$

Equating these,

$$\begin{aligned}
 -\frac{E^2}{c^2} + \mathbf{p}^2 &= -m^2 c^2 \\
 E^2 &= \mathbf{p}^2 c^2 + m^2 c^4
 \end{aligned}$$

Now suppose the electron is described by a plane wave, in which we replace $(\frac{\omega}{c}, \mathbf{k})$ using the Planck and deBroglie relations

$$\begin{aligned}\psi &= Ae^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \\ &= Ae^{\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{x}-Et)}\end{aligned}$$

Then we may recover the energy and momentum by differentiation,

$$\begin{aligned}-i\hbar\nabla\psi &= \nabla\left[Ae^{\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{x}-Et)}\right] \\ &= \mathbf{p}\psi \\ i\hbar\frac{\partial}{\partial t}\psi &= \frac{\partial}{\partial t}\left[Ae^{\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{x}-Et)}\right] \\ &= E\psi\end{aligned}$$

Identifying the energy and momentum *operators*,

$$\begin{aligned}p_\alpha &= \left(-\frac{E}{c}, \mathbf{p}\right) = -i\hbar\left(\frac{1}{c}\frac{\partial}{\partial t}, \nabla\right) \\ &= -i\hbar\frac{\partial}{\partial x^\alpha}\end{aligned}\tag{1}$$

and substituting into the energy-momentum relation,

$$\begin{aligned}E^2 &= \mathbf{p}^2c^2 + m^2c^4 \\ \left(i\hbar\frac{\partial}{\partial t}\right)^2 &= (i\hbar\nabla)^2c^2 + m^2c^4\end{aligned}$$

gives a differential operator. Allowing this operator to act on a “wave function”, ψ ,

$$\begin{aligned}-\hbar^2\frac{\partial^2\psi}{\partial t^2} &= -\hbar^2c^2\nabla^2\psi + m^2c^4\psi \\ -\frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} + \nabla^2\psi &= \frac{m^2c^2}{\hbar^2}\psi\end{aligned}\tag{2}$$

The differential operator

$$\begin{aligned}\square &\equiv -\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \nabla^2 \\ &= \eta^{\alpha\beta}\frac{\partial}{\partial x^\alpha}\frac{\partial}{\partial x^\beta}\end{aligned}$$

is the spacetime generalization of the Laplacian, $\nabla^2 = \delta^{ij}\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}$. Notice that the wave speed, c , must be the speed of light for $\eta^{\alpha\beta}\frac{\partial}{\partial x^\alpha}\frac{\partial}{\partial x^\beta}$ to be a Lorentz invariant operator.

The time dependence makes eq.(2) a wave operator, but because of the Planck and deBroglie relationships, it also depends on particle-like energy and momentum. Indeed, the plane-wave solutions may be written as

$$\psi = Ae^{\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{x}-Et)}$$

Eq.(2),

$$\square\psi = \frac{m^2c^2}{\hbar^2}\psi$$

is called the *Klein-Gordon equation*. It first appears in Schrödinger’s notes in 1925 before being published the next year first by Oskar Klein and Walter Gordon, but also the same year by Vladimir Fock, Johann Kudar, Théophile de Donder and Frans-H. van den Dungen, and Louis de Broglie. It is the obvious relativistic generalization of the Schrödinger equation but fails to describe electron spin.

2.3 A first order equation: the Schrödinger equation

Because the Klein-Gordon equation is second order in time derivatives, it requires both initial position and velocity specifications, and this is forbidden by the uncertainty principle. Additionally, it is found that the equation leads to negative probability states.

In 1925, Schrödinger took a different approach. The problems arising from the second order time derivatives may be avoided by first solving for the energy, then taking a non-relativistic approximation, and finally substituting the operator forms, eq.(1), for E and \mathbf{p} .

With the total energy of our electron including a potential, we may write

$$\begin{aligned} E &= \sqrt{\mathbf{p}^2 c^2 + m^2 c^4} + V \\ &= mc^2 \sqrt{1 + \frac{\mathbf{p}^2}{m^2 c^2}} + V \end{aligned}$$

For $v \ll c$ we may expand $\sqrt{1 + \frac{\mathbf{p}^2}{m^2 c^2}}$ in a Taylor series,

$$\begin{aligned} \sqrt{1 + \frac{\mathbf{p}^2}{m^2 c^2}} &= 1 + \frac{\mathbf{p}^2}{2m^2 c^2} - \frac{1}{8} \left(\frac{\mathbf{p}^2}{m^2 c^2} \right)^2 + \dots \\ &\approx 1 + \frac{\mathbf{p}^2}{2m^2 c^2} \end{aligned}$$

This lets us find the usual Newtonian expression, together with the rest energy:

$$E \approx mc^2 \left(1 + \frac{\mathbf{p}^2}{2m^2 c^2} \right) + V$$

Making the same operator substitutions that led us to the Klein-Gordon equation, and allowing it to operate on a function, ϕ , gives

$$i\hbar \frac{\partial \phi}{\partial t} = mc^2 \phi - \frac{\hbar^2}{2m} \nabla^2 \phi + V \phi$$

The constant mass term may be removed by the replacement

$$\phi = \psi e^{-\frac{i}{\hbar} mc^2 t}$$

Then we find

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \left(\psi e^{-\frac{i}{\hbar} mc^2 t} \right) &= mc^2 \psi e^{-\frac{i}{\hbar} mc^2 t} - \frac{\hbar^2}{2m} \nabla^2 \left(\psi e^{-\frac{i}{\hbar} mc^2 t} \right) + V \psi e^{-\frac{i}{\hbar} mc^2 t} \\ i\hbar \left(\frac{\partial \psi}{\partial t} - \frac{i}{\hbar} mc^2 \psi \right) e^{-\frac{i}{\hbar} mc^2 t} &= \left(mc^2 \psi - \frac{\hbar^2}{2m} \nabla^2 \psi + V \psi \right) e^{-\frac{i}{\hbar} mc^2 t} \end{aligned}$$

resulting in the familiar form of the *time-dependent Schrödinger equation*,

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V \psi \quad (3)$$

Notice that it is possible to generate a relativistic correction to the Schrödinger equation by keeping the next order term in the Taylor series. The result is

$$\begin{aligned} i\hbar \frac{\partial \psi}{\partial t} &= -\frac{\hbar^2}{2m} \nabla^2 \psi - \frac{1}{8} mc^2 \left(\frac{\mathbf{p}^2}{m^2 c^2} \right)^2 \psi + V \psi \\ i\hbar \frac{\partial \psi}{\partial t} &= -\frac{\hbar^2}{2m} \nabla^2 \psi - \frac{1}{8} \frac{(-\hbar^2 \nabla^2)^2}{m^3 c^2} \psi + V \psi \end{aligned}$$

and therefore

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi - \frac{1}{8} \frac{\hbar^4}{m^3 c^2} (\nabla^2)^2 \psi + V \psi$$

This relativistic correction contributes to the fine structure of hydrogenic spectra, along with *spin-orbit coupling* and the *Darwin term*.

3 Spin

Stern and Gerlach performed a series of experiments through the 1920s that together show an unexpected feature first interpreted as the intrinsic spin of the electron by Uhlenbeck and Goudsmit.

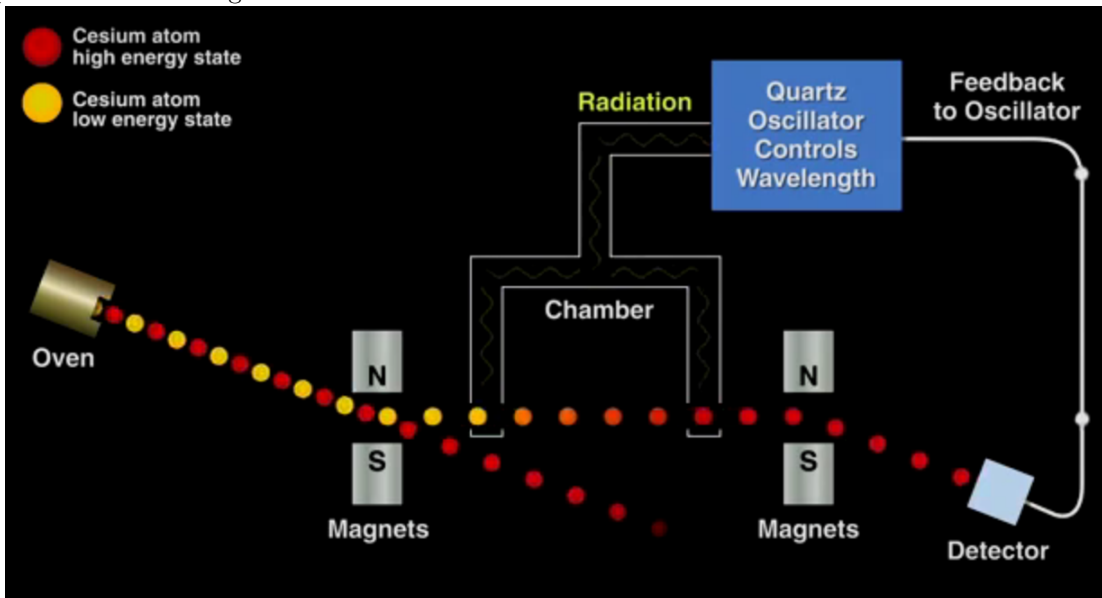
3.1 The Stern-Gerlach experiment

In an attempt to measure the angular momentum predicted for atomic systems, Stern and Gerlach passed silver atoms through an inhomogeneous magnetic field. The field gave a force on each atom proportional to its angular momentum. The orbital angular momentum of the electrons in silver depends only on the state of the outer electron. With an orbital total angular momentum of $L = l\hbar$ there should be $2l + 1$ distinct values of angular momentum, and the beam of silver atoms should split into $2l + 1$ separate beams. However, by 1925, some experiments showed an even number - two separate beams - which cannot be explained by integers l . The result was explained by Goudsmit and Uhlenbeck, by assigning the electron an *intrinsic spin* of $\frac{1}{2}\hbar$.

Some of the history of this period is wonderfully described by Goudsmit (see Notes or the homework page).

3.1.1 Cesium clock

A pair of Stern-Gerlach devices (using cesium instead of silver) is the basis of the atomic clock. An atomic beam is split into two, dependent on the outer electron's spin orientation. The "spin down" beam is irradiated at a frequency ω with energy $E = \hbar\omega$ exactly sufficient to flip it to the "spin up" state. This beam then passes through a second Stern-Gerlach magnet and the "spin up" component measured. With a feedback loop, the frequency ω is tuned to maximize this spin up component. The entire device highly stabilizes the frequency ω to about one part in 10^{14} . Counting the waves gives accurate timing, with the second *defined* as the duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium 133 atom.



3.2 The Pauli equation

Although the Schrödinger equation fails to describe the electron spin, Pauli generalized it in 1927. The first insight is to write the total classical energy of a point charge in an electromagnetic field. If the electric and

magnetic fields are found from scalar and vector potentials, $A_\alpha = (\phi, \mathbf{A})$, then this energy is given by the Hamiltonian,

$$H = \frac{1}{2m} (\mathbf{p} - e\mathbf{A})^2 + qV$$

Replacing the energy with the operator form of this expression gives,

$$\left(\frac{1}{2m} (-i\hbar\nabla - e\mathbf{A})^2 + qV \right) \psi = i\hbar \frac{\partial\psi}{\partial t}$$

but this still does not include the coupling of the intrinsic spin of the electron to the magnetic field.

Pauli introduced two component *spinors* to correspond to the two quantum spin states of the electron,

$$\Psi = \begin{pmatrix} \psi_1(\mathbf{x}, t) \\ \psi_2(\mathbf{x}, t) \end{pmatrix}$$

then modified the equation to include two-dimensional operators. Since these operators must be Hermitian in order to give real values for the energy, they must be linear combinations of the identity and the Pauli matrices,

$$\begin{aligned} \boldsymbol{\sigma} &= (\sigma_x, \sigma_y, \sigma_z) \\ &= \left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right) \end{aligned}$$

We will explore these in much more detail in later Sections. Pauli took the relevant linear combination to be

$$\boldsymbol{\sigma} \cdot (i\hbar\nabla - e\mathbf{A})$$

along with the 2×2 identity for the scalar potential, $e\phi\mathbf{1}$. Replacing $\mathbf{p} - e\mathbf{A}$ in the Hamiltonian then gives

$$i\hbar \frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m} [\boldsymbol{\sigma} \cdot (i\hbar\nabla - e\mathbf{A})]^2 \Psi + e\phi\mathbf{1}\Psi$$

The resulting Pauli equation applies to a 2-component *spinor* and, when the potential for the electromagnetic field is included using the Pauli matrices, allows for the correct description of non-relativistic spin, including the Stern-Gerlach results. If we let

$$\Psi = \begin{pmatrix} \psi_1(\mathbf{x}, t) \\ \psi_2(\mathbf{x}, t) \end{pmatrix}$$

and (ϕ, \mathbf{A}) be the scalar and vector potentials of electrodynamics, then the Pauli equation is for a spin- $\frac{1}{2}$ particle with charge e . Here, the Pauli matrices are given by

$$\begin{aligned} \boldsymbol{\sigma} &= (\sigma_x, \sigma_y, \sigma_z) \\ &= \left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right) \end{aligned}$$

Exercise: Show that the Pauli matrices satisfy

$$\sigma_i \sigma_j = \delta_{ij} \mathbf{1} + i \varepsilon_{ijk} \sigma_k$$

where $i, j, k = 1, 2, 3$, δ_{ij} is the Kronecker delta, ε_{ijk} is the Levi-Civita tensor defined by

$$\varepsilon_{ijk} = \begin{cases} +1 & ijk \text{ an even permutation of } 123 \\ -1 & ijk \text{ an odd permutation of } 123 \\ 0 & \text{otherwise} \end{cases}$$

and there is an implicit sum on k (Einstein summation convention). The Levi-Civita is antisymmetric under interchange of any pair of indices, $\varepsilon_{ijk} = -\varepsilon_{ikj}$ and so on.

Exercise: Let S_{ij} be any symmetric matrix, $S_{ij} = S_{ji}$, and A_{ij} be any antisymmetric matrix, $A_{ij} = -A_{ji}$. Show that the trace of their product vanishes,

$$\sum_{i,j} A_{ij} S_{ji} = A_{ij} S_{ji} = 0$$

It follows that

$$\varepsilon_{ijk} S_{jk} = 0$$

for any symmetric matrix.

Exercise: Show that the i^{th} component of the cross product of two vectors, A_i, B_j may be written as

$$[\mathbf{A} \times \mathbf{B}]_i = \varepsilon_{ijk} A_j B_k$$

and the curl of a vector field as

$$[\nabla \times \mathbf{A}]_i = \varepsilon_{ijk} \nabla_j B_k$$

Using this expression for the product of Pauli matrices, and the quantity $[\boldsymbol{\sigma} \cdot (i\hbar\nabla - e\mathbf{A})]^2$ works out to

$$[\boldsymbol{\sigma} \cdot (i\hbar\nabla - e\mathbf{A})]^2 \psi$$

$$\begin{aligned} [\boldsymbol{\sigma} \cdot (i\hbar\nabla - e\mathbf{A})]^2 \Psi &= \left(i\hbar\sigma_i \frac{\partial}{\partial x^i} - e\sigma_i A_i \right)^2 \Psi \\ &= \left(i\hbar\sigma_i \frac{\partial}{\partial x^i} - e\sigma_i A_i \right) \left(i\hbar\sigma_j \frac{\partial}{\partial x^j} - e\sigma_j A_j \right) \Psi \\ &= \sigma_i \sigma_j \left(i\hbar \frac{\partial}{\partial x^i} - eA_i \right) \left(i\hbar \frac{\partial}{\partial x^j} - eA_j \right) \Psi \\ &= (\delta_{ij} + i\varepsilon_{ijk} \sigma_k) \left(i\hbar \frac{\partial}{\partial x^i} - eA_i \right) \left(i\hbar \frac{\partial \Psi}{\partial x^j} - eA_j \Psi \right) \\ &= (\delta_{ij} + i\varepsilon_{ijk} \sigma_k) \left(-\hbar^2 \frac{\partial^2 \Psi}{\partial x^i \partial x^j} - i\hbar e \left(A_i \frac{\partial \Psi}{\partial x^j} + A_j \frac{\partial \Psi}{\partial x^i} \right) + e^2 A_i A_j \Psi - e \left(i\hbar \frac{\partial}{\partial x^i} A_j \right) \Psi \right) \\ &= \delta_{ij} \left(-\hbar^2 \frac{\partial^2 \Psi}{\partial x^i \partial x^j} - i\hbar e \left(A_i \frac{\partial \Psi}{\partial x^j} + A_j \frac{\partial \Psi}{\partial x^i} \right) + e^2 A_i A_j \Psi - e \left(i\hbar \frac{\partial}{\partial x^i} A_j \right) \Psi \right) + i\varepsilon_{ijk} \sigma_k \left(-\hbar^2 \frac{\partial^2 \Psi}{\partial x^i \partial x^j} \right) \\ &= (-\hbar^2 \nabla^2 \Psi - 2i\hbar e \mathbf{A} \cdot \nabla \Psi + e^2 \mathbf{A}^2 \Psi - e (i\hbar \nabla \cdot \mathbf{A}) \Psi) + \hbar e \boldsymbol{\sigma} \cdot (\nabla \times \mathbf{A}) \Psi \\ &= (i\hbar \nabla - e\mathbf{A})^2 \Psi + \hbar e \boldsymbol{\sigma} \cdot \mathbf{B} \Psi \end{aligned}$$

This gives the Schrödinger-Pauli equation,

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} (i\hbar \nabla - e\mathbf{A})^2 \Psi - \frac{\hbar e}{2m} \boldsymbol{\sigma} \cdot \mathbf{B} \Psi + e\varphi \Psi$$