

1 Introduction to quantum mechanics

The theory of quantum mechanics explores the behavior of matter and energy at atomic scales. At these scales, the classical rules of physics we are familiar with no longer apply. This introduction is designed for graduate students in mathematics with little to no background in physics. We will use the famous double-slit experiment as a gateway to uncover key ideas in quantum mechanics, culminating in the electronic Schrödinger equation, a fundamental tool for studying particles at the atomic level. We will come to the conclusion that electrons behave like both particles and waves which is known as the particle-wave duality. As most readers will be – at least at the intuitive level – familiar with classical mechanics, we will begin our exposition with classical physics and then transition into the quantum theory of particles.

Point particles vs classical waves

In classical physics, waves and point particles are two very different models for physical systems, each with an exceptionally large range of applications.

Point particles which are idealized particles with no spatial extension are very commonly used in physics to describe the behavior of (uniform) rigid body dynamics, such as the “projectile motion” or “billiard mechanics”. The notion of (point) particles is governed by Newtonian mechanics going back to the 17th century, which is manifested in three Newtonian Law’s of motion

1. A body remains at rest, or in motion at a constant speed in a straight line, except insofar as it is acted upon by a force.
2. At any instant of time, the net force on a body is equal to the rate at which the body’s momentum is changing with time.

$$F = \frac{dp}{dt} = m \frac{dv}{dt} = ma \quad (1)$$

3. If two bodies exert forces on each other, these forces have the same magnitude but opposite directions.

The describing quantities of (point) particles are (1) their mass, (2) their position, and (3) their forces.

Classical waves are propagating dynamic disturbances, i.e., changes from an equilibrium state. Periodic waves oscillate repeatedly about an equilibrium (resting) value at some frequency. When the entire waveform moves in one direction, it is said to be a traveling wave; by contrast, a pair of superimposed periodic waves traveling in opposite directions makes a standing wave. In a standing wave, the amplitude of vibration has nulls at some positions where the wave amplitude appears smaller or even zero. Commonly, a wave is described by a scalar field

$$u : \mathbb{R}^3 \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R} ; (\mathbf{r}, t) \mapsto u(\mathbf{r}, t) \quad (2)$$

assigning at any point in time $t \in \mathbb{R}$ and any point in space $\mathbf{r} \in \mathbb{R}^3$ a magnitude $u(\mathbf{r}, t)$. Classical waves are governed by the wave equation, i.e.,

$$\frac{\partial^2 u}{\partial t^2} = c^2 \Delta_{\mathbf{r}} u, \quad (3)$$

where $c \in \mathbb{R}_{>0}$. Unlike point particles, waves have continuous values at many points in space that vary with time. The spatial extent of waves can vary with time due to diffraction or wave interferences, see Fig. 1. These are characteristic observations of waves.

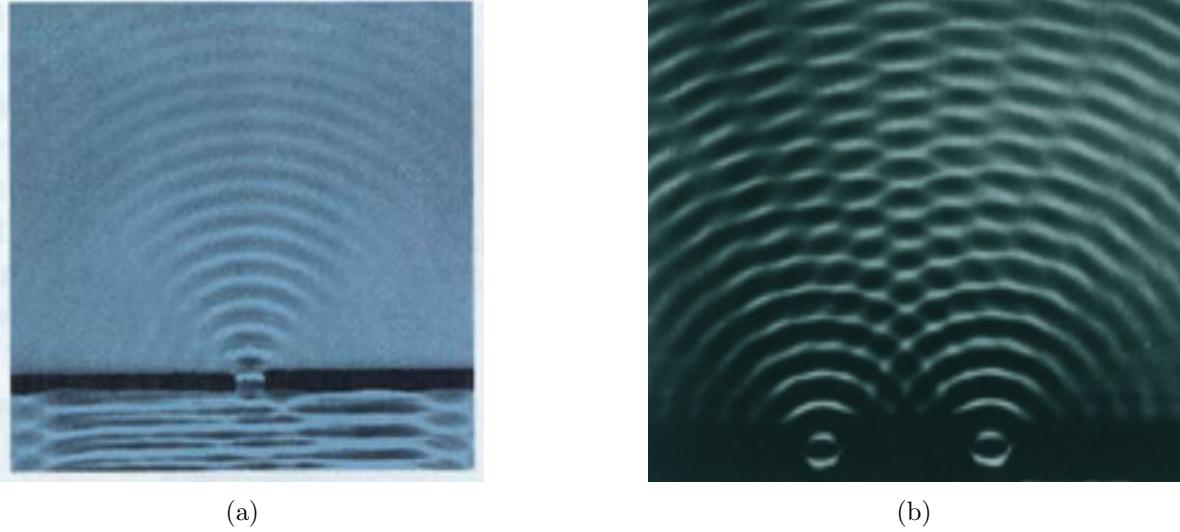


Figure 1: (a) Left panel shows diffraction of water wave on a single slit. (b) Right panel shows interference pattern of concentric waves.

Unlike waves, point particles do not exhibit diffraction or interference.

The Double-Slit Experiment

We will transition to atomic scales; imagine an apparatus where particles at the atomic scale, such as electrons, are fired onto a detector, see Fig. 2

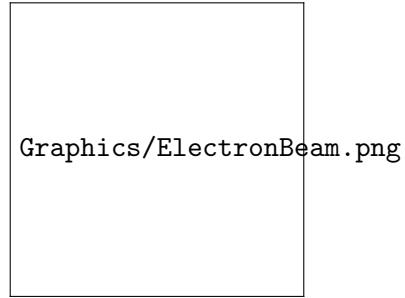


Figure 2: electron beam

The particles can be individually detected and appear to behave like point particles in classical physics. Now, imagine a setup where the particles are sent toward a barrier with two small openings, called slits, see 3. Beyond the slits is a screen that detects where the particles land.

We now consider two scenarios:

1. **Blocking one slit:** If we close one slit, the particles form a simple pattern on the screen directly behind the open slit, see ???. This result again aligns with our expectation from the point particle picture.

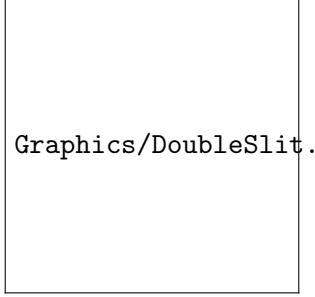


Figure 3: electron beam

2. **Opening both slits:** With both slits open, we expect the result to be the sum of the patterns from each slit. However, what we observe is far more surprising; a series of bright and dark bands, called an interference pattern, similar to what we see if waves (not particles) were passing through the slits, see 3.

This experiment shows that particles such as electrons behave like waves under certain conditions. More surprisingly, if we place detectors at the slits to observe which slit the individual electron passes through, see ??, the interference pattern disappears, and the electrons behave like particles again. This puzzling behavior is at the heart of quantum mechanics.

But, how do we interpret this result and what does it mean for our understanding of physics? The double-slit experiment suggests that particles have a dual nature: they can behave like both, particles and waves. This is called the *wave-particle duality* challenging our classical picture of physics in which these concepts are fundamentally different. The quantum mechanical framework allows us to describe the interference pattern arising in the dual slit experiment. Quantum mechanically, particles are characterized by a *wave function*, which determines the probability of finding the particle in a particular location. As the name suggests, the propagation of the particle's wave function is governed by “some type of wave equation”. In particular, when no measurement is made, the particle's wave function spreads out like a wave producing the wave-like interference patterns observed in the double-slit experiment. However, when we measure which slit the particle goes through, the wave function “collapses” into a specific location, and the interference pattern vanishes – the electron now behaves like a classical particle.

Mathematically, the wave function is a function in space and time, denoted by

$$\psi : \mathbb{R}^m \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{C} ; (\mathbf{x}, t) \mapsto \psi(\mathbf{x}, t) \quad (4)$$

where $m \in \mathbb{N}$. The wave function is the central object in quantum mechanics that encapsulates all the information about the system's state. Its squared modulus is the probability density of finding the particles at positions $\mathbf{x} \in \mathbb{R}^m$ at time $t \in \mathbb{R}_{\geq 0}$, hence, $\psi \in L^2(\mathbb{R}^m)$ for all $t \in \mathbb{R}_{\geq 0}$. The governing equation for the wave function is the Schrödinger equation, proposed by Erwin Schrödinger in 1925 [?]. In its time-dependent form, the Schrödinger equation reads

$$i \frac{\partial \psi(\mathbf{x}, t)}{\partial t} = H\psi(\mathbf{x}, t), \quad (5)$$

where i is the imaginary unit and H is the Hamiltonian operator, representing the total energy

of the system of interest. For a particle moving in a potential $V(\mathbf{x})$, the Hamiltonian is

$$H = -\frac{1}{2}\nabla^2 + V(\mathbf{x}), \quad (6)$$

where $-\frac{1}{2}\nabla^2$ describes the kinetic energy, and $V(\mathbf{x})$ describes the potential energy. In many cases, however, we are interested in systems that do not change in time, such as electrons in a stable atom or molecule. Here, the Schrödinger equation simplifies to

$$H\psi(x) = E\psi(x), \quad (7)$$

where E describes the system's energy. Narrowing down on the application to atoms and molecules, we can specify the potential further. Consider a molecule with N_{nuc} nuclei and N electrons, the wave function depends on both nuclear positions $\mathbf{R} = (\mathbf{R}_1, \dots, \mathbf{R}_{N_{\text{nuc}}}) \in \mathbb{R}^{3N_{\text{nuc}}}$ with $\mathbf{R}_i \in \mathbb{R}^3$ and the electronic positions $\mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_N) \in \mathbb{R}^{3N}$ with $\mathbf{r}_i \in \mathbb{R}^3$

$$H\Psi(\mathbf{r}, \mathbf{R}) = E\Psi(\mathbf{r}, \mathbf{R}) \quad (8)$$

where

$$\begin{aligned} H = & - \sum_{i \in [N]} \frac{1}{2} \Delta_{\mathbf{r}_i} - \sum_{i \in [N]} \sum_{j \in [N_{\text{nuc}}]} \frac{Z_j}{\|\mathbf{r}_i - \mathbf{R}_j\|} + \frac{1}{2} \sum_{i \in [N]} \sum_{i \neq j \in [N]} \frac{1}{\|\mathbf{r}_i - \mathbf{r}_j\|} \\ & - \sum_{i \in [N_{\text{nuc}}]} \frac{1}{2M_i} \Delta_{\mathbf{R}_i} + \frac{1}{2} \sum_{i \in [N_{\text{nuc}}]} \sum_{i \neq j \in [N_{\text{nuc}}]} \frac{Z_i Z_j}{\|\mathbf{R}_i - \mathbf{R}_j\|} \end{aligned} \quad (9)$$

where M_i is the nuclear mass relative to the electronic mass, and Z_i are the atomic numbers found on the periodic table. A common simplification is the Born-Oppenheimer approximation, which assumes that nuclei move much more slowly than electrons, meaning that the kinetic potential energy coming exclusively from nuclei simply enter as a constant. This reduces the problem to the electronic Schrödinger equation:

$$H(\mathbf{R})\Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \quad (10)$$

where

$$H(\mathbf{R}) = - \sum_{i \in [N]} \frac{1}{2} \Delta_{\mathbf{r}_i} - \sum_{i \in [N]} \sum_{j \in [N_{\text{nuc}}]} \frac{Z_j}{\|\mathbf{r}_i - \mathbf{R}_j\|} + \frac{1}{2} \sum_{i \in [N]} \sum_{i \neq j \in [N]} \frac{1}{\|\mathbf{r}_i - \mathbf{r}_j\|}. \quad (11)$$

Subject of this topic course will be to solve this innocent looking eigenvalue problem.

2 The hydrogen atom

For the hydrogen atom, the (stationary) Schrödinger equation takes the simple form

$$\left(-\frac{1}{2}\Delta - \frac{1}{\|\mathbf{R} - \mathbf{r}\|}\right)\Psi(\mathbf{r}) = E\Psi(\mathbf{r}), \quad (12)$$

where $\mathbf{R} \in \mathbb{R}^3$ denotes the position of the hydrogen nucleus. Without loss of generality, we can define our coordinate system such that the nucleus defines the coordinate origin simplifying the equation to

$$\left(-\frac{1}{2}\Delta - \frac{1}{\|\mathbf{r}\|}\right)\Psi(\mathbf{r}) = E\Psi(\mathbf{r}). \quad (13)$$

We note that this is a spherically symmetric problem. This can be seen as follows: Consider the rotation matrix R , then

$$\mathbf{r} \mapsto R\mathbf{r} = \mathbf{r}' \quad (14)$$

describes a rotation of the frame. Since R is unitary and the 2-norm is unitarily invariant, we have

$$\|\mathbf{r}'\| = \|R\mathbf{r}\| = \|\mathbf{r}\|. \quad (15)$$

Moreover, in the rotated frame the partial derivative is given by

$$\partial_{x'_i} = \sum_j R_{i,j} \partial_{x_j} \quad \Rightarrow \quad \partial_{x'_i}^2 = \sum_{k,j} R_{i,k} R_{j,k} \partial_{x_j} \partial_{x_k}. \quad (16)$$

Then

$$\Delta' f = \sum_i \sum_{k,j} R_{i,k} R_{j,k} \partial_{x_j} \partial_{x_k} f = \sum_{k,j} \left(\sum_i R_{i,k} R_{j,k} \right) \partial_{x_j} \partial_{x_k} f = \sum_k \partial_{x_k}^2 f = \Delta f. \quad (17)$$

This suggests that we investigate this problem in spherical coordinates. Recall the spherical coordinate transformation

$$s : [0, \pi] \times [0, 2\pi] \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^3 ; \quad \begin{bmatrix} \theta \\ \phi \\ r \end{bmatrix} \mapsto \begin{bmatrix} r \sin(\theta) \cos(\phi) \\ r \sin(\theta) \sin(\phi) \\ r \cos(\theta) \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (18)$$

and

$$s^{-1} : \mathbb{R}^3 \rightarrow [0, \pi] \times [0, 2\pi] \times \mathbb{R}_{\geq 0} ; \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} \phi \\ \theta \\ r \end{bmatrix} = \begin{bmatrix} \phi \\ \theta \\ r \end{bmatrix} \quad (19)$$

The Laplacian in spherical coordinates is given by

$$\begin{aligned} \Delta &= \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \\ &=: \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{1}{r^2} \mathcal{L}^2 \end{aligned} \quad (20)$$

where we defined

$$\mathcal{L}^2 = -\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) - \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2}. \quad (21)$$

The separation of the radial and angular part in Eq. (20) suggests to use the separation of variables ansatz, i.e.,

$$\Psi(r, \theta, \phi) = R(r)Y(\theta, \phi). \quad (22)$$

Inserting this ansatz into the Schrödinger equation (12), this yields

$$\begin{aligned} & \left(-\frac{1}{2} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{2} \frac{1}{r^2} \mathcal{L}^2 - \frac{1}{|r|} \right) R(r)Y(\theta, \phi) = ER(r)Y(\theta, \phi) \\ \Leftrightarrow & \left(-\frac{1}{2} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{1}{|r|} - E \right) R(r) + \frac{R(r)}{2r^2} \frac{\mathcal{L}^2 Y(\theta, \phi)}{Y(\theta, \phi)} = 0 \end{aligned} \quad (23)$$

We see that solving this equation is subject to characterizing the spectrum of \mathcal{L}^2 .

2.0.1 Spectrum of \mathcal{L}^2

We seek to solve the eigenvalue problem

$$\mathcal{L}^2 Y(\theta, \phi) = k Y(\theta, \phi). \quad (24)$$

Given the structure of the operator \mathcal{L}^2 , see Eq. (21), we again propose the separation of variables ansatz

$$Y(\theta, \phi) = \Theta(\theta)\Phi(\phi), \quad (25)$$

which yields

$$\begin{aligned} & \left(-\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) - \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right) \Theta(\theta)\Phi(\phi) = k\Theta(\theta)\Phi(\phi) \\ \Leftrightarrow & \frac{\sin^2 \theta}{\Theta(\theta)} \left(-\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) - k \right) \Theta(\theta) - \frac{1}{\Phi(\phi)} \frac{\partial^2}{\partial \phi^2} \Phi(\phi) = 0 \end{aligned} \quad (26)$$

We first separate $\Phi(\phi)$ which yields

$$\begin{cases} -\frac{\partial^2}{\partial \phi^2} \Phi(\phi) = m^2 \Phi(\phi) \\ \Phi(0) = \Phi(2\pi) \end{cases} \quad (27)$$

with separation constant m^2 , and periodic boundary condition because of the spherical symmetry of the problem. The solution to Eq. (27) is given by

$$\Phi(\phi) = A e^{im\phi}, \quad (28)$$

with $m \in \mathbb{Z}$ to fulfill the azimuthal periodic boundary condition.

For $\Theta(\theta)$ we then obtain

$$\left(-\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{m^2}{\sin^2 \theta} \right) \Theta(\theta) = k \Theta(\theta) \quad (29)$$

Performing the coordinate transformation

$$\zeta := \cos(\theta) \quad \text{and} \quad \xi(\cos(\theta)) := \Theta(\theta) \quad (30)$$

yields

$$\frac{d}{d\zeta} \left((1 - \zeta)^2 \frac{d\xi}{d\zeta} \right) + \left(k - \frac{m^2}{1 - \zeta^2} \right) \xi = 0 \quad (31)$$

with $|\xi(1)|, |\xi(-1)| < \infty$. This differential equation is a (well-)known differential equation, namely, the generalized Legendre differential equation [?]. Recall the conventional Legendre equation, i.e.,

$$(1 - x^2) \frac{d^2y}{dx^2} - 2x \frac{dy}{dx} + \ell(\ell + 1)y = 0, \quad (32)$$

where ℓ is the degree. This equation describes solutions to Laplace's equation for azimuthally symmetric boundary conditions, i.e., when the solution has no dependence on the azimuthal angle ϕ . However, to account for non-zero azimuthal dependence, particularly in gravitational and electromagnetic potentials, the equation was extended to the generalized Legendre differential equation, i.e.,

$$(1 - x^2) \frac{d^2y}{dx^2} - 2x \frac{dy}{dx} + \left[\ell(\ell + 1) - \frac{m^2}{1 - x^2} \right] y = 0, \quad (33)$$

introducing the order m , which corresponds to the projection of angular momentum along the z -axis. The solutions to this equation are the associated Legendre functions $P_\ell^m(x)$, which are derived from the Legendre polynomials $P_\ell(x)$ using the relation

$$P_\ell^m(x) = (1 - x^2)^{\frac{m}{2}} \frac{d^m}{dx^m} P_\ell(x). \quad (34)$$

These functions generalize $P_\ell(x)$ by incorporating azimuthal dependence and are crucial in describing solutions to Laplace's equation in spherical coordinates for problems with full rotational symmetry. This formalism, developed to handle problems involving angular momentum and wave-like behavior, became essential in physics, geophysics, and engineering.

The Eq. (31) is a Legendre differential equation on $[-1, 1]$, which yields

$$k = \ell(\ell + 1), \quad \ell \in \mathbb{Z}_{\geq 0} \quad (35)$$

Each eigenvalue $\ell(\ell + 1)$ corresponds to $\ell + 1$ degenerate, orthogonal eigenfunctions, denoted

$$\xi(\zeta) = P_\ell^m(\zeta) := (-1)^m (1 - \zeta^2)^{\frac{m}{2}} \frac{d^m}{d\zeta^m} P_\ell(\zeta), \quad m \in \{0, \dots, \ell\} \quad (36)$$

where P_ℓ^m are the associated Legendre polynomials and

$$P_\ell(\zeta) = \frac{(-1)^\ell}{2^\ell \ell!} \frac{d^\ell}{d\zeta^\ell} (1 - \zeta^2)^\ell \quad (37)$$

are the Legendre polynomials [?]. We also define the *negative associated Legendre polynomials* as

$$P_\ell^{-m}(\zeta) = (-1)^m \frac{(\ell - m)!}{(\ell + m)!} P_\ell^m(\zeta) \quad (38)$$

Example 2.0.1. The first six Legendre polynomials are given by

$$\begin{aligned} P_0(\zeta) &= 1 & P_1(\zeta) &= \zeta & P_2(\zeta) &= \frac{1}{2}(3\zeta^2 - 1) \\ P_3(\zeta) &= \frac{1}{2}(5\zeta^3 - 3\zeta) & P_4(\zeta) &= \frac{1}{8}(35\zeta^4 - 30\zeta^2 + 3) & P_5(\zeta) &= \frac{1}{8}(63\zeta^5 - 70\zeta^3 + 15\zeta) \end{aligned}$$

The first few associated Legendre polynomials are give by

$$\begin{aligned} P_0^0(x) &= 1 \\ P_1^{-1}(x) &= \frac{(1-x^2)^{1/2}}{2} & P_1^0(x) &= x & P_1^1(x) &= -(1-x^2)^{1/2} \\ P_2^{-2}(x) &= \frac{3(1-x^2)}{24} & P_2^{-1}(x) &= \frac{3x(1-x^2)^{1/2}}{6} & P_2^0(x) &= \frac{(3x^2-1)}{2} & P_2^1(x) &= -3x(1-x^2)^{1/2} & P_2^2(x) &= 3(1-x^2) \end{aligned}$$

Changing the variables back to θ and $\Theta(\theta)$, this yields that

$$\Theta(\theta) = P_\ell^m(\cos(\theta)) \quad (39)$$

Hence

$$Y_{\ell,m}(\theta, \phi) = C_{\ell,m} P_\ell^m(\cos(\theta)) e^{im\phi} \quad (40)$$

where $C_{\ell,m}$ a normalization constant which is chosen such that

$$\begin{aligned} &\int_{[0,\phi] \times [0,2\pi)} Y_{\ell,m}(\theta, \phi) Y_{\ell',m'}(\theta, \phi) \sin^2(\theta) d\theta d\phi = \delta_{\ell,\ell'} \delta_{m,m'} \\ \Leftrightarrow \quad C_{\ell,m} &= (-1)^m \sqrt{\frac{(2\ell+1)}{4\pi} \frac{(\ell-m)!}{(\ell+m)!}} \end{aligned} \quad (41)$$

The functions $Y_{\ell,m}(\theta, \phi)$ are the *spherical harmonics* fulfilling

$$\mathcal{L}^2 Y_{\ell,m} = \ell(\ell+1) Y_{\ell,m} \quad (42)$$

where the eigenvalue had geometric multiplicity $2\ell+1$.

Example 2.0.2. Let's consider $\ell = 1$ and $m = 0$, then

$$Y_{1,0}(\theta, \phi) = \sqrt{\frac{3}{4\pi}} \cos(\theta). \quad (43)$$

We see that $Y_{1,0}$ only depends on the polar angle θ and not on the azimuthal angle ϕ . A possible way to visualize this function is via the map

$$f_\ell^m : (\theta, \phi) \mapsto P_\ell^m(\cos(\theta)) e^{im\phi} \begin{bmatrix} \sin(\theta) \sin(\phi) \\ \sin(\theta) \cos(\phi) \\ \cos(\theta) \end{bmatrix} \quad (44)$$

Using this procedure, we can visualize $Y_{1,0}$ as shown in Fig. 4.

This procedure can similarly be applied to the first few spherical harmonics in Fig. 5.

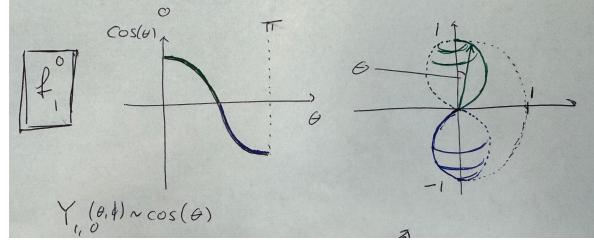


Figure 4: Construction of a visualization of $Y_{1,0}$.

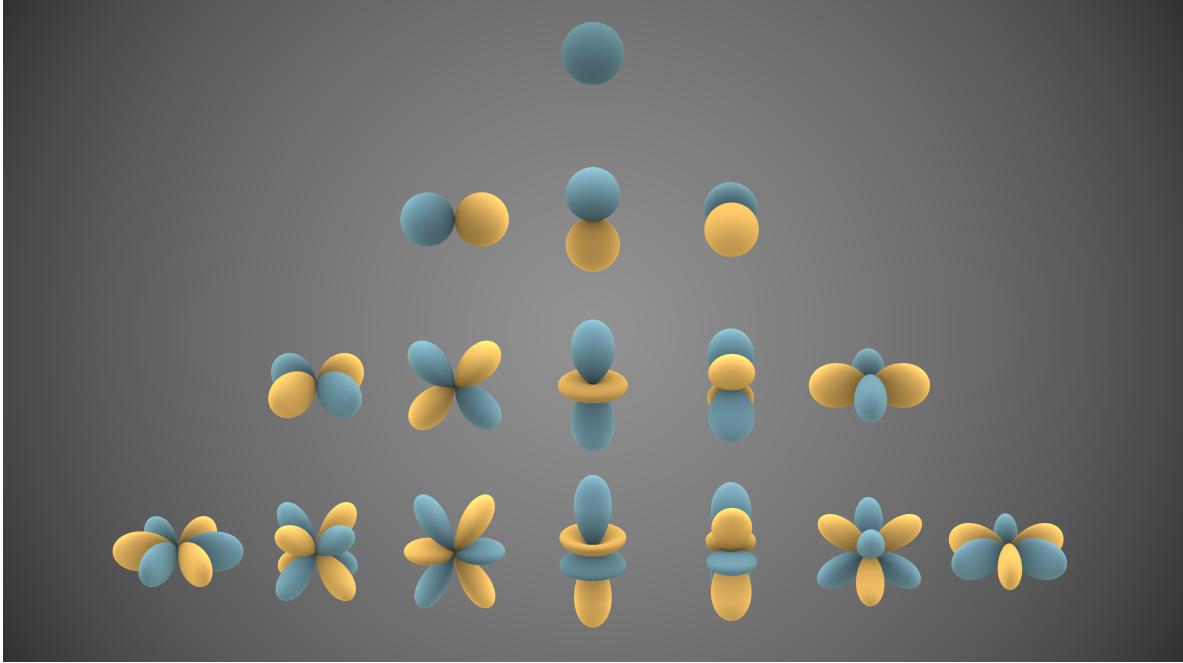


Figure 5: Graphical representation of the first few spherical harmonics, using the function f_ℓ^m in Eq. (44). Blue portions represent regions where the function is positive, and yellow portions represent where it is negative. The rows correspond to $\ell = 0$ (s), $\ell = 1$ (p), $\ell = 2$ (d), $\ell = 3$ (f); the columns correspond to $m \in \{-\ell, \dots, \ell\}$.

We shall now return to Eq. (23). Having found the spectrum of \mathcal{L}^2 , Eq. (23) simplifies to

$$-\frac{1}{2} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R(r)}{\partial r} \right) - \frac{1}{r} R(r) + \frac{R(r)}{2r^2} \ell(\ell+1) = E R(r), \quad r > 0, \quad (45)$$

which is the radial equation. Note that

$$\begin{aligned} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R(r)}{\partial r} \right) &= \frac{1}{r^2} \left(2r \frac{\partial R(r)}{\partial r} + r^2 \frac{\partial^2 R(r)}{\partial r^2} \right) = \left(\frac{2}{r} \frac{\partial R(r)}{\partial r} + \frac{\partial^2 R(r)}{\partial r^2} \right) \\ &= \frac{1}{r} \frac{\partial}{\partial r} \left(R(r) + r \frac{\partial R(r)}{\partial r} \right) = \frac{1}{r} \frac{\partial^2}{\partial r^2} r R(r) \end{aligned} \quad (46)$$

Substituting

$$u(r) = r R(r) \quad (47)$$

yields

$$\begin{aligned} & -\frac{1}{2} \frac{\partial^2}{\partial r^2} r R(r) - R(r) + \frac{R(r)}{2r} \ell(\ell+1) = Er R(r) \\ \Leftrightarrow \quad & -\frac{1}{2} \frac{\partial^2}{\partial r^2} u(r) + \tilde{V}(r) u(r) = Eu(r) \end{aligned} \quad (48)$$

where

$$\tilde{V}(r) = \frac{\ell(\ell+1)}{2r^2} - \frac{1}{r}. \quad (49)$$

We will now investigate the solution to Eq. (48) in two limiting cases. First, in the far field. We note that

$$\tilde{V}(r) \xrightarrow{r \rightarrow \infty} 0 \quad (50)$$

which simplifies Eq. (48) to

$$-\frac{1}{2} \frac{\partial^2}{\partial r^2} u(r) = Eu(r). \quad (51)$$

For $E > 0$ the solution to Eq. (51) yields

$$u(r) = c_1 e^{i\sqrt{E}2r} + c_2 e^{-i\sqrt{E}2r}, \quad (52)$$

which cannot be square integrable, hence, any value $E > 0$ cannot be an isolated eigenvalue. For $E \leq 0$ the solution is given by

$$u(r) = c_1 e^{\sqrt{|E|}2r} + c_2 e^{-\sqrt{|E|}2r}. \quad (53)$$

Since this solution must remain finite as $r \rightarrow \infty$, c_1 must be zero, hence,

$$u(r) = ce^{-\sqrt{|E|}2r}. \quad (54)$$

Second, in the near field. We note that

$$\tilde{V}(r) \xrightarrow{r \rightarrow 0} -\frac{\ell(\ell+1)}{2r^2} \quad \text{and} \quad u(r) \xrightarrow{r \rightarrow 0} 0 \quad (55)$$

which yields

$$-\frac{1}{2} \frac{\partial^2}{\partial r^2} u(r) - \frac{\ell(\ell+1)}{r^2} u(r) = 0. \quad (56)$$

Equation (56) has the solution

$$u(r) = c_1 r^{\ell+1} + c_2 r^{-\ell} \quad (57)$$

but since $r^{-\ell} \rightarrow \infty$ for $r \rightarrow 0$, c_2 must be zero, hence,

$$u(r) = cr^{\ell+1}. \quad (58)$$

Combining the near field and far field solution then yields the general solution

$$u(r) = Cr^{\ell+1} e^{-\sqrt{\frac{|E|}{2}}r} G(r) \quad (59)$$

where

$$G(r) = A_0 + A_1 r + A_2 r^2 + \dots \quad (60)$$

“stitches” the near field and far field solution together. Inserting this ansatz into Eq. (48) yields the recursion

$$A_k = -2A_{k-1} \frac{1 - (\ell + k)\sqrt{-E2}}{(\ell + k)(\ell + k + 1) - \ell(\ell + 1)}. \quad (61)$$

For this series to terminate we require

$$0 = 1 - (\ell + k)\sqrt{-2E} \quad \Leftrightarrow \quad E_{k,\ell} = -\frac{1}{2(\ell + k)^2} \quad (62)$$

3 Spin

We now want to transition to multi-electron systems. To that end, we need to address a fundamental quantity that is – surprisingly – not explicitly incorporated into the electronic Schrödinger equation, namely, the spin. Understanding the spin of an electron is a cornerstone of quantum mechanics, yet it can be an abstract concept for learners. The Stern-Gerlach experiment provides an intuitive and visualizable way to grasp the key ideas surrounding spin. In the early 20th century, physicists were intrigued by the idea that electrons and other particles possess intrinsic angular momentum beyond what classical physics could explain. This intrinsic property, later named *spin*, was suspected but not clearly understood. The Stern-Gerlach experiment, conducted in 1922 by Otto Stern and Walther Gerlach, aimed to probe the nature of this intrinsic property by observing the deflection of particles in a magnetic field.

The core idea of the experiment is deceptively simple. Here's a breakdown of the essential components, see Fig. 6:

- Source of Silver Atoms: The experiment initially used silver atoms – which contain a single valence electron in an unpaired state (Ag has 47 electrons) – as a convenient proxy for electrons.
- Collimated Beam: The silver atoms are heated and directed into a narrow beam.
- Non-Uniform Magnetic Field: The atoms pass through a region of non-uniform magnetic field, which exerts a different force depending on the magnetic moment of each atom.
- Detection Screen: After passing through the magnetic field, the atoms strike a detection screen, revealing their distribution.

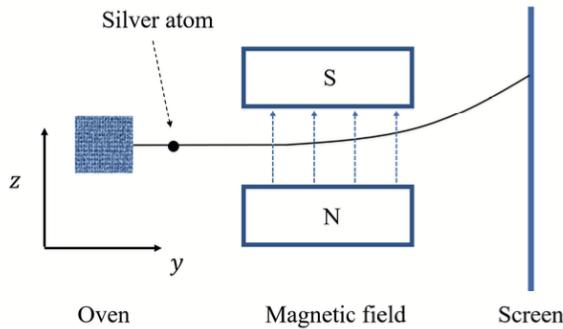


Figure 6: Setup of the Stern-Gerlach experiment

Hot silver atoms emerge from an oven with an initial velocity directed along the y -axis. These atoms then pass through a region containing an inhomogeneous magnetic field, created by combining a uniform magnetic field oriented along the z -axis with a small perturbation that introduces a gradient. As the silver atoms traverse this magnetic field, their final positions along the z -axis are recorded on a detection screen placed to the right. Inside the oven, individual silver atoms are produced carrying one unpaired valence electron that imparts a net magnetic moment. This magnetic moment, referred to as the electron's spin, is represented by a vector $\mu \in \mathbb{R}^3$. The direction of this vector determines whether the atom's trajectory bends upwards or downwards within the magnetic field gradient.

Before the experiment was conducted, classical physics predicted that the magnetic moments of the atoms would orient randomly. As a result, a continuous smear of detection points was expected on the screen, see Fig. 7.

However, quantum mechanics made a strikingly different prediction: due to the quantized nature of spin, the beam should split into discrete parts, see Fig. 7.

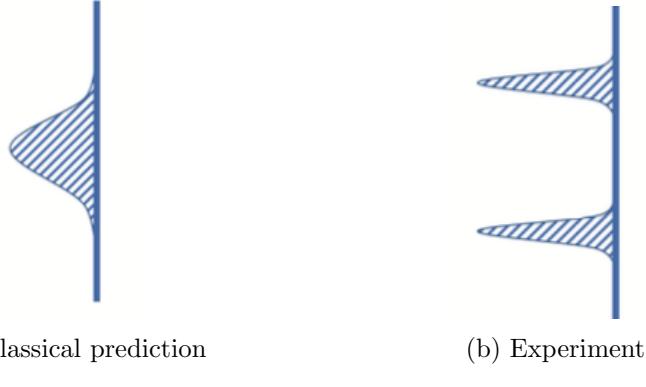


Figure 7: (a) The prediction of the result of the Stern–Gerlach experiment from classical theory and (b) the experimental result.

By blocking one particle beam, we may use the Stern–Gerlach experiment to generate filters, see Fig. 8. Embracing the Dirac notation, we denote the two states that pass through a Stern–Gerlach apparatus with magnetic field in z -direction (SG_z apparatus) by $|+_z\rangle$ and $|-_z\rangle$, respectively, see Fig. 8.



Figure 8: (a) The Stern–Gerlach apparatus along the z -direction and (b) a filtering apparatus.

Note that the spatial orientation of the magnetic field does not have any fundamental effect on the outcome. More precisely, we may rotate the apparatus in a way that the magnetic field is oriented along the x -axis instead, and the outcome would be the same as above, however, rotated around the y -axis. The resulting apparatus would be a Stern–Gerlach apparatus with a magnetic field in x -direction (SG_x apparatus) and the out-coming states would be $|+_x\rangle$ and $|-_x\rangle$, respectively.

Having the ability to combine different filters may allow us to perform a number of experiments. First, we would combine two SG_z apparatuses – or equivalently two SG_x apparatuses. Once a spin is filtered out, we do not observe a beam split if we filter twice or n times in fact, see Fig. 9. On the other hand, filtering one spin out of the beam does not affect the other spin, i.e., once an x -spin is filtered out, we still observe a beam split when transitioning through a SG_z apparatus.

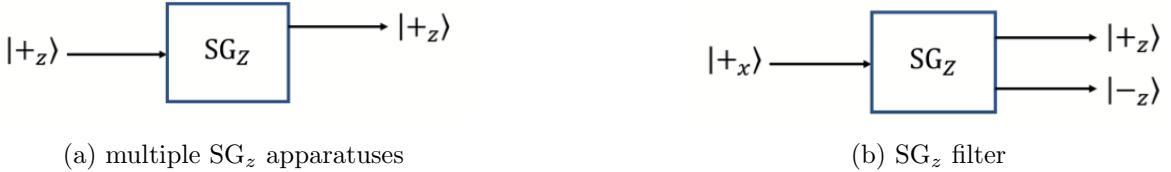


Figure 9: (a) The Stern–Gerlach apparatus along the z -direction and (b) a filtering apparatus.

An interesting observation happens when concatenating three SG filters, $\text{SZ}_z\text{--}\text{SZ}_x\text{--}\text{SZ}_z$. Intuitively, the first filter would filter out a spin component in z -direction, the second filter would filter out a spin component in x -direction, and we should no observe a beam split when the beam transitions the third SZ filter. However, the experiment shows a different outcome, see Fig. 10.

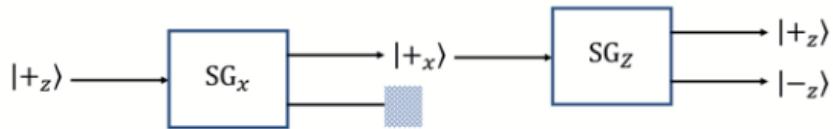


Figure 10: Sequence of three SG filters $\text{SZ}_z\text{--}\text{SZ}_x\text{--}\text{SZ}_z$.

The solution to this mysterious outcome can be obtained using only 2×2 matrices and will lead to the Heisenberg uncertainty principle. The theory of a single spin- $\frac{1}{2}$ particle can be described using a two-dimensional vector space $\mathcal{H} \simeq \mathbb{C}^2$ (isometric isomorphic) called the state vector space. The states $|\pm_z\rangle$ form a basis of \mathcal{H} , i.e., for any $|\psi\rangle \in \mathcal{H}$ there exist $c_1, c_2 \in \mathbb{C}$, s.t.

$$|\psi\rangle = c_1|+z\rangle + c_2|-z\rangle. \quad (63)$$

In particular, $|\pm_x\rangle$ are also states in \mathcal{H} and can be extended using $|\pm_z\rangle$. Moreover, \mathcal{H} is a Hilbert space, and $|\pm_z\rangle$ are orthonormal w.r.t. the inner product structure. In Dirac notation, this means

$$\langle +z|+z\rangle = \langle -z|-z\rangle = 1 \quad \text{and} \quad \langle +z|-z\rangle = 0. \quad (64)$$

Now, since the orientation of the Stern-Gerlach experiment is somewhat arbitrary, we may change $|\pm_z\rangle$ to $|\pm_x\rangle$ in the above discussion. This in turn means that $|\pm_x\rangle$ are linearly independent and therewith form a basis of \mathcal{H} .

In quantum mechanics, a measurement corresponds to a linear operator that is self-adjoint on \mathcal{H} , and the possible outcomes correspond to the operator's spectrum. The Stern-Gerlach apparatus SG_z corresponds to a measurement that measures the spin in z -direction. According to the above postulate, this means that it corresponds to a linear and self-adjoint operator on \mathcal{H} , we shall denote it S_z , and we know the eigenstates of S_z are $|\pm_z\rangle$. Similarly, we may define S_x and S_y together with their spectra $|\pm_x\rangle$ and $|\pm_y\rangle$, respectively. Since $|\pm_x\rangle \in \mathcal{H}$ and $|\pm_y\rangle \in \mathcal{H}$ we may wonder about their expansion using the basis $|\pm_z\rangle$. The Stern-Gerlach experiment provides some insight into this expansion. We first note that when $|+x\rangle$ passes through SG_z the outcome is $|\pm_x\rangle$ with equal probability, i.e., the outcome suggests a bimodal symmetric distribution. Since we moreover have a phase that can be introduced, we find

$$|+x\rangle = \frac{1}{\sqrt{2}} (|+z\rangle + e^{i\alpha} |-z\rangle) \quad (65)$$

and we find

$$|-x\rangle = \frac{1}{\sqrt{2}} (|+z\rangle - e^{i\alpha} |-z\rangle) \quad (66)$$

Similarly we find

$$|+y\rangle = \frac{1}{\sqrt{2}} (|+z\rangle + e^{i\beta} |-z\rangle) \quad (67)$$

and

$$|-y\rangle = \frac{1}{\sqrt{2}} (|+z\rangle - e^{i\beta} |-z\rangle) \quad (68)$$

In order to relate α and β we consider passing $|\pm_x\rangle$ through SG_y ; which yields a bimodal distribution! Hence

$$|\langle \pm_y | \pm_x \rangle|^2 = \frac{1}{2} \quad (69)$$

Since

$$\begin{aligned} \langle \pm_y | \pm_x \rangle &= \frac{1}{2} (\langle +z | \pm e^{-i\beta} \langle -z |) (|+z\rangle \pm e^{i\alpha} |-z\rangle) \\ &= \frac{1}{2} (\langle +z | +z \rangle \pm e^{-i\beta} \langle -z | +z \rangle \pm e^{i\alpha} \langle +z | -z \rangle + e^{-i(\beta-\alpha)} \langle -z | -z \rangle) \\ &= \frac{1}{2} (1 + e^{-i(\beta-\alpha)}) \end{aligned} \quad (70)$$

this yields

$$\frac{1}{2} = |\langle \pm_y | \pm_x \rangle|^2 = \frac{1}{2} (1 + \cos(\beta - \alpha)) \quad (71)$$

and therewith

$$\alpha - \beta = \pm \frac{\pi}{2} + 2\pi n, \quad n \in \mathbb{Z}. \quad (72)$$

We introduce the convention that $\alpha = 0$ and $\beta = \pi/2$ which yields that

$$|\pm_x\rangle = \frac{1}{\sqrt{2}} (|+z\rangle \pm |-z\rangle) \quad \text{and} \quad |\pm_y\rangle = \frac{1}{\sqrt{2}} (|+z\rangle \pm i|-z\rangle) \quad (73)$$

Since $\mathcal{H} \simeq \mathbb{C}^2$ we may perform the identification

$$|+z\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad |-z\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (74)$$

which yields

$$|+x\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad |-x\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (75)$$

and

$$|+_y\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} \quad \text{and} \quad |-_y\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix} \quad (76)$$

This choice of functions allows us to formulate S_x , S_z , and S_z in matrix form. To that end, we use the spectral theorem which yields

$$S_z = \frac{1}{2} (|+_z\rangle\langle +_z| - |-_z\rangle\langle -_z|) = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (77)$$

as $|\pm_z\rangle$ are eigenvectors corresponding to eigenvalues $\pm\frac{1}{2}$. The choice of these eigenvalues. Similarly we find

$$\begin{aligned} S_x &= \frac{1}{2} (|+_x\rangle\langle +_x| - |-_x\rangle\langle -_x|) = \frac{1}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ S_y &= \frac{1}{2} (|+_y\rangle\langle +_y| - |-_y\rangle\langle -_y|) = \frac{1}{2} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \end{aligned} \quad (78)$$

Following the convention that $S = (S_x, S_y, S_z) = \frac{1}{2}(\sigma_x, \sigma_y, \sigma_z)$ we define the Pauli matrices as

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (79)$$

One immediate observation is that linear operators for different observables do not necessarily commute, e.g.,

$$\begin{aligned} S_x S_z &= \frac{1}{4} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \\ S_z S_x &= \frac{1}{4} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \end{aligned} \quad (80)$$

which yields

$$[S_z, S_x] = S_z S_x - S_x S_z = \frac{1}{2} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = i S_y \quad (81)$$

This means that S_x and S_z are not simultaneously diagonalizable and we refer to them as incompatible in a quantum mechanical context. Similarly, we find that

$$[S_x, S_y] = i S_z \quad \text{and} \quad [S_y, S_z] = i S_x \quad (82)$$

Hence, S_x , S_y and S_z are mutually incompatible. On the other hand

$$S^2 = S_x^2 + S_y^2 + S_z^2 \quad (83)$$

is compatible with all spin operators along any individual direction.