Quantum Mechanics Oral Exam Study Guide

Notes covering typical questions from the oral exam required to pass the third-year undergraduate course *Kvanta mehanika* (Quantum Mechanics), given by professor Anton Ramšak at the Faculty of Mathematics and Physics at the University of Ljubljana in the academic year 2020-21. Credit for the material covered in these notes is due to professor Ramšak, while the voice, typesetting, and translation to English in this document are my own.

Disclaimer: Mistakes—both simple typos and legitimate errors—are likely. Keep in mind that these are the notes of an undergraduate student in the process of learning the material himself—take what you read with a grain of salt. If you find mistakes and feel like telling me, for example by email, I'll be happy to hear from you, even for the most trivial of errors.

How the notes are organized

- The sections follow the order in which the course material was taught.

- Each typical exam question is covered in a dedicated subsection, where the question is written in *italics* at the beginning of the subsection.

- Each question subsection may potentially be further divided into subsubsections to better organize the answer material.

For more notes like this, visit `ejmastnak.github.io/fmf`

Contents

1 Fundamentals of Quantum Mechanics

1.1 The Schrödinger Equation

What is the Schrödinger equation? Explain why the Schrödinger equation (and not some other differential equation) forms the basis of quantum mechanics.

- In three dimensions, the Schrödinger equation reads

$$i\hbar\frac{\partial\Psi(\boldsymbol{r},t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi(\boldsymbol{r},t) + V(\boldsymbol{r},t)\Psi(\boldsymbol{r},t).$$

  In one dimension, the Schrödinger equation simplifies to

$$i\hbar\frac{\partial\Psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,t)}{\partial x^2} + V(x,t)\Psi(x,t).$$

- Interpreted from a physical perspective, the Schrödinger equation is the simplest equation whose solution satsifies the following quantum-mechanical properties:

  - The basic solution for a particle is $\Psi(\boldsymbol{r},t) \sim e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}$, meaning particle has wave characteristics—a wavelength $\lambda = 2\pi/k$ and frequency $\nu = 2\pi/\omega$.
  - A particle's energy is proportional to its frequency, i.e. $E = \hbar\omega$, as in, for example, the photoelectric effect.
  - Momentum is related to a wave vector via $\boldsymbol{p} = \hbar\boldsymbol{k}$, as in the de Broglie relation.
  - A free particle has the classical energy $E = \frac{p^2}{2m}$ and dispersion relation $\omega \propto k^2$.

- To show the Schrödinger equation corresponds to oscillation, we decompose Ψ into real and imaginary components via

$$\Psi = \Psi_{\mathrm{Re}} + i\Psi_{\mathrm{Im}}.$$

  Substituting this decomposition into the Schrödinger equation produces

$$i\hbar(\dot{\Psi}_{\mathrm{Re}} + i\dot{\Psi}_{\mathrm{Im}}) = -\frac{\hbar^2}{2m}(\Psi''_{\mathrm{Re}} + i\Psi''_{\mathrm{Im}}) + V(\Psi_{\mathrm{Re}} + i\Psi_{\mathrm{Im}}).$$

  Separating the real and imaginary parts produces the coupled system of real equations

$$-\hbar\dot{\Psi}_{\mathrm{Im}} = -\frac{\hbar^2}{2m}\Psi''_{\mathrm{Re}} + V\Psi_{\mathrm{Re}}$$

$$-\hbar\dot{\Psi}_{\mathrm{Re}} = -\frac{\hbar^2}{2m}\Psi''_{\mathrm{Im}} + V\Psi_{\mathrm{Im}}.$$

  This coupled system leads to the desired oscillation and wavelike behavior of the wavefunction Ψ, even though the Schrödinger equation is first degree in time (unlike e.g. the wave equation, which is second order in time).

Note: On the Diffusion Equation
Note the similarity of the Schrödinger equation to the diffusion equation

$$\frac{\partial T}{\partial t} = D\frac{\partial^2 T}{\partial x^2}.$$

Both are first degree in time and second degree in position. The wave-like ansatz $T(x,t) \propto e^{i(kx-\omega t)}$ solves the diffusion equation with a quadratic dispersion relation

$$\omega = -iDk^2,$$

as desired for a wave-like quantum particle. However, the desired energy relationship $E = \frac{p^2}{2m}$ holds only for $D = \frac{i\hbar}{2m} \in \mathbb{C}$, and a complex diffusion constant is non-physical. We thus reject the diffusion equation.

1.2 The Wave Function, Probability and the Continuity Equation

What is a wave function and what is its role in quantum mechanics? How is the wavefunction related to the probability of detecting a particle in a region of space?

- The wavefunction $\Psi(\boldsymbol{r},t)$ is the fundamental quantity used to describe a particle in quantum mechanics. The wavefunction is used to defined probability density $\rho(\boldsymbol{r},t)$ according to

$$\rho(\boldsymbol{r},t) = |\Psi(\boldsymbol{r},t)|^2,$$

  and the probability $\mathrm{d}P$ of finding a particle in the region of space $\mathrm{d}V$ is

$$\mathrm{d}P = \rho(\boldsymbol{r},t)\,\mathrm{d}V.$$

- Wavefunctions are expected to obey the normalization condition

$$\int_V |\Psi(\boldsymbol{r},t)|^2\,\mathrm{d}^3\boldsymbol{r} \equiv 1.$$

  Interpreted physically, the normalization condition means the probability of finding the particle somewhere in all of space V is one:

- The probability current density $\boldsymbol{j}$ associated with a particle with wavefunction Ψ is

$$\boldsymbol{j}(\boldsymbol{r},t) = \frac{\hbar}{2im}\big[\Psi^*(\boldsymbol{r},t)\nabla\Psi(\boldsymbol{r},t) - \Psi(\boldsymbol{r},t)\nabla\Psi^*(\boldsymbol{r},t)\big],$$

  and the continuity equation encoding conservation or probability reads

$$\frac{\partial\rho(\boldsymbol{r},t)}{\partial t} + \nabla\cdot\boldsymbol{j}(\boldsymbol{r},t) = 0.$$

Derivation: Probability Current and the Continuity Equation

- We begin by differentiating probability density with the product rule, which produces

$$\frac{\partial\rho(\boldsymbol{r},t)}{\partial t} = \frac{\partial|\Psi(\boldsymbol{r},t)|^2}{\partial t} = \frac{\partial\Psi^*(\boldsymbol{r},t)}{\partial t}\Psi(\boldsymbol{r},t) + \frac{\partial\Psi(\boldsymbol{r},t)}{\partial t}\Psi^*(\boldsymbol{r},t).$$

  We then substitute in $\dot{\Psi}$ from the Schrödinger equation to get

$$\frac{\partial\rho(\boldsymbol{r},t)}{\partial t} = \left(-\frac{\hbar i}{2m}\nabla^2\Psi^*(\boldsymbol{r},t) + \frac{i}{\hbar}V^*(\boldsymbol{r},t)\Psi^*(\boldsymbol{r},t)\right)\Psi(\boldsymbol{r},t)$$

$$+ \left(\frac{\hbar i}{2m}\nabla^2\Psi(\boldsymbol{r},t) - \frac{i}{\hbar}V(\boldsymbol{r},t)\Psi(\boldsymbol{r},t)\right)\Psi^*(\boldsymbol{r},t),$$

  where we have allowed the possibility of complex potential $V(\boldsymbol{r},t) \in \mathbb{C}$ when conjugating the Schrödinger equation.

6

- We then use the identity

$$\Psi\nabla^2\Psi^* = \nabla\cdot(\Psi\nabla\Psi^*) - \nabla\Psi\cdot\nabla\Psi^*$$

  to write

$$\frac{\partial\rho(\boldsymbol{r},t)}{\partial t} + \nabla\cdot\boldsymbol{j}(\boldsymbol{r},t) = q(\boldsymbol{r},t),$$

  which motivates the definition of *probability current density* as

$$\boldsymbol{j}(\boldsymbol{r},t) = \frac{\hbar}{2im}\big[\Psi^*(\boldsymbol{r},t)\nabla\Psi(\boldsymbol{r},t) - \Psi(\boldsymbol{r},t)\nabla\Psi^*(\boldsymbol{r},t)\big],$$

  and the *probability source density* as

$$q(\boldsymbol{r},t) = 2\,\mathrm{Im}\big[V(\boldsymbol{r},t)\rho(\boldsymbol{r},t)\big].$$

  The source density q is zero when V is a real function, which is the case in most physical situations.

- Probability is conserved when $q(\boldsymbol{r},t) = 0$ (i.e. when V is real), which results in the continuity equation

$$\frac{\partial\rho(\boldsymbol{r},t)}{\partial t} + \nabla\cdot\boldsymbol{j}(\boldsymbol{r},t) = 0.$$

1.3 Stationary States and Time Evolution

What are stationary states? Include a physical interpretation. State and derive the stationary Schrödinger equation and discuss its relationship to stationary states. Discuss the relationship between the stationary Schrödinger equation and time evolution.

- Stationary state solutions of the Schrödinger equation are a quantum-mechanical analog of standing wave solutions of the wave equation. "Standing wavefunctions" in the Schrödinger equation occur when the wavefunction can be factored into the product of a position-dependent and time-dependent wavefunction in the form

$$\Psi(\boldsymbol{r},t) = \psi(\boldsymbol{r})f(t).$$

  Solutions of this form are called *stationary states*.

- The stationary Schrödinger equation reads

$$-\frac{\hbar^2}{2m}\nabla^2\psi_n(\boldsymbol{r}) + V(\boldsymbol{r})\psi_n(\boldsymbol{r}) = E_n\psi_n(\boldsymbol{r}),$$

  and is used to solve for a stationary state ψ_n with energy E_n.

Derivation: Stationary Schrödinger Equation

- To derive the stationary Schrödinger equation, we assume the potential is independent of time, i.e. $V = V(\boldsymbol{r})$. We then substitute the stationary ansatz $\Psi(\boldsymbol{r},t) = \psi(\boldsymbol{r})f(t)$ into the Schrödinger equation to get

$$i\hbar\psi(\boldsymbol{r})\frac{\partial f(t)}{\partial t} = -\frac{\hbar^2}{2m}f(t)\nabla^2\psi(\boldsymbol{r}) + V(\boldsymbol{r})f(t)\psi(\boldsymbol{r}).$$

  Next, we divide through by $\psi(\boldsymbol{r})f(t)$, which simplifies things to

$$\frac{i\hbar}{f(t)}\frac{\partial f(t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\nabla^2\psi(\boldsymbol{r})}{\psi(\boldsymbol{r})} + V(\boldsymbol{r}).$$

- Since the left-hand side of the equation depends only on time, and the right-hand side only on position, the equality holds for all t and $\boldsymbol{r}$ only if both sides are constant. We make this requirement explicit by writing

$$\frac{i\hbar}{f(t)}\frac{\partial f(t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\nabla^2\psi(\boldsymbol{r})}{\psi(\boldsymbol{r})} + V(\boldsymbol{r}) \equiv E,$$

where the constant E represents the stationary state's energy.

- The position-dependent portion of the separated equation produces the stationary Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2\psi_n(\boldsymbol{r}) + V(\boldsymbol{r})\psi_n(\boldsymbol{r}) = E_n\psi_n(\boldsymbol{r}), \qquad n \in \mathbb{N}.$$

Note that this is an eigenvalue equation for the stationary state eigenfunctions ψ_n and energy eigenvalues E_n. An energy eigenvalue E is called *degenerate* if their exist multiple linearly independent eigenfunctions, e.g. ψ_1, ψ_2, with the same energy eigenvalue E.

- *Important*: The complete set of stationary state eigenfunctions $\{\psi_n(\boldsymbol{r})\}$ form an orthonormal basis of the wavefunction solution space and satisfy the relation

$$\langle\psi_n|\psi_m\rangle \equiv \int \psi_n^*(\boldsymbol{r})\psi_m(\boldsymbol{r})\,\mathrm{d}^3\boldsymbol{r} = \delta_{nm},$$

where δ_{nm} is the Kronecker delta.

- Finally, with the energy eigenvalues E_n known, we solve the time-dependent portion of the separated Schrödinger equation to get

$$f(t) = e^{-i\frac{E_n}{\hbar}t} \equiv e^{-i\omega_n t},$$

which represents oscillation in time with at the frequency ω_n, and thus satisfies the familiar quantum-mechanical relation $E_n = \hbar\omega_n$.

Time Evolution

- Time evolution is the process of writing a time-dependent solution $\Psi(\boldsymbol{r}, t)$ Schrödinger equation in terms of the stationary state eigenfunction basis, given an initial state $\Psi(\boldsymbol{r}, 0)$. The time evolution formula reads

$$\Psi(r, t) = \sum_n c_n e^{-i\frac{E_n}{\hbar}t}\psi_n(\boldsymbol{r}), \qquad c_n = \int_V \psi_n^*\Psi(\boldsymbol{r}, 0)\,\mathrm{d}^3\boldsymbol{r},$$

where E_n and $\psi_n(\boldsymbol{r})$ are the energy eigenvalues and eigenfunctions of the system's Hamiltonian, which we find by solving the stationary Schrödinger equation.

Derivation: Time Evolution Formula

- We first expand the initial state $\Psi(\boldsymbol{r}, 0)$ in the eigenfunction basis $\{\psi_n\}$ to get

$$\Psi(\boldsymbol{r}, 0) = \sum_n c_n\psi_n(\boldsymbol{r}).$$

To find the coefficients c_n, we take the inner product of the above equation with an arbitrary eigenstate ψ_m, which results in

$$\langle \psi_m | \Psi(\boldsymbol{r}, 0) \rangle \equiv \int_V \psi_m^* \Psi(\boldsymbol{r}, 0) \, \mathrm{d}^3 \boldsymbol{r} = \sum_n c_n \langle \psi_m | \psi_n \rangle .$$

We then apply the orthogonality relation $\langle \psi_n | \psi_m \rangle = \delta_{mn}$, which results in

$$\int_V \psi_m^* \Psi(\boldsymbol{r}, 0) \, \mathrm{d}^3 \boldsymbol{r} = \sum_n c_n \delta_{mn} = c_m \implies c_n = \int_V \psi_n^* \Psi(\boldsymbol{r}, 0) \, \mathrm{d}^3 \boldsymbol{r},$$

which is the expression for the coefficients c_n as quoted above.

- We derive the complete expression for $\Psi(\boldsymbol{r}, t)$ from the separated equation

$$\Psi(\boldsymbol{r}, t) = \psi(\boldsymbol{r}) f(t).$$

For the position-dependent term $\psi(\boldsymbol{r})$ we use the known initial state $\Psi(\boldsymbol{r}, t)$, i.e.

$$\psi(\boldsymbol{r}) \to \Psi(\boldsymbol{r}, 0) = \sum_n c_n \psi_n(\boldsymbol{r}).$$

where ψ_n are the eigenfunctions found from solving the stationary Schrödinger equation. For the time-dependent term $f(t)$, we use the solution to the time-dependent portion of the separated Schrödinger equation, which was

$$f(t) = e^{-i \frac{E_n}{\hbar} t}.$$

Substituting $\psi(\boldsymbol{r}) = \Psi(\boldsymbol{r}, 0)$ and $f(t) = e^{-i \frac{E_n}{\hbar} t}$ into $\Psi(\boldsymbol{r}, t) = \psi(\boldsymbol{r}) f(t)$ produces the time evolution formula

$$\Psi(\boldsymbol{r}, t) = \Psi(\boldsymbol{r}, 0) e^{-i \frac{E_n}{\hbar} t} = \sum_n c_n e^{-i \frac{E_n}{\hbar} t} \psi_n(\boldsymbol{r}).$$

1.4 Properties of the Wave Function

Discuss some of the wave function's most important properties.

1.4.1 Normalization, Continuity and Differentiability

- Wavefunctions (are expected to) obey the normalization condition

$$\int_V |\Psi(\boldsymbol{r}, t)|^2 \equiv 1.$$

Physically, this requirement means there is unit probability of detecting a particle somewhere in all of space.

- Wavefunctions are continuous on their entire domain.

- Generally, the wavefunction's first derivative Ψ' is also continuous on its domain. This is derived later in this subsection.

- Points of inflection (zeros of the second derivative Ψ'') occur at the classically-expected turning points where $E = V$. The wavefunction must be smooth at the turning points to satisfy the continuity conditions on Ψ and Ψ'.

  To analyze the second derivative, we write the Schrödinger equation in the form

  $$\frac{1}{\psi(x)} \frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2} = \frac{2m}{\hbar^2}\big[V(x) - E\big].$$

  By observing the sign of $\psi''(x)$ as a function of the value of E, we see that ψ is concave where $E > V$ and convex where $E < V$.

Derivation: Continuity of the First Derivative

- We begin by integrating the stationary Schrödinger equation on an arbitrary interval $x \in [a, b]$, which produces

  $$-\frac{\hbar^2}{2m}\int_a^b \psi''(x)\,\mathrm{d}x + \int_a^b V(x)\psi(x)\,\mathrm{d}x = E\int_a^b \psi(x)\,\mathrm{d}x.$$

  Next, we evaluate the integral of $\psi''(x)$ and rearrange to get

  $$\psi'(b) - \psi'(a) = \frac{2m}{\hbar^2}\int_a^b V(x)\psi(x)\,\mathrm{d}x - \frac{2mE}{\hbar^2}\int_a^b \psi(x)\,\mathrm{d}x.$$

- We then consider the limit behavior $a \to b$. From introductory real analysis, since ψ is continuous, we have $\int_a^b \psi(x)\,\mathrm{d}x \to 0$ as $a \to b$. As long as $V(x)$ is continuous, then $V(x)\psi(x)$ is also continuous, which implies $\int_a^b V(x)\psi(x)\,\mathrm{d}x \to 0$ as $a \to b$ and thus

  $$\lim_{a \to b}\big[\psi'(b) - \psi'(a)\big] = \frac{2m}{\hbar^2}\cdot 0 - \frac{2mE}{\hbar^2}\cdot 0 = 0.$$

  The resulting equality $\lim_{a \to b}\big[\psi'(b) - \psi'(a)\big] = 0$ implies the wavefunction derivative ψ' is also a continuous function.

- Note, however, that the first wavefunction derivative will be discontinuous is the Schrödinger equation contains a discontinuous potential. As an example, if the potential takes the form of a delta function, i.e. $V(x) = \lambda\delta(x)$ where λ is a constant, the wavefunction's first derivative has a discontinuity of the form

  $$\lim_{a \to b}\big[\psi'(b) - \psi'(a)\big] = \frac{2m\lambda}{\hbar^2}\psi(a).$$

1.4.2 Degeneracy and the Nondegeneracy Theorem

In one dimension, all stationary wavefunctions ψ_n that vanish at $\pm\infty$ obey the nondegeneracy theorem, which states that

> The energy eigenvalue spectrum $\{E_n\}$ of a one-dimensional quantum system is degenerate as long as the wavefunctions ψ_n vanish at $\pm\infty$.

For review, an energy eigenvalue E is degenerate if their exist multiple linearly independent eigenfunctions, e.g. ψ_1, ψ_2, with the same energy eigenvalue E.

Derivation

- Consider a one-dimensional quantum system with two eigenfunctions $\psi_1(x)$ and $\psi_2(x)$, both with energy eigenvalue E. The stationary Schrödinger equation for the two eigenfunctions reads

$$-\frac{\hbar^2}{2m}\psi_1''(x) + \left[V(x) - E\right]\psi_1(x) = 0$$
$$-\frac{\hbar^2}{2m}\psi_2''(x) + \left[V(x) - E\right]\psi_2(x) = 0.$$

We multiply the first equation by ψ_1, the second by ψ_2 and subtract the equations to get

$$\psi_1\frac{\mathrm{d}^2\psi_2}{\mathrm{d}x^2} - \psi_2\frac{\mathrm{d}^2\psi_1}{\mathrm{d}x^2} = 0.$$

Mathematical aside: the Wronskian determinant of the wavefunctions ψ_1 and ψ_2 is

$$W_{12} \equiv \det\begin{pmatrix} \psi_1 & \psi_2 \\ \psi_1' & \psi_2' \end{pmatrix} = \psi_1\psi_2' - \psi_2\psi_1'.$$

- In terms of the Wronskian, the above equation relating ψ_1, ψ_2 and their second derivatives reads

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\psi_1\frac{\mathrm{d}\psi_2}{\mathrm{d}x} - \psi_2\frac{\mathrm{d}\psi_1}{\mathrm{d}x}\right) = \frac{\mathrm{d}W_{12}}{\mathrm{d}x} = 0,$$

which implies the Wronskian is constant with respect to x.

- Next, we apply the condition $\psi_{1,2} \to 0$ and $\psi_{1,2}' \to 0$ as $|x| \to \infty$, which implies $W_{12} \to 0$ as $|x| \to \infty$. The result $W_{12} \to 0$ as $|x| \to 0$ in turn implies $W_{12} = 0$ for all x, since W is constant with respect to x—if W_{12} equals zero at infinity, it must equal zero everywhere else on the real line as well. The result $W_{12} = 0$ implies

$$\psi_1\frac{\mathrm{d}\psi_2}{\mathrm{d}x} = \psi_2\frac{\mathrm{d}\psi_1}{\mathrm{d}x} \implies \frac{1}{\psi_1}\frac{\mathrm{d}\psi_1}{\mathrm{d}x} - \frac{1}{\psi_2}\frac{\mathrm{d}\psi_2}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}(\ln\psi_1 - \ln\psi_2) = 0.$$

Integrating the final equality produces

$$\ln\psi_1 - \ln\psi_2 = \ln\frac{\psi_1}{\psi_2} = C \implies \psi_1(x) = D\psi_2(x).$$

where D is a constant. In other words, ψ_1 and ψ_2 are linearly dependent, implying the one-dimensional energy spectrum $\{E_n\}$ is nondegenerate (under the condition that $\psi_{1,2}$ vanish at infinity).

1.5 Properties of Operators

Explain the role of operators in quantum mechanics and state some of their important properties.. Discuss the momentum operator.

- In quantum mechanics, every measurable quantity—called an *observable*—is assigned a corresponding operator. Some common operators are

$$\hat{x} \to x\mathrm{I} \qquad \hat{\boldsymbol{r}} \to \boldsymbol{r}\,\mathrm{I} \qquad \hat{V} \to V(\boldsymbol{r}, t)\mathrm{I},$$

where I is the identity operator. We typically leave the identity operator implicit and write e.g. $\hat{x} \to x$. The momentum operator in various forms reads

$$\hat{p}_\alpha = -i\hbar\frac{\partial}{\partial\alpha} \qquad \hat{p}_\alpha = (-i\hbar)^n\frac{\partial^n}{\partial\alpha^n} \qquad \hat{\boldsymbol{p}} = \sum_{\alpha=x,y,z} \hat{p}_\alpha = -i\hbar\nabla \qquad \hat{\boldsymbol{p}}^2 = (-i\hbar)^2\nabla^2.$$

Notation: In this section I will intermittently write operators with a hat, i.e. $\hat{p}$. However, by convention we usually write operators without the hat symbol and distinguish between operators and scalar quantities based on context.

Functions of Operators

- For an analytic complex function $f(x)$ with the power series definition

$$f(z) = \sum_{n=0}^{\infty} c_n z^n,$$

the function of an operator $\mathcal{O}$, which is itself an operator, is defined as

$$f(\mathcal{O}) = \sum_{n=0}^{\infty} c_n \mathcal{O}^n.$$

For example, the exponential function of an operator $\mathcal{O}$ is defined as via the exponential function's Taylor series as

$$e^{\mathcal{O}} = I + \mathcal{O} + \frac{\mathcal{O}^2}{2!} + \frac{\mathcal{O}^3}{3!} + \cdots + \frac{\mathcal{O}^n}{n!} + \cdots$$

- Functions of operators give simple results when applied to eigenvalue relations. Consider for example $\mathcal{O}$ for which we know the eigenvalue relation $\mathcal{O}\psi = \lambda\psi$. In this case the operator function $f(\mathcal{O})$ applied to ψ reads

$$f(\mathcal{O})\psi \equiv \left(\sum_{n=0}^{\infty} c_n \mathcal{O}^n\right)\psi = \sum_{n=0}^{\infty} c_n \left(\mathcal{O}^n\psi\right) = \sum_{n=0}^{\infty} c_n \lambda^n\psi = f(\lambda)\psi.$$

In other words, the operator expression $f(\mathcal{O})\psi$ reduces to the scalar expression $f(\lambda)\psi$.

1.6 Expectation Values

Define and explain the concept of an expectation value of an operator in the context of quantum mechanics. How do we find the time derivative of an expectation value?

- The expectation value of an operator $\mathcal{O}$ for system described by the wavefunction Ψ with associated probability density $\rho = |\Psi|^2$ is defined as

$$\langle\mathcal{O}\rangle = \int_V \mathcal{O}\rho(\boldsymbol{r}, t)\,\mathrm{d}^3\boldsymbol{r} = \int_V \Psi^*(x, t)\mathcal{O}\Psi(x, t)\,\mathrm{d}^3\boldsymbol{r}.$$

Physically, the expectation value represents the statistically expected average value of many measurements of the quantity associated with the operator $\mathcal{O}$ for the system described by the wavefunciton Ψ.

- The time-derivative of a time-dependent expectation value $\langle \mathcal{O}, t \rangle$ of an operator $\mathcal{O}$ for a quantum system with the wavefunction $\Psi(\boldsymbol{r}, t)$ can be written in the form

$$\frac{\mathrm{d} \langle \mathcal{O}, t \rangle}{\mathrm{d} t} = \left\langle \frac{\partial \mathcal{O}}{\partial t} \right\rangle + \frac{1}{i\hbar} \langle [\mathcal{O}, H] \rangle,$$

where $[\mathcal{O}, H]$ is the commutator of the operator and the system's Hamiltonian.

Note the similarity to an analogous result from classical mechanics for a function $f(p, q)$ of the canonical coordinates, in terms of Poisson brackets, which reads

$$\frac{\mathrm{d} f}{\mathrm{d} t} = \frac{\partial f}{\partial t} + \{f, H\}.$$

Derivation: Time-Dependent Expectation Values

- We begin with the definition of a time-dependent expectation value of the operator $\mathcal{O}$, which reads

$$\langle \mathcal{O}, t \rangle = \int_V \Psi^*(\boldsymbol{r}, t) \mathcal{O} \Psi(\boldsymbol{r}, t) \, \mathrm{d}^3 \boldsymbol{r}.$$

The corresponding time derivative of $\langle \mathcal{O}, t \rangle$, using the product rule, is

$$\frac{\mathrm{d} \langle \mathcal{O}, t \rangle}{\mathrm{d} t} = \int_V \left(\frac{\partial \Psi^*}{\partial t} \mathcal{O} \Psi + \Psi^* \frac{\partial \mathcal{O}}{\partial t} \Psi + \Psi^* \mathcal{O} \frac{\partial \Psi}{\partial t} \right) \mathrm{d}^3 \boldsymbol{r}.$$

- Using the Schrödinger equation, the time derivatives of Ψ can be written in the form

$$\frac{\partial \Psi}{\partial t} = \frac{1}{i\hbar} H \Psi \qquad \text{and} \qquad \frac{\partial \Psi^*}{\partial t} = -\frac{1}{i\hbar} (H\Psi)^*.$$

Substituting these expressions into the time derivative of $\langle \mathcal{O}, t \rangle$ gives

$$\frac{\mathrm{d} \langle \mathcal{O}, t \rangle}{\mathrm{d} t} = \left\langle \frac{\partial \mathcal{O}}{\partial t} \right\rangle + \frac{1}{i\hbar} \int_V \left[-(H\Psi)^* \mathcal{O} \Psi + \Psi^* \mathcal{O} H \Psi \right] \mathrm{d}^3 \boldsymbol{r}$$

$$= \left\langle \frac{\partial \mathcal{O}}{\partial t} \right\rangle + \frac{1}{i\hbar} \int_V (\Psi^* \mathcal{O} H \Psi - \Psi^* H \mathcal{O} \Psi) \, \mathrm{d}^3 \boldsymbol{r},$$

where we have used $(H\Psi)^* = \Psi^* H^*$ and applied the Hermitian identity $H^* = H$.

- Finally, we use a commutator to compactly write the above result for time derivative of $\langle \mathcal{O}, t \rangle$ in the form

$$\frac{\mathrm{d} \langle \mathcal{O}, t \rangle}{\mathrm{d} t} = \left\langle \frac{\partial \mathcal{O}}{\partial t} \right\rangle + \frac{1}{i\hbar} \langle [\mathcal{O}, H] \rangle.$$

1.7 The Momentum Operator

How is the momentum operator defined in one and three dimensions? How is the momentum operator related to probability current? Explain the motivation behind the definition of the momentum operator.

- In one dimension and three dimensions, the momentum operator is defined as

$$p \rightarrow -i\hbar \frac{\partial}{\partial x} \qquad \text{and} \qquad \boldsymbol{p} \rightarrow -i\hbar \nabla.$$

The momentum operator is related to the position operator by

$$\langle \boldsymbol{p} \rangle = \frac{1}{m} \frac{\mathrm{d} \langle \boldsymbol{r} \rangle}{\mathrm{d} t}.$$

- For a particle of mass m with wave function Ψ, the momentum operator and probability current density $\boldsymbol{j}$ are related by

$$\boldsymbol{j}(\boldsymbol{r}, t) = \frac{1}{m} \operatorname{Re} \left[\Psi^*(\boldsymbol{r}, t) \boldsymbol{p} \Psi(\boldsymbol{r}, t) \right] \qquad \text{and} \qquad \langle \boldsymbol{p} \rangle = m \int_V \boldsymbol{j}(\boldsymbol{r}, t) \, \mathrm{d}^3 \boldsymbol{r}.$$

Derivation: Motivation for the Momentum Operator's Definition

- We begin by finding the time derivative of the position expectation value, i.e.

$$\frac{\mathrm{d} \langle x \rangle}{\mathrm{d}t} = \frac{\partial}{\partial t} \int_{-\infty}^{\infty} \Psi^*(x, t) x \Psi(x, t) \, \mathrm{d}x$$

$$= \int_{-\infty}^{\infty} \left(\frac{\partial \Psi^*(x, t)}{\partial t} x \Psi(x, t) + \Psi^*(x, t) x \frac{\partial \Psi(x, t)}{\partial t} \right) \mathrm{d}x.$$

Assuming a real potential $V(x)$ so that $V = V^*$, we then express $\frac{\partial \Psi^*}{\partial t}$ and $\frac{\partial \Psi}{\partial t}$ in terms of $\frac{\partial^2 \Psi^*}{\partial x^2}$ and $\frac{\partial^2 \Psi}{\partial x^2}$ using the Schrödinger equation, substitute these expressions in to the above expression for $\frac{\mathrm{d} \langle x \rangle}{\mathrm{d}t}$, and simplify like terms to get

$$\frac{\mathrm{d} \langle x \rangle}{\mathrm{d}t} = \frac{\hbar}{2im} \int_{-\infty}^{\infty} \left(\frac{\partial^2 \Psi^*(x, t)}{\partial x^2} x \Psi(x, t) - \frac{\partial^2 \Psi(x, t)}{\partial x^2} x \Psi^*(x, t) \right).$$

We then rewrite this result with a reverse-engineered derivative with respect to x:

$$\frac{\mathrm{d} \langle x \rangle}{\mathrm{d}t} = \frac{\hbar}{2im} \int_{-\infty}^{\infty} \frac{\partial}{\partial x} \left(\frac{\partial \Psi^*}{\partial x} x \Psi - |\Psi|^2 - \Psi^* x \frac{\partial \Psi}{\partial x} \right) \mathrm{d}x$$

$$+ \frac{1}{m} \int_{-\infty}^{\infty} \Psi^* \left(-i\hbar \frac{\partial}{\partial x} \Psi \right) \mathrm{d}x.$$

For rapidly falling wavefunctions in the Schwartz space, the first integral evaluates to zero. We are left with

$$\frac{\mathrm{d} \langle x \rangle}{\mathrm{d}t} = \frac{1}{m} \int_{-\infty}^{\infty} \Psi^* \left(-i\hbar \frac{\partial}{\partial x} \Psi \right) \mathrm{d}x$$

- The above result for $\frac{\mathrm{d} \langle x \rangle}{\mathrm{d}t}$, written in the momentum-like form

$$m \frac{\mathrm{d} \langle x \rangle}{\mathrm{d}t} = \langle p \rangle = \int_{-\infty}^{\infty} \Psi^* \left(-i\hbar \frac{\partial}{\partial x} \Psi \right) \mathrm{d}x,$$

motivates the introduction of the momentum operator

$$p \to -\hbar \frac{\partial}{\partial x} \implies \langle p \rangle = \int_{-\infty}^{\infty} \Psi p \Psi \, \mathrm{d}x.$$

1.8 The Ehrenfest Theorem

State and derive the Ehernfest theorem and explain the theorem's physical significance.

- The Ehrenfest theorem relates a quantum particle's momentum $\boldsymbol{p}$ to the potential V felt by the particle. With operators written explicitly, the Eherenfest theorem reads

$$\frac{\mathrm{d} \langle \hat{\boldsymbol{p}}, t \rangle}{\mathrm{d}t} = \langle -\nabla \hat{V} \rangle = \langle \hat{\boldsymbol{F}} \rangle, \qquad \text{where } \hat{\boldsymbol{F}}(\boldsymbol{r}) = -\nabla \hat{V}(\boldsymbol{r}).$$

- The Ehrenfest theorem can be interpreted as an quantum-mechanical analog of Newton's second law, which reads

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = -\nabla V = \boldsymbol{F}.$$

In the Ehrenfest theorem, the classical quantites $\boldsymbol{p}$, V and $\boldsymbol{F}$ are replaced by the expectation values of the quantum mechanical operators $\hat{\boldsymbol{p}}$, $\hat{V}$ and $\hat{\boldsymbol{F}}$.

Derivation

- We first consider the time-dependent expectation value $\langle x, t \rangle$. Using the identity

$$\frac{\mathrm{d}\langle \mathcal{O}, t \rangle}{\mathrm{d}t} = \left\langle \frac{\partial \mathcal{O}}{\partial t} \right\rangle + \frac{1}{i\hbar} \langle [\mathcal{O}, H] \rangle$$

with $\mathcal{O} = x$ and the identity $\frac{\partial x}{\partial t} = 0$ produces the relationship

$$\frac{\mathrm{d}\langle x, t \rangle}{\mathrm{d}t} = \frac{1}{i\hbar} \langle [x, H] \rangle = \frac{1}{i\hbar} \left\langle \left[x, \frac{p_x^2}{2m} + V \right] \right\rangle$$
$$= \frac{1}{2i\hbar m} \langle [x, p_x^2] \rangle + \frac{1}{i\hbar} \langle [x, V] \rangle .$$

The first commutator in the above expression evaluates to

$$[x, p_x^2] = p_x[x, p_x] + [x, p_x]p_x = p_x(i\hbar) + (i\hbar)p_x = 2i\hbar p_x$$

The second is simply $[x, V] = 0$, since x and V commute.

- Using the just-derived intermediate results $[x, p_x^2] = 2i\hbar p_x$ and $[x, V(x,t)] = 0$, the time derivative of $\langle x, t \rangle$ is

$$\frac{\mathrm{d}\langle x, t \rangle}{\mathrm{d}t} = \frac{1}{2i\hbar m} \langle 2i\hbar p_x \rangle + \frac{1}{i\hbar} \langle 0 \rangle = \frac{1}{m} \langle p_x, t \rangle ,$$

in analogy with the classical result $m\dot{x} = p_x$.

- Next, we find the time derivative of $\langle p_x, t \rangle$. Using the general result for the time derivative of an expectation value and implicitly recognizing $\frac{\partial p_x}{\partial t} = 0$, we have

$$\frac{\mathrm{d}\langle p_x, t \rangle}{\mathrm{d}t} = \frac{1}{i\hbar} \langle [p_x, H] \rangle = \frac{1}{i\hbar} \left\langle \left[p_x, \frac{p_x^2}{2m} + V \right] \right\rangle$$
$$= \frac{1}{2i\hbar m} \langle [p_x, p_x^2] \rangle + \frac{1}{i\hbar} \langle [p_x, V] \rangle .$$

The first commutator is simply $[p_x, p_x^2] = 0$, which follows from $[p_x, p_x] = 0$ and $[A, BC] = B[A, C] + [A, B]C$. We find the second as follows:

$$[p_x, V]\psi_x \equiv \left[\left(-i\hbar \frac{\partial}{\partial x} \right) V - V \left(-i\hbar \frac{\partial}{\partial x} \right) \right] \psi = -i\hbar f \frac{\partial V}{\partial x} - i\hbar V \frac{\partial f}{\partial x} + i\hbar V \frac{\partial f}{\partial x}$$
$$= -i\hbar \frac{\partial V}{\partial x} f \implies [p_x, V] = -i\hbar \frac{\partial V}{\partial x}.$$

- Using the just derived intermediate results $[p_x, p_x^2] = 0$ and $[p_x, V] = -i\hbar\frac{\partial V}{\partial x}$, the time derivative of $\langle p_x, t \rangle$ is

$$\frac{\mathrm{d}\langle p_x, t\rangle}{\mathrm{d}t} = \frac{1}{2i\hbar m}\langle 0 \rangle + \frac{1}{i\hbar}\left\langle -i\hbar\frac{\partial V}{\partial x} \right\rangle = \left\langle -\frac{\partial V}{\partial x} \right\rangle.$$

We could then apply an analogous derivation for the coordinates y and z to get

$$\frac{\mathrm{d}\langle p_y, t\rangle}{\mathrm{d}t} = \left\langle -\frac{\partial V}{\partial y} \right\rangle \qquad \text{and} \qquad \frac{\mathrm{d}\langle p_z, t\rangle}{\mathrm{d}t} = \left\langle -\frac{\partial V}{\partial z} \right\rangle$$

Putting the x, y and z results together and combining the three position derivatives into the single gradient operator gives the Ehrenfest theorem:

$$\frac{\mathrm{d}\langle \boldsymbol{p}, t\rangle}{\mathrm{d}t} = \langle -\nabla V \rangle = \langle \boldsymbol{F} \rangle, \qquad \text{where } \boldsymbol{F}(\boldsymbol{r}) = -\nabla V(\boldsymbol{r}).$$

1.9 The Virial Theorem

State and derive the virial theorem in quantum mechanics.

- The virial theorem in quantum mechanics reads

$$\frac{\mathrm{d}\langle \boldsymbol{r} \cdot \boldsymbol{p}\rangle}{\mathrm{d}t} = 2\langle T \rangle + \langle \boldsymbol{r} \cdot \boldsymbol{F} \rangle, \qquad \text{where } T = \frac{p^2}{2m}.$$

For a stationary state, in which $\frac{\mathrm{d}\langle \boldsymbol{r} \cdot \boldsymbol{p}\rangle}{\mathrm{d}t} = 0$, the virial theorem simplifies to

$$2\langle T \rangle = -\langle \boldsymbol{r} \cdot \boldsymbol{F} \rangle,$$

which is reminiscent of the classical result

$$2\langle T \rangle = -\left\langle \sum_i^N \boldsymbol{F}_i \cdot \boldsymbol{r}_i \right\rangle$$

for a system of N particles.

Derivation: The Quantum Virial Theorem

- We derive the virial theorem in quantum mechanics by finding the time derivative of the expectation value $\langle \boldsymbol{r} \cdot \boldsymbol{p}, t \rangle$. Using the general result for the time derivative of an expectation value and recognizing $\frac{\partial \boldsymbol{r} \cdot \boldsymbol{p}}{\partial t} = 0$, we have

$$\frac{\mathrm{d}\langle \boldsymbol{r} \cdot \boldsymbol{p}\rangle}{\mathrm{d}t} = \frac{1}{i\hbar}\langle [\boldsymbol{r} \cdot \boldsymbol{p}, H] \rangle.$$

We evaluate the commutator by components, starting with

$$[x_\alpha p_\alpha, H] = \left[x_\alpha p_\alpha, \frac{p_\alpha^2}{2m} + V\right] = \frac{x_\alpha}{2m}[p_\alpha, p_\alpha^2] + [x_\alpha, p_\alpha^2]\frac{p_\alpha}{2m} + x_\alpha[p_\alpha, V] + [x_\alpha, V]p_\alpha.$$

We use the results $[p_\alpha, p_\alpha^2] = [x_\alpha, V] = 0$ and expand $[x_\alpha, p_\alpha^2]$ to get

$$[x_\alpha p_\alpha, H] = \frac{p_\alpha}{2m}[x_\alpha, p_\alpha]p_\alpha + [x_\alpha, p_\alpha]\frac{p_\alpha^2}{2m} + x[p_\alpha, V].$$

Reusing the earlier results $[x_\alpha, p_\alpha] = i\hbar$ and $[p_\alpha, V] = -i\hbar\frac{\partial V}{\partial x_\alpha}$ gives

$$[x_\alpha p_\alpha, H] = 2i\hbar\frac{p_\alpha^2}{2m} - i\hbar x_\alpha\frac{\partial V}{\partial x_\alpha}.$$

- If we substitute the above result into the time derivative of $\langle \boldsymbol{r} \cdot \boldsymbol{p} \rangle$, write the components in vector form, and use $\boldsymbol{F} = -\nabla V$, we get the virial theorem:

$$\frac{\mathrm{d}\langle \boldsymbol{r} \cdot \boldsymbol{p} \rangle}{\mathrm{d}t} = 2\frac{\langle p^2 \rangle}{2m} + \langle \boldsymbol{r} \cdot \boldsymbol{F} \rangle = 2\langle T \rangle + \langle \boldsymbol{r} \cdot \boldsymbol{F} \rangle$$

where we have defined the kinetic energy operator $T = \frac{p^2}{2m}$. For a stationary state in which $\frac{\mathrm{d}\langle \boldsymbol{r} \cdot \boldsymbol{p} \rangle}{\mathrm{d}t} = 0$, the virial theorem simplifies to the above quoted expression

$$2\langle T \rangle = -\langle \boldsymbol{r} \cdot \boldsymbol{F} \rangle.$$

2 The Dirac Formalism

2.1 The Coopenhagen Interpretation

State the postulates of the Coopenhagen interpretation of quantum mechanics.

1. A quantum system is described by a state vector $|\psi\rangle$ in a function Hilbert space.

2. Every physically observable quantity is associated with a Hermitian operator.

3. The expectation value of an observable with operator $\mathcal{O}$ for a system in the state $|\psi\rangle$ is $\langle\psi|\mathcal{O}|\psi\rangle$.

4. The time evolution of a state $|\psi\rangle$ is determined by the Schrödinger equation

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi\rangle = H|\psi\rangle\,,$$

   where H is the Hamiltonian operator.

5. When measuring an observable corresponding to the operator $\mathcal{O}$, the result of a single measurement is an eigenvalue of $\mathcal{O}$ (e.g. the eigenvalue $\lambda \in \mathbb{R}$). The probability of this measurement result is $|\langle\lambda|\psi\rangle|^2$, where $|\lambda\rangle$ is $\mathcal{O}$'s eigenstate corresponding to the eigenvalue λ. After a measurement, the system's wavefunction "collapses" into the state $|\lambda\rangle$.

2.2 Braket Notation and Hilbert Space

Explain Dirac braket notation and its relationship to Hilbert spaces. State the definition and notation of the inner product in the Hilbert space of wave functions, along with some of the inner product's most important properties.

2.2.1 Braket Notation: Ket

- In Dirac braket notation, a wavefunction Ψ describing a quantum system is generalized to an abstract vector in a Hilbert space of all possible wavefunctions. When considered as an element of a Hilbert space, the wave function is written $|\Psi\rangle$ and is called a *ket*.

  As a ket, the wavefunction is not tied to a particular choice of coordinates, but is viewed more generally as a state encoding a quantum system. In this sense, $|\Psi\rangle$ is more general than a coordinate representation $\Psi(x, y, z, t)$, which is tied to Cartesian coordinates.

- In Dirac notation, a set of eigenfunctions $\{\psi_n\}$ of a quantum operator is usually written either in terms of the index or the associated eigenvalue. Some examples:

  - A generic eigenfunction ψ_n with index n would be written $|n\rangle$, and a hypothetical eigenfuncion ψ_2 would be written $|2\rangle$.
  - A momentum eigenfunction ψ_p with eigenvalue p_0 might be written $|p_0\rangle$.

- For the purposes of this course, we will take our Hilbert space to be L^2, i.e. the space of all complex functions $\psi : \mathbb{R}^3 \to \mathbb{C}$ for which

$$\int_V |\psi|^2 \, \mathrm{d}^3\boldsymbol{r} < \infty.$$

In this case, the inner product $\langle\phi|\psi\rangle$ of two wave functions $\phi,\psi \in L^2$ is

$$\langle\phi|\psi\rangle \equiv \int_V \psi^*(\boldsymbol{r})\psi(\boldsymbol{r})\,\mathrm{d}^3\boldsymbol{r}.$$

The orthonormality relation for a set of orthonormal eigenstates $\{|\psi_n\rangle\}$ is written

$$\langle m|n\rangle = \delta_{mn}.$$

- Some important properties of the inner product include

    - $\langle\lambda\psi + \mu\chi|\phi\rangle = \lambda^* \langle\psi|\phi\rangle + \mu^* \langle\chi|\phi\rangle$ where $\lambda,\mu \in \mathbb{C}$ are scalars.
    - $\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*$
    - $\langle\psi|\psi\rangle \geq 0 \qquad$ and $\qquad \langle\psi|\psi\rangle = 0 \iff \psi \equiv 0$
    - $|\langle\phi|\psi\rangle|^2 \leq \langle\phi|\phi\rangle\langle\psi|\psi\rangle$.

2.2.2 Braket Notation: Bra

- In braket notation, a *bra*, written $\langle\psi|$, represents an element of the dual space of linear functionals $f : L^2 \to \mathbb{C}$ mapping wave functions in the Hilbert space L^2 to scalars in $\mathbb{C}$.

- Mathematical background: By the Riesz representation theorem, for each linear functional $f : L^2 \to \mathbb{C}$ there exists an associated vector $|\phi_f\rangle \in L^2$ for which

$$f\,|\psi\rangle = \langle\phi_f|\psi\rangle \equiv \int_V \phi_f^*\psi\,\mathrm{d}^3\boldsymbol{r} \quad \text{for all } \psi \in L^2.$$

Because every linear functional f is associated with a vector $|\phi_f\rangle$, we can interpret the action of a linear functional f on a wavefunction $|\psi\rangle$ as the expression

$$f\,|\psi\rangle = \int_V \phi_f^*\psi\,\mathrm{d}^3\boldsymbol{r} \in \mathbb{C}.$$

In braket notation, the action of the linear functional f, associated with a vector $|\phi_f\rangle$ by the Riesz representation theorem, is written $\langle\phi_f|$.

- In terms of bras and kets, the action of a linear functional represented by $\langle\phi_f|$ on a vector represented by $|\psi\rangle$ is written

$$f\,|\psi\rangle \equiv \langle\phi_f||\psi\rangle = \langle\phi_f|\psi\rangle \in \mathbb{C}.$$

Technicality: Although $\langle\phi_f||\psi\rangle$ and $\langle\phi_f|\psi\rangle$ are numerically equal, they technically denote different things:

    - $\langle\phi_f||\psi\rangle$ represents the action of a linear functional f on the vector in L^2,
    - and the result is the scalar product $\langle\phi_f|\psi\rangle \in \mathbb{C}$.

2.3 Expansion in an Orthonormal Basis

State and discuss how a wavefunction and an operator are expanded in a given orthonormal basis. How is the identity operator written in terms of an orthonormal basis? How is an operator equation written in an orthonormal basis? Provide a matrix interpretation.

- An arbitrary state $|\psi\rangle$ is written in the orthonormal basis $\{|n\rangle\}$ as

$$|\psi\rangle = \sum_n c_n |n\rangle, \qquad \text{where } c_n = \langle n|\psi\rangle.$$

- In terms of an arbitrary orthonormal basis $\{|n\rangle\}$, the identity operator is written

$$\mathrm{I} = \sum_n |n\rangle\langle n|.$$

  Interpretation: The identity operator is written as a sum of projections onto the individual basis vectors, i.e.

$$\mathrm{I} = \sum_n P_n,$$

  where $P_n = |n\rangle\langle n|$ is the projection operator onto the basis vector $|n\rangle$.

- An operotor $\mathcal{O}$ is written in the orthormal basis $\{|n\rangle\}$ as

$$\mathcal{O} = \sum_{mn} |m\rangle\, \mathcal{O}_{mn}\, \langle n|, \qquad \text{where } \mathcal{O}_{mn} = \langle m|\mathcal{O}|n\rangle \in \mathbb{C}.$$

  In other words, an operator $\mathcal{O}$ can be represented in an arbitrary orthonormal basis $\{|n\rangle\}$ in terms of a matrix $\mathcal{O}_{mn}$ with matrix elements

$$\mathcal{O}_{mn} = \langle m|\mathcal{O}|n\rangle \equiv \int_V \psi_m^* \mathcal{O} \psi_n \, \mathrm{d}^3\boldsymbol{r}.$$

- The operator equation $\mathcal{O}|\psi\rangle = |\varphi\rangle$, where $\mathcal{O}$ acts on $|\psi\rangle$ to produce $|\varphi\rangle$, and where $|\psi\rangle$ is written in the orthonormal basis $\{|n\rangle\}$ as

$$|\psi\rangle = \sum_n c_n |n\rangle,$$

  can be written in the form

$$|\varphi\rangle = \sum_m d_m |m\rangle, \qquad \text{where } d_m = \sum_n \mathcal{O}_{mn} c_n.$$

- In matrix form, the operator equation $\mathcal{O}|\psi\rangle = |\varphi\rangle$ reads

$$\boldsymbol{\mathcal{O}} \boldsymbol{c} = \boldsymbol{d},$$

  where the (potentially infinite) dimensions of the matrix $\boldsymbol{\mathcal{O}}$ and vectors $\boldsymbol{c}$ and $\boldsymbol{d}$ are determined by the number of basis vectors in the basis $\{|n\rangle\}$, and the matrix elements $\mathcal{O}_{mn}$ are given with the usual formula $\mathcal{O}_{mn} = \langle m|\mathcal{O}|n\rangle$.

- If we expand an operator in a orthonormal basis consisting of the operator's eigenstates, the operator's matrix $\boldsymbol{O}$ is diagonal, with matrix elements

$$\mathcal{O}_{mn} = \langle m|\mathcal{O}|n\rangle = \lambda_n \delta_{mn}.$$

2.3.1 Discussion: Expanding a Wave Function in a Basis

Consider an orthonormal basis $\{|n\rangle\}$ consisting of the eigenstates $|n\rangle$ of some operator and spanning the Hilbert space of possible wave functions.

- Every such basis $\{|n\rangle\}$ has a corresponding basis $\{\langle n|\}$ of the Hilbert space's dual space of linear functionals.

- In Dirac notation, the expansion of a state $|\psi\rangle$ in a basis $\{|n\rangle\}$ takes the general form

$$|\psi\rangle = \sum_n c_n |n\rangle$$

  We find the coefficients c_n by acting on the basis expansion with a generic eigenstate $\langle m|$ and applying the basis' orthonormality identity $\langle n|m\rangle = \delta_{nm}$ to get

$$\langle m|n\rangle = \sum_n c_n \langle m|n\rangle = \sum_n c_n \delta_{mn} = c_m.$$

  The result is $c_m = \langle m|\psi\rangle$. We conventionally switch indices from m to n to get

$$c_n = \langle n|\psi\rangle \qquad \text{and} \qquad |\psi\rangle = \sum_n \langle n|\psi\rangle |n\rangle.$$

2.3.2 Derivation: Identity Operator in an Orthonormal Basis

- As quoted above, the identity operator in the basis $\{|n\rangle\}$ reads

$$\mathrm{I} = \sum_n |n\rangle \langle n|.$$

  To derive this expression, we begin with the wave function expansion equation

$$|\psi\rangle = \sum_n c_n |n\rangle.$$

- Since $\langle n|\psi\rangle$ is a scalar, we can rewrite the above expansion of $|\psi\rangle$ in the basis $\{|n\rangle\}$ and apply $\langle n|\psi\rangle = \langle n|\psi\rangle$ to get

$$|\psi\rangle = \sum_n \langle n|\psi\rangle |n\rangle = \sum_n |n\rangle \langle n|\psi\rangle = \sum_n |n\rangle \langle n||\psi\rangle = \left(\sum_n |n\rangle \langle n|\right)|\psi\rangle.$$

  Comparing the first and last term gives an important identity:

$$|\psi\rangle = \left(\sum_n |n\rangle \langle n|\right)|\psi\rangle \implies \sum_n |n\rangle \langle n| = \mathrm{I}.$$

  where I is the identity operator.

2.3.3 Derivation: Writing an Operator in an Orthonormal Basis

- Using the previous identity for the identity operator we rewrite $\mathcal{O}\ket{\psi}$ to get

$$\mathcal{O}\ket{\psi} \equiv (\mathrm{I}\,\mathcal{O}\,\mathrm{I})\ket{\psi} = \left(\sum_m \ket{m}\bra{m}\right)\mathcal{O}\left(\sum_n \ket{n}\bra{n}\right)\ket{\psi}$$
$$= \sum_m \ket{m}\bra{m}\mathcal{O}\sum_n \ket{n}\braket{n|\psi}.$$

- Next, we introduce the *matrix element* $\mathcal{O}_{mn}$, defined as

$$\mathcal{O}_{mn} = \braket{m|\mathcal{O}|n} \equiv \int_V \psi_m^* \mathcal{O}\psi_n \,\mathrm{d}^3\boldsymbol{r} \in \mathbb{C}.$$

In terms of this matrix element, we can then write $\mathcal{O}$ in the basis $\{\ket{n}\}$ as

$$\mathcal{O}\ket{\psi} = \sum_m \ket{m}\bra{m}\mathcal{O}\sum_n \ket{n}\braket{n|\psi}$$
$$= \sum_{mn} \ket{m}\mathcal{O}_{mn}\braket{n|\psi}.$$

Comparing the first and last term gives us the desired expression

$$\mathcal{O} = \sum_{mn} \ket{m}\mathcal{O}_{mn}\bra{n}.$$

2.3.4 Derivation: Operator Equation in an Orthonormal Basis

- As quoted above, for a wave function $\ket{\psi}$ with the basis expansion

$$\ket{\psi} = \sum_n c_n \ket{n} = \sum_n \braket{n|\psi}\ket{n},$$

the operator equation $\mathcal{O}\ket{\psi} = \ket{\varphi}$ is written in the basis $\{\ket{n}\}$ as

$$\ket{\varphi} = \sum_m d_m \ket{m}, \qquad \text{where } d_m = \sum_n \mathcal{O}_{mn}c_n.$$

- To derive this expression, we begin with the operator equation $\mathcal{O}\ket{\psi} = \ket{\varphi}$, write $\mathcal{O}$ in the basis $\{\ket{n}\}$ and apply $c_n = \braket{n|\psi}$ to get

$$\mathcal{O}\ket{\psi} \equiv \sum_{mn} \ket{m}\mathcal{O}_{mn}\braket{n|\psi} = \sum_{mn} \ket{m}\mathcal{O}_{mn}c_n$$
$$= \sum_m \left(\sum_n \mathcal{O}_{mn}c_n\right)\ket{m}$$
$$\equiv \sum_m d_m \ket{m}$$
$$= \ket{\varphi}.$$

The last equality is the desired result, i.e.

$$\ket{\varphi} = \sum_m d_m \ket{m}, \qquad \text{where } d_m = \sum_n \mathcal{O}_{mn}c_n$$

2.4 Hermitian Operators

What is the definition of a Hermitian and anti-Hermitian operator? Discuss the important properties of Hermitian operators.

- An operator $\mathcal{O}$ is symmetric, also called Hermitian, if for all $\phi, \psi \in L^2$

$$\langle \phi | \mathcal{O}\psi \rangle = \langle \mathcal{O}\phi | \psi \rangle .$$

  Similarly, the operator $\mathcal{O}$ is antisymmetric, or anti-Hermitian, if

$$\langle \phi | \mathcal{O}\psi \rangle = - \langle \mathcal{O}\phi | \psi \rangle .$$

- The expectation values of Hermitian operators are real.

  Derivation: We begin with the definition $\langle \psi | \mathcal{O}\psi \rangle = \langle \mathcal{O}\psi | \psi \rangle$, followed by the general identity $\langle \psi | \mathcal{O}\psi \rangle = \langle \mathcal{O}\psi | \psi \rangle^*$. The result is

$$\langle \mathcal{O} \rangle \equiv \langle \psi | \mathcal{O}\psi \rangle = \langle \mathcal{O}\psi | \psi \rangle = \langle \psi | \mathcal{O}\psi \rangle^* = \langle \mathcal{O} \rangle^* .$$

  The equality $\langle \mathcal{O} \rangle = \langle \mathcal{O} \rangle^*$ implies $\langle \mathcal{O} \rangle$ is real.

- The expectation value of a squared Hermitian operator is positive, i.e.

$$\langle \mathcal{O}^2 \rangle = \langle \psi | \mathcal{O}^2 | \psi \rangle = \langle \mathcal{O}\psi | \mathcal{O}\psi \rangle \geq 0.$$

  The last equality follows from the inner product identity $\langle \psi | \psi \rangle \geq 0$ for all ψ.

- The square $\mathcal{O}^2$ of a Hermitian operator is positive definite.

  Derivation: We apply the identity $\langle \mathcal{O}^2 \rangle \geq 0$ to the eigenvalue equation $\mathcal{O}^2 | \psi_n \rangle = \lambda_n | \psi_n \rangle$, followed by the inner product identity $\langle \psi_n | \psi_n \rangle$. The result is

$$\langle \psi_n | \mathcal{O}^2 | \psi_n \rangle = \lambda_n \langle \psi_n | \psi_n \rangle \implies \lambda_n \geq 0.$$

  The result $\lambda_n \geq 0$ means $\mathcal{O}^2$ is positive-definite.

- The eigenvalues of a Hermitian operator are real. To show this, we start with a generic Hermitian operator with the eigenvalues relation $\mathcal{O} | \psi_n \rangle = \lambda_n | \psi_n \rangle$. We then act on both sides of the equation with $\langle \psi_n |$ and apply the eigenvalue relation to get

$$\mathcal{O} | \psi_n \rangle = \lambda_n | \psi_n \rangle \implies \langle \psi_n | \mathcal{O} | \psi_n \rangle = \lambda_n \langle \psi_n | \psi_n \rangle$$

  We then apply $\langle \psi_n | \mathcal{O} | \psi_n \rangle \in \mathbb{R}$ (expectation value of a Hermitian operator is real) and $\langle \psi_n | \psi_n \rangle = 1 \in \mathbb{R}$ (the eigenstate normalization condition) to get $\lambda_n \in \mathbb{R}$.

- For a Hermitian operator, eigenfunctions corresponding to different eigenvalues are orthogonal.

Derivation: Eigenfunctions with Different Eigenvalues are Orthogonal

- Begin by considering Hermitian operator $\mathcal{O}$ and two eigenfunctions $|a\rangle$ and $|b\rangle$, with corresponding eigenvalues λ_a and λ_b, and eigenvalue equations

$$\mathcal{O}\,|a\rangle = \lambda_a\,|a\rangle \qquad \text{and} \qquad \mathcal{O}\,|b\rangle = \lambda_b\,|b\rangle\,.$$

  We act on the two equations with $\langle b|$ and $\langle a|$, respectively, to get

$$\langle b|\mathcal{O}|a\rangle = \lambda_a\,\langle b|a\rangle \qquad \text{and} \qquad \langle a|\mathcal{O}|b\rangle = \lambda_b\,\langle a|b\rangle\,,$$

  and take the complex conjugate of the second equation and apply $\lambda_n = \lambda_n^*$ for a Hermitian operator to get

$$\lambda_b\,\langle a|b\rangle^* = \langle a|\mathcal{O}|b\rangle^*\,.$$

- The rest is just playing around with the inner product identity $\langle\psi|\phi\rangle = \langle\phi|\psi\rangle^*$, the Hermitian identity $\langle a|\mathcal{O}b\rangle = \langle\mathcal{O}a|b\rangle$, and the eigenvalue equations for $|a\rangle$ and $|b\rangle$.

$$\lambda_b\,\langle a|b\rangle^* = \langle a|\mathcal{O}|b\rangle^* \equiv \langle a|\mathcal{O}b\rangle^* = \langle\mathcal{O}b|a\rangle = \langle b|\mathcal{O}a\rangle = \lambda_a\,\langle b|a\rangle\,.$$

  The first and last term imply $\lambda_b\,\langle a|b\rangle^* = \lambda_a\,\langle b|a\rangle$, and applying $\langle a|b\rangle^* = \langle b|a\rangle$ gives

$$\lambda_b\,\langle b|a\rangle = \lambda_a\,\langle b|a\rangle \implies (\lambda_b - \lambda_a)\,\langle b|a\rangle = 0,$$

  which implies $\langle b|a\rangle = \langle a|b\rangle = 0$ if $\lambda_a \neq \lambda_b$.

2.5 Adjoint and Self-Adjoint Operators

What is the definition of an adjoint operator? Discuss some the important properties of adjoint operators. How is an adjoint operator written in an orthonormal basis?

What is the definition of a self-adjoint operator? Explain the relationship between self-adjoint and Hermitian operators.

2.5.1 Adjoint Operators and Their Properties

- The adjoint $\mathcal{O}^\dagger$ of an operator $\mathcal{O}$ is defined implicitly by the relationship

$$\langle\phi|\mathcal{O}\psi\rangle = \langle\mathcal{O}^\dagger\phi|\psi\rangle.$$

- Two self-adjoint operators A and B obey

$$\left(AB\right)^\dagger = B^\dagger A^\dagger,$$

  which follows from comparing the first and last ket terms in the expression

$$\langle(AB)^\dagger\phi|\psi\rangle = \langle\phi|AB\psi\rangle = \langle A^\dagger\phi|B\psi\rangle = \langle B^\dagger A^\dagger\phi|\psi\rangle.$$

- The adjoint of an operator $\mathcal{O} = |m\rangle\langle n|$ is $\mathcal{O}^\dagger = |n\rangle\langle m|$, which follows from

$$\langle\phi|\mathcal{O}\psi\rangle = \langle\phi|m\rangle\langle n|\psi\rangle = \langle n|\psi\rangle\langle\phi|m\rangle = \big(\,\langle\psi|n\rangle\langle m|\phi\rangle\,\big)^*$$

  Similarly, $(\langle\psi|\,\mathcal{O})^\dagger = \mathcal{O}^\dagger\,|\psi\rangle$.

- Consider an operator $\mathcal{O}$ written in some generic orthonormal basis $\{|n\rangle\}$:

$$\mathcal{O} = \sum_{mn} |m\rangle\,\mathcal{O}_{mn}\,\langle n|\,.$$

The adjoint operator $\mathcal{O}^\dagger$ is then written in the basis as

$$\mathcal{O}^\dagger = \sum_{mn} |n\rangle\,\mathcal{O}_{mn}^*\,\langle m| = \sum_{mn} |m\rangle\,\mathcal{O}_{nm}^*\,\langle n|$$

The matrix elements of an operator and its adjoint are thus related by

$$\left(\mathcal{O}^\dagger\right)_{mn} = \mathcal{O}_{nm}^*$$

- Any operator $\mathcal{O}$ obeys $\left(\mathcal{O}^\dagger\right)^\dagger = \mathcal{O}$, which implies:

  - the operator $\mathcal{O} + \mathcal{O}^\dagger$ is Hermitian,
  - the operator $\mathcal{O} - \mathcal{O}^\dagger$ is anti-Hermitian,
  - if $\mathcal{O}$ is Hermitian, then $i\mathcal{O}$ is anti-Hermitian.

- The expectation values of an operator $\mathcal{O}$ obey the convenient identities

$$
\begin{array}{rclcl}
2\,\mathrm{Re}\,\langle\mathcal{O}\rangle & \equiv & 2\,\mathrm{Re}\,\langle\psi|\mathcal{O}|\psi\rangle & = & \langle\psi|(\mathcal{O} + \mathcal{O}^\dagger)|\psi\rangle \\
2i\,\mathrm{Im}\,\langle\mathcal{O}\rangle & \equiv & 2i\,\mathrm{Im}\,\langle\psi|\mathcal{O}|\psi\rangle & = & \langle\psi|(\mathcal{O} - \mathcal{O}^\dagger)|\psi\rangle\,.
\end{array}
$$

- Consider two operators A and B related by $A = \lambda B$ where $\lambda \in \mathbb{C}$ is a constant. The operators' adjoint are then related by

$$A^\dagger = \lambda^* B^\dagger,$$

which follows from comparing the first and last ket terms in the expression

$$\langle A^\dagger \phi | \psi \rangle = \langle \phi | A\psi \rangle = \langle \phi | \lambda B\psi \rangle = \langle \lambda^* B^\dagger \phi | \psi \rangle\,.$$

- The projection operator $P_n \equiv |n\rangle\,\langle n|$ equals its adjoint, i.e. $P_n = P_n^\dagger$. More so, $P_n = P_n^2$, which follows from

$$P_n^2 = |n\rangle\,\langle n|n\rangle\,\langle n| = |n\rangle\,\langle n| = P_n$$

and the normalization condition $\langle n|n\rangle = 1$.

2.5.2 Self-Adjoint Operators

- An operator $\mathcal{O}$ is self-adjoint if:

  1. Both $\mathcal{O}$ and $\mathcal{O}^\dagger$ are Hermitian, i.e.

  $$\langle\phi|\mathcal{O}\psi\rangle = \langle\mathcal{O}\phi|\psi\rangle \qquad \text{and} \qquad \langle\phi|\mathcal{O}^\dagger\psi\rangle = \langle\mathcal{O}^\dagger\phi|\psi\rangle \text{ for all } \phi, \psi \in L^2,$$

  2. $\mathcal{O}$ and $\mathcal{O}^\dagger$ act on the same domain (in our case generally the Hilbert space L^2).

A self-adjoint operator obeys $\mathcal{O} = \mathcal{O}^\dagger$, which makes sense from the name—a self-adjoint operator $\mathcal{O}$ equals its adjoint $\mathcal{O}^\dagger$.

- Every self-adjoint operator is Hermitian, but in general not every Hermitian operator is self-adjoint. However (without proof), in finite N-dimensional vector spaces $\mathbb{C}^N$ and in the Schwartz space of rapidly falling functions, Hermitian and self-adjoint operators are equivalent. Since physicists typically work only with quantities in $\mathbb{R}^N$ or functions in the Schwartz space, we tend to (incorrectly) use the terms Hermitian and self-adjoint interchangeably.

2.6 Unitary Operators

How are unitary and anti-unitary operators defined? State some of their important properties, and discuss the role of unitary operators in quantum mechanics.

Discuss unitary operators in the context of a unitary change of basis.

What is the generator of a unitary transformation? Show that all single-parameter unitary operators can be written in terms of a self-adjoint generator.

2.6.1 Unitary Operators and Their Properties

- A unitary operator is any operator U that obeys the relationship

$$UU^\dagger = U^\dagger U = \mathrm{I} \implies U^{-1} = U^\dagger,$$

where I is the identity operator. Unitary operators in quantum mechanics are analogous to orthogonal transformations in classical mechanics.

- An anti-unitary operator is an operator U that obeys the relationship

$$\langle U\phi | U\psi \rangle = \langle \phi | \psi \rangle^* = \langle \psi | \phi \rangle.$$

Anti-unitary operators are antilinear, i.e.

$$U\big(\lambda\,|\phi\rangle + \mu\,|\psi\rangle\big) = \lambda^* U\,|\phi\rangle + \mu^* U\,|\psi\rangle.$$

- Unitary and adjoint operators are related by the identity

$$|\tilde{\psi}\rangle = U\,|\psi\rangle \implies \langle \tilde{\psi}| = \langle U\psi| = \langle \psi|\,U^\dagger.$$

- Unitary operators preserve the inner product. In symbols, this means that for any unitary operator U and any two functions $|\tilde{\phi}\rangle = U\,|\phi\rangle$ and $|\tilde{\psi}\rangle = U\,|\psi\rangle$, we have

$$\langle \phi | \psi \rangle = \langle \tilde{\phi} | \tilde{\psi} \rangle.$$

The above follows from comparing the first and last terms in the expression

$$\langle \tilde{\phi} | \tilde{\psi} \rangle = \langle U\phi | U\psi \rangle = \langle UU^\dagger \phi | \psi \rangle = \langle \mathrm{I}\,\phi | \psi \rangle = \langle \phi | \psi \rangle.$$

- Unitary transformations preserve matrix elements. In symbols, this means that for any unitary operator U, any two functions $|\tilde{\phi}\rangle = U\,|\phi\rangle$ and $|\tilde{\psi}\rangle = U\,|\psi\rangle$, and any operator $\tilde{\mathcal{O}} = U\mathcal{O}U^\dagger$, we have

$$\langle \phi | \mathcal{O} | \psi \rangle = \langle \tilde{\phi} | \tilde{\mathcal{O}} | \tilde{\psi} \rangle.$$

The above follows from comparing the first and last terms in the expression

$$\langle \phi | \mathcal{O} | \psi \rangle = \langle U^{-1}\tilde{\phi} | \mathcal{O} | U^{-1}\tilde{\psi} \rangle = \langle U^{\dagger}\tilde{\phi} | \mathcal{O} | U^{\dagger}\tilde{\psi} \rangle = \langle \tilde{\phi} | U\mathcal{O}U^{\dagger} | \tilde{\psi} \rangle \equiv \langle \tilde{\phi} | \tilde{\mathcal{O}} | \tilde{\psi} \rangle \,.$$

In other words, the matrix element of $\mathcal{O}$ found with the wavefunctions $|\phi\rangle$ and $|\psi\rangle$ equal the matrix elements of the transformed operator $\tilde{\mathcal{O}} = U\mathcal{O}U^{\dagger}$ found with the transformed wavefunctions $|\tilde{\phi}\rangle$ and $|\tilde{\psi}\rangle$.

- Unitary transformations preserve eigenvalue equations:

$$\mathcal{O} \, |\psi_n\rangle = \lambda_n \, |\psi_n\rangle \implies U\mathcal{O}\,\mathrm{I} \, |\psi_n\rangle = U\mathcal{O}U^{\dagger}U \, |\psi_n\rangle = \lambda_n U \, |\psi_n\rangle$$

$$\tilde{\mathcal{O}} \, |U\psi_n\rangle = \lambda_n \, |U\psi_n\rangle$$

$$\tilde{\mathcal{O}} = |\tilde{\psi}_n\rangle = \lambda_n \, |\tilde{\psi}_n\rangle \,.$$

2.6.2   Unitary Change of Basis

- Consider an orthonormal basis $\{|\psi_n\rangle\}$ and the transformed basis $\{\, |\tilde{\psi}_n\rangle \} = \{|U\psi_n\rangle\}$ where U is a unitary operator. The operator U can then be written as

$$U = \sum_n |\tilde{\psi}_n\rangle \langle \psi | = \sum_{mn} |\psi_m\rangle \, U_{mn} \, \langle \psi_n| \,, \qquad \text{where } U_{mn} = \langle \psi_m | \tilde{\psi}_n \rangle \,.$$

  Derivation: The first equality follows from comparing the first and last terms in the expression

$$U = U\,\mathrm{I} = U \sum_n |\psi_n\rangle \langle \psi_n| = \sum_n U \, |\psi_n\rangle \langle \psi_n| = \sum_n |\tilde{\psi}_n\rangle \langle \psi_n| \,.$$

  To derive the second equality, we substitute $\mathrm{I} = \sum_m |\psi_m\rangle \langle \psi_m|$ into the first equality and define the matrix elements $U_{mn} = \langle \psi_m | \tilde{\psi}_n \rangle$ to get

$$U = \sum_n |\tilde{\psi}_n\rangle \langle \psi_n| = \sum_n \mathrm{I} \, |\tilde{\psi}_n\rangle \langle \psi_n| = \sum_n \left(\sum_m |\psi_m\rangle \langle \psi_m| \right) |\tilde{\psi}_n\rangle \langle \psi_n|$$

$$= \sum_{mn} |\psi_m\rangle \, U_{mn} \, \langle \psi_n| \,.$$

- The identity operator takes the same form in an orthonormal basis $\{|\psi_n\rangle\}$ as in the basis $\{\, |\tilde{\psi}_n\rangle \} = \{U \, |\psi_n\rangle\}$ transformed by a unitary operator U, i.e.

$$\sum_n |n\rangle \langle n| = \mathrm{I} = \sum_n |\tilde{\psi}_n\rangle \langle \tilde{\psi}_n| \,,$$

  which follows from comparing the first and last terms in the expression

$$\sum_n |n\rangle \langle n| = \mathrm{I} = UU^{\dagger} = \sum_{mn} |\tilde{\psi}_m\rangle \langle \psi_m | \psi_n \rangle \langle \tilde{\psi}_n| = \sum_n |\tilde{\psi}_n\rangle \langle \tilde{\psi}_n| \,.$$

- In a unitary change of basis $\{|\psi_n\rangle\} \to \{\, |\tilde{\psi}_n\rangle \}$, the coefficients transform as

$$|\phi\rangle = \sum_n c_n \, |\psi_n\rangle = \sum_{mn} |\tilde{\psi}_m\rangle \langle \tilde{\psi}_m | c_n | \psi_n \rangle = \sum_n d_n |\tilde{\psi}_n\rangle,$$

  where the new coefficients are

$$d_n = \sum_m U^{\dagger}_{nm} c_m.$$

2.6.3 Generators of Unitary Transformations

- If K is Hermitian, then the operator $U = e^{iK}$ is unitary, which follows from

$$e^{iK}e^{-iK} = \mathrm{I} \implies UU^\dagger = \mathrm{I}.$$

  The multiplication of the exponents relies on the Baker-Campbell-Hausdorff formula.

- Every single-parameter unitary operator $U(s)$, where $s \in \mathbb{R}$ is a real constant, can be written in the form

$$U(s) = e^{isK},$$

  where K is a self-adjoint operator called the *generator* of the unitary operator U. Often, for infinitesimal parameters $s \to 0$, we work in the first-order approximation $U(s) \approx \mathrm{I} + isK$.

Derivation:

- To derive the expression $U(s) = e^{isK}$, we first expand an arbitarary unitary operator U in powers of s for vanishingly small $s \to 0$, which gives

$$U(s) = \mathrm{I} + \frac{\mathrm{d}U}{\mathrm{d}s}s + \mathcal{O}(s^2) \qquad \text{and} \qquad U^\dagger(s) = \mathrm{I} + \frac{\mathrm{d}U^\dagger}{\mathrm{d}s} + \mathcal{O}(s^2).$$

  We then write U in terms of the identity operator, i.e.

$$UU^\dagger = \mathrm{I} + \left(\frac{\mathrm{d}U}{\mathrm{d}s} + \frac{\mathrm{d}U^\dagger}{\mathrm{d}s} \right)s + \mathcal{O}(s^2).$$

  To satisfy the unitary identity $UU^\dagger \equiv \mathrm{I}$ to first order in s, the quantity in parentheses must equal zero, i.e.

$$\frac{\mathrm{d}U^\dagger}{\mathrm{d}s} = \left(\frac{\mathrm{d}U}{\mathrm{d}s} \right)^\dagger = -\frac{\mathrm{d}U}{\mathrm{d}s}.$$

  Using the above result, we can write $\frac{\mathrm{d}U}{\mathrm{d}s}$ in the form

$$\frac{\mathrm{d}U}{\mathrm{d}s} = iK,$$

  where K is a Hermitian operator.

- Next, we divide the parameter s into $N \to \infty$ equal subintervals and apply the operator $U(\frac{s}{N})$ N times, which produces the desired result

$$U(s) = \lim_{N \to \infty} U\left(\frac{s}{N}\right) U\left(\frac{s}{N}\right) \cdots U\left(\frac{s}{N}\right) = \lim_{N \to \infty} \left(\mathrm{I} + iK\frac{s}{N}\right)^N = e^{isK}.$$

2.7 The Time Evolution Operator

Define the time evolution operator for the time evolution of stationary states and explain the motivation for its definition. How is the time evolution operator related to the Schrödinger equation?

- The time evolution operator for a system with Hamiltonian H is defined as

$$U(t) = e^{-i\frac{H}{\hbar}t}.$$

  The time evolution operator is unitary with generator H. Because U is unitary, it preserves the inner product.

- The time evolution operator gives an alternate way to find the time evolution of an initial state $|\psi(0)\rangle$ via the formula

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle.$$

Derivation: The Time Evolution Operator

- We begin by expanding an arbitrary, time-dependent state $|\psi(t)\rangle$ with the known initial state $|\psi(0)\rangle$ in a basis formed of the energy eigenstates $\{|n\rangle\}$, in the form

$$|\psi(t)\rangle = \sum_m \langle n|\psi(0)\rangle\, e^{-i\frac{E_n}{\hbar}t}\,|n\rangle.$$

- We then use the operator function identity $f(\mathcal{O})\varphi_n = f(\lambda_n)\varphi_n$ with $f(x) = e^{-i\frac{x}{\hbar}t}$, to replace the energy eigenvalues E_n with the Hamiltonian operator H, producing

$$|\psi(t)\rangle = \sum_n \langle n|\psi(0)\rangle\, e^{-i\frac{E_n}{\hbar}t}\,|n\rangle = \sum_n \langle n|\psi(0)\rangle\, e^{-i\frac{H}{\hbar}t}\,|n\rangle.$$

  Since $e^{-i\frac{H}{\hbar}t}$ is independent of n, we factor it out of the sum to get

$$\begin{aligned}
|\psi(t)\rangle &= e^{-i\frac{H}{\hbar}t}\sum_n \langle n|\psi(0)|n\rangle = e^{-i\frac{H}{\hbar}t}\left(\sum_n |n\rangle\langle n|\right)|\psi(0)\rangle \\
&= e^{-i\frac{H}{\hbar}t}\,\mathrm{I}\,|\psi(0)\rangle \\
&\equiv U(t)|\psi(0)\rangle,
\end{aligned}$$

  where we have defined the time evolution operator $U(t) \equiv e^{-i\frac{H}{\hbar}t}$.

- Applying $U(t)$ to an infinitesimal time step $\mathrm{d}t$ in the evolution of a wavefunction $|\psi\rangle$ gives

$$|\delta\psi\rangle = |\psi(t+\mathrm{d}t)\rangle - |\psi(t)\rangle = -i\frac{H}{\hbar}\,\mathrm{d}t\,|\psi(t)\rangle$$

  "Dividing" by $\mathrm{d}t$ and rearranging produces the Schrödinger equation

$$i\hbar\frac{|\psi(t+\mathrm{d}t)\rangle - |\psi(t)\rangle}{\mathrm{d}t} = i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H|\psi(t)\rangle$$

2.8   The Position and Momentum Representations

What are the momentum eigenfunctions? How are they normalized?

Explain the position- and momentum-domain representations of quantum mechanics and provide a brief background of the relevant mathematics.

State and derive how differentiation in x space is related to multiplication in p space, and vice versa.

- The momentum eigenfunctions are solutions to the momentum eigenvalue equation

$$\hat{p}\,|\varphi_p\rangle = p\,|\varphi_p\rangle \qquad \text{or} \qquad -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\varphi_p = p\varphi_p,$$

  where p is the momentum eigenvalue corresponding to the eigenfunction φ_p. Up to a constant factor, the momentum eigenstates are plane waves of the form

$$\varphi_p(x) = Ce^{i\frac{p}{\hbar}x}.$$

- The momentum eigenfunctions are normalized with the Dirac normalization

$$\langle \varphi_{p'}|\varphi_p\rangle = \delta(p - p'),$$

  in terms of which the momentum eigenfunctions read

$$\varphi_p(x) = \frac{1}{\sqrt{2\pi\hbar}}e^{i\frac{p}{\hbar}x}.$$

- The position and momentum representations are two equivalent ways to write the same generic wavefunction ψ, once in position space x and once in momentum space p. The two representations are related by the Fourier transform and read

$$\psi(x) = \int_{-\infty}^{\infty} \widehat{\psi}(p)\varphi_p(x)\,\mathrm{d}p, \quad \text{and} \quad \widehat{\psi}(p) = \int_{-\infty}^{\infty} \psi(x)\varphi_p^*\,\mathrm{d}x,$$

  where $\varphi_p(x)$ is the momentum eigenfunction with eigenvalue p.

- The relevant mathematics is described in the subsubsection immediately below.

- The action of the momentum operator $\hat{p}$ (i.e. differentiation by x) on the x-domain wave function $\psi(x)$ is equivalent to multiplication of $\widehat{\psi}(p)$ by the eigenvalue p in the momentum domain, encoded by the formula

$$(-i\hbar)^n\frac{\mathrm{d}^n}{\mathrm{d}x^n}\psi(x) \iff p^n\widehat{\psi}(p).$$

  Symmetrically, acting on the x-domain wave function with the operator $\hat{x}$ (i.e. multiplication by x) corresponds differentiation of $\widehat{\psi}(p)$ by p:

$$x^n\psi(x) \iff (i\hbar)^n\frac{\mathrm{d}^n}{\mathrm{d}p^n}\widehat{\psi}(p).$$

2.8.1 Quick Review of the Relevant Mathematics

- A function $f(x)$ and its Fourier transform to k space, where $\hbar k = p$, are

$$f(x) = \int_{-\infty}^{\infty} \widehat{f}(k)e^{ikx}\,\mathrm{d}k \qquad \text{and} \qquad \widehat{f}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-ikx}\,\mathrm{d}x.$$

  Next, we substitue the Fourier transform $\widehat{f}(k)$ into the expression for $f(x)$ and switch the order of integration to get

$$\begin{aligned}
f(x) &= \int_{-\infty}^{\infty}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} f(x')e^{-ikx'}\,\mathrm{d}x'\right]e^{ikx}\,\mathrm{d}k \\
&= \int_{-\infty}^{\infty}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ik(x'-x)}\,\mathrm{d}k\right]f(x')\,\mathrm{d}x'.
\end{aligned}$$

- For shorthand, we then define the Dirac delta function as

$$\delta(x - x') \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ik(x'-x)} \, \mathrm{d}k,$$

which simplifies the expression for $f(x)$ to

$$f(x) = \int_{-\infty}^{\infty} \delta(x' - x) f(x') \, \mathrm{d}x'.$$

Keep in mind that the delta function at this point is just a placeholder for an integral. By itself it is not manifestly convergent—it is just shorthand. In practice, it will always be multiplied by a sufficiently rapidly-falling function, and the resulting integral will then converge.

- Next, setting $x' = 0$ in the expression for $\delta(x - x')$ gives us the identity

$$\delta(x) = \frac{1}{2\pi} \int e^{ikx} \, \mathrm{d}k.$$

Finally, we note an important delta function property, namely

$$\delta(ax) = \frac{1}{|a|} \delta(x).$$

2.8.2   Discussion: Eigenvalues and Eigenstates of the Momentum Operator

- The eigenvalue equation for the momentum operator $\hat{p}$ reads

$$\hat{p} \, |\varphi_p\rangle = p \, |\varphi_p\rangle \qquad \text{or} \qquad -i\hbar \frac{\mathrm{d}}{\mathrm{d}x} \varphi_p = p \varphi_p,$$

where $p \in \mathbb{R}$ is a momentum eigenvalue. Note that the equation itself permits $p \in \mathbb{C}$, but we require real momentum eigenvalues for normalization and on a physical basis.

- The momentum eigenfunctions are plane waves, which follows from solving the momentum eigenvalue equation:

$$\hat{p}\varphi_p = -i\hbar \frac{\mathrm{d}}{\mathrm{d}x} \varphi_p(x) = p\varphi_p(x) \implies \varphi_p(x) = C e^{i\frac{p}{\hbar}x}.$$

- Momentum eigenfunctions are not normalizable in the usual sense, i.e.

$$\int_{-\infty}^{\infty} |\varphi_p|^2 \, \mathrm{d}x \equiv 1,$$

since the integral of the plane wave would diverge. Thus, momentum eigenfunctions are not elements of the Hilbert space L^2, and are called pseudo-vectors.

Instead, we normalize plane wave momentum eigenfunctions with the Dirac delta function, via

$$\begin{aligned}
\langle \varphi_{p'} | \varphi_p \rangle = \int_{-\infty}^{\infty} \varphi_{p'}^*(x) \varphi_p(x) \, \mathrm{d}x &= C^2 \int_{-\infty}^{\infty} e^{-i\frac{x}{\hbar}(p'-p)} \, \mathrm{d}x \\
&= 2\pi C^2 \delta \left[\tfrac{1}{\hbar}(p - p') \right] \\
&= 2\pi \hbar C^2 \delta(p - p'),
\end{aligned}$$

where the last line uses $\delta(ax) = \frac{1}{|a|}\delta(x)$. The relationship $\langle \varphi_{p'} | \varphi_p \rangle = 2\pi\hbar C^2 \delta(p - p')$ motivates a normalization constant $C = \frac{1}{\sqrt{2\pi\hbar}}$, which leads to

$$\varphi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{i\frac{p}{\hbar}x} \qquad \text{and} \qquad \langle \varphi_{p'} | \varphi_p \rangle = \delta(p - p').$$

This plane wave normalization convention is called Dirac normalization.

- Finally, note that the momentum eigenstates $|\varphi_p\rangle$ are simultaneously eigenstates of the stationary Schrödinger equation, in which case the eigenvalue equation reads

$$\frac{\hat{p}^2}{2m} |\varphi_p\rangle = E |\varphi_p\rangle, \qquad E = \frac{p^2}{2m}.$$

Because the momentum eigenstates $|\varphi_p\rangle$ solve the Schrödinger equation, they form a convenient basis in which to find the time evolution of an arbitrary state $|t\rangle$.

2.8.3   Expansion in the Momentum Eigenbasis

- Since the momentum operator has a continuous spectrum of eigenvalues $p \in \mathbb{R}$, the corresponding momentum eigenbasis $\{\varphi_p\}$ is infinite-dimensional. Because the eigenvalues and basis functions are continuously spaced, we expand an arbitrary function $\psi(x)$ in the momentum eigenbasis using an integral instead of a sum. The expansion reads

$$\psi(x) = \int_{-\infty}^{\infty} \widehat{\psi}(p) \varphi_p(x) \, \mathrm{d}p, \quad \text{where} \quad \widehat{\psi}(p) = \int_{-\infty}^{\infty} \psi(x) \varphi_p^* \, \mathrm{d}x.$$

Note that $\widehat{\psi}(p)$ is a Fourier transform of the wave function $\psi(x)$ from the position domain to the momentum domain. The usual coefficient $\frac{1}{2\pi}$ is accounted for in the plane waves' normalization constant.

- The above two relationships are important, so I'll write them again:

$$\psi(x) = \int_{-\infty}^{\infty} \widehat{\psi}(p) \varphi_p(x) \, \mathrm{d}p, \quad \text{and} \quad \widehat{\psi}(p) = \int_{-\infty}^{\infty} \psi(x) \varphi_p^* \, \mathrm{d}x.$$

These two expressions are called the position and momentum representations of ψ, respectively. One uses x space and one uses p space, and they provide physically equivalent ways of analyzing the wavefunction ψ. Note that as long as one of $\psi(x)$ or $\widehat{\psi}(p)$ is known, we can always find the other.

- The wavefunction ψ is normalized by the Parseval equality, which reads

$$\int_{-\infty}^{\infty} |\psi(x)|^2 \, \mathrm{d}x = \int_{-\infty}^{\infty} \left|\widehat{\psi}(p)\right|^2 \, \mathrm{d}p = 1$$

and rests on the Dirac normalization of the momentum eigenstates used to write either $\psi(x)$ or $\widehat{\psi}(p)$. The Parseval equation guarantees that both the position and momentum representations are equal and normalized, as they must be.

Interpretation: the Parseval equation can be thought of as a continuous analog of the discrete normalization condition

$$\sum_n = |\langle \varphi_n | \psi \rangle|^2 = 1.$$

- The momentum eigenfunctions φ_p are related to the delta function by the following completeness relation:

$$\int_{-\infty}^{\infty} \varphi_p^*(x')\varphi_p(x)\,\mathrm{d}p = \delta(x' - x),$$

  where the integral represents a "sum" over all possible momentum eigenfunctions.

- Finally, note that when a particle occurs in an external potential with $V(x) \neq 0$, the momentum eigenstates $|\varphi_p\rangle$ are no longer stationary states of the stationary Schrödinger equation. However, even for $V(x) \neq 0$, the momentum eigenstates still form a valid basis in which to expand an arbitrary wave function $\psi(x)$.

2.8.4   Differentiation and Multiplication in x and p Space

- We derive the first differentiation-multiplication identity via

$$\hat{p}\psi(x) = -i\hbar \frac{\mathrm{d}}{\mathrm{d}x}\psi(x) = \int_{-\infty}^{\infty} \widehat{\psi}(p)\left(-i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\varphi_p(x)\right)\mathrm{d}p = \int_{-\infty}^{\infty}\left(p\widehat{\psi}(p)\right)\varphi_p(x)\,\mathrm{d}p.$$

  Without proof, this relationship generalizes to higher-order derivatives according to

$$(-i\hbar)^n \frac{\mathrm{d}^n}{\mathrm{d}x^n}\psi(x) \iff p^n\widehat{\psi}(p).$$

- We derive the second differentiation-multiplication identity via

$$\hat{x}\psi(x) = x\psi(x) = \int_{-\infty}^{\infty} \widehat{\psi}(p)x\varphi_p(x)\,\mathrm{d}p = \int_{-\infty}^{\infty} \widehat{\psi}(p)\left(-i\hbar\frac{\mathrm{d}}{\mathrm{d}p}\varphi_p(x)\right)\mathrm{d}p$$

$$= \int_{-\infty}^{\infty}\left(i\hbar\frac{\mathrm{d}}{\mathrm{d}p}\widehat{\psi}(p)\right)\varphi_p(x)\,\mathrm{d}p.$$

  Again without proof, this result generalizes to arbitrary order n according to

$$x^n\psi(x) \iff (i\hbar)^n\frac{\mathrm{d}^n}{\mathrm{d}p^n}\widehat{\psi}(p).$$

2.8.5   Discussion: Eigenvalues and Eigenstates of the Position Operator

- We now consider the position operator $\hat{x}$, for which the eigenvalue relation reads

$$\hat{x}\psi_0(x) = x_0\psi_0(x),$$

  where $x_0 \in \mathbb{R}$ is the position eigenvalue. Transferring to momentum space and using the fact that multiplication by x in x space corresponds to applying $i\hbar\frac{\mathrm{d}}{\mathrm{d}p}$ in p space produces

$$\hat{x}\psi_0(x) = x\int_{-\infty}^{\infty}\widehat{\psi}_0(p)\varphi_p(x)\,\mathrm{d}p = \int_{-\infty}^{\infty}\left(i\hbar\frac{\mathrm{d}}{\mathrm{d}p}\widehat{\psi}_0(p)\right)\varphi_p(x)\,\mathrm{d}p$$

$$= x_0\int_{-\infty}^{\infty}\widehat{\psi}_0(p)\varphi_p(x)\,\mathrm{d}p$$

  The last equality implies the position eigenfunctions ψ_0 obey the relationship

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}p}\widehat{\psi}_0(p) = x_0\widehat{\psi}_0(p) \implies \widehat{\psi}_0(p) = \frac{1}{\sqrt{2\pi\hbar}}e^{-i\frac{p}{\hbar}x_0} = \varphi_p^*(x_0).$$

  In other words, the Fourier transform of a position eigenfunction with eigenvalue x_0 is a plane wave in p space.

- Next, we substitute in $\widehat{\psi_0}(p) = \varphi_p^*(x_0)$ to the expanded eigenvalue equation to get

$$\psi_0(x) = \int_{-\infty}^{\infty} \varphi_p^*(x_0)\varphi_p(x)\,\mathrm{d}p = \delta(x - x_0).$$

Interpretation: An eigenfunction of the position operator with eigenvalue x_0 is a delta function centered at x_0, which can be physically interpreted as a localized function centered around the eigenvalue $x = x_0$.

Using $\psi_0(x) = \delta(x - x_0)$, the position eigenvalue equation $\hat{x}\psi_0(x) = x_0\psi_0(x)$ reads

$$\hat{x}\delta(x - x_0) = x_0\delta(x - x_0).$$

2.9 The Probability Amplitudes $\langle p|\psi\rangle$ and $\langle x|\psi\rangle$

State, derive, and interpret the meaning of the quantities $\langle p|\psi\rangle$ and $\langle x|\psi\rangle$.

How is a state $|\psi\rangle$ expanded in the momentum and position eigenstates in Dirac notation?

- The quantity $\langle p|\psi\rangle$ is related to a quantum state with wavefunction ψ via

$$\langle p|\psi\rangle = \widehat{\psi}(p).$$

which can be interpreted as a projection of the momentum eigenstate $|p\rangle$ onto the quantum state $|\psi\rangle$. In this probability interpretation, the quantity

$$\rho_p(p_0) = |\langle p_0|\psi\rangle|^2$$

encodes the probability of detecting a particle in the quantum state $|\psi\rangle$ in the momentum eigenstate p_0, and the associated probability is $\mathrm{d}P = \rho_p\,\mathrm{d}p$.

- The quantity $\langle x|\psi\rangle$ is related to a quantum state with wavefunction ψ via

$$\langle x|\psi\rangle = \psi(x),$$

which can be interpreted as a projection of the position eigenstate $|x\rangle$ onto the quantum state $|\psi\rangle$. In this probability interpretation, the quantity

$$\rho_x(x_0) = |\langle x_0|\psi\rangle|^2$$

encodes the probability of detecting a particle in the quantum state $|\psi\rangle$ in the position eigenstate x_0, and the associated probability is $\mathrm{d}P = \rho_x\,\mathrm{d}x$.

- A generic state $|\psi\rangle$ is expanded in the position and momentum eigenstates via

$$|p\rangle = \int_{-\infty}^{\infty} \widehat{\psi}(p)\,|p\rangle\,\mathrm{d}p \qquad \text{and} \qquad |\psi\rangle = \int_{-\infty}^{\infty} \psi(x)\,|x\rangle\,\mathrm{d}x.$$

2.9.1 Derivation: Momentum Probability Amplitude

- We begin by expanding are arbitrary state $|\psi\rangle$ in the momentum eigenbasis $\{|p\rangle\}$, which reads

$$|\psi\rangle = \int_{-\infty}^{\infty} \widehat{\psi}(p)\,|p\rangle\,\mathrm{d}p.$$

  This expression is just a braket notation version of the position representation

$$\psi(x) = \int_{-\infty}^{\infty} \widehat{\psi}(p)\varphi_p(x)\,\mathrm{d}p.$$

- We then multiply the equation for $|\psi\rangle$ by $\langle p'|$ to get

$$\langle p'|\psi\rangle = \int_{-\infty}^{\infty} \widehat{\psi}(p)\,\langle p'|p\rangle\,\mathrm{d}p = \int_{-\infty}^{\infty} \widehat{\psi}(p)\delta(p-p')\,\mathrm{d}p = \widehat{\psi}(p'),$$

  where we have used $\langle p|p'\rangle = \delta(p-p')$ and the delta function identity

$$f(x) = \int_{-\infty}^{\infty} \delta(x'-x)f(x')\,\mathrm{d}x'.$$

  The end result is the desired probability amplitude

$$\widehat{\psi}(p) = \langle p|\psi\rangle$$

2.9.2 Derivation: Position Probability Ampitude

- As for for derivation of momentum probability amplitude, we begin by expanding an arbitrary state $|\psi\rangle$ in the momentum eigenbasis via

$$|\psi\rangle = \int_{-\infty}^{\infty} \psi(x)\,|p\rangle\,\mathrm{d}x.$$

  We then multiply the equation through by $\langle x_0|$, which produces

$$\langle x_0|\psi\rangle = \int_{-\infty}^{\infty} \widehat{\psi}(p)\,\langle x_0|p\rangle\,\mathrm{d}p.$$

- Next, we use $|x_0\rangle \to \delta(x-x_0)$ and $|p\rangle \to \frac{1}{\sqrt{2\pi\hbar}}e^{i\frac{p}{\hbar}x}$ to write $\langle x_0|p\rangle$ as

$$\langle x_0|p\rangle = \int_{-\infty}^{\infty} \delta(x-x_0)\frac{1}{\sqrt{2\pi\hbar}}e^{i\frac{p}{\hbar}x}\,\mathrm{d}x = \frac{1}{\sqrt{2\pi\hbar}}e^{i\frac{p}{\hbar}x_0} = \varphi_p(x_0),$$

  which we substitute into the expresion for $\langle x_0|p\rangle$ to get

$$\langle x_0|\psi\rangle = \int_{-\infty}^{\infty} \widehat{\psi}(p)\,\langle x_0|p\rangle\,\mathrm{d}p = \int_{-\infty}^{\infty} \widehat{\psi}(p)\varphi_p(x_0)\,\mathrm{d}p = \psi(x_0),$$

  where the last line follows from the general position representation

$$\psi(x) = \int_{-\infty}^{\infty} \widehat{\psi}(p)\varphi_p(x)\,\mathrm{d}p.$$

  The end result is $\langle x_0|\psi\rangle = \psi(x_0)$, which implies the position probability amplitude

$$\psi(x) = \langle x|\psi\rangle\,.$$

- The probability amplitude $\psi(x) = \langle x|\psi \rangle$ motivates the definition for expanding an arbitrary wavefunction $|\psi \rangle$ in the position eigenbasis $\{|x\rangle\}$ as

$$|\psi \rangle = \int_{-\infty}^{\infty} \psi(x) |x\rangle \, \mathrm{d}x.$$

To show this is correct, we can multiply through by $\langle x_0|$ to recover the known identity

$$\langle x_0|\psi \rangle = \psi(x_0).$$

Derivation:

$$\langle x_0|\psi \rangle = \int_{-\infty}^{\infty} \psi(x) \langle x_0|x\rangle \, \mathrm{d}x = \int_{-\infty}^{\infty} \psi(x)\delta(x - x_0) \, \mathrm{d}x = \psi(x_0).$$

Two More Notes

- The position eigenstates are orthonormalized in terms of the delta function via $\langle x_0|x\rangle = \delta(x - x_0)$, which implies

$$|x_0\rangle = \int_{-\infty}^{\infty} \delta(x - x_0) |x\rangle \, \mathrm{d}x.$$

We can then multiply the above equation through by $\langle x_1|$, which produces

$$\int_{-\infty}^{\infty} \delta(x - x_0)\delta(x - x_1) \, \mathrm{d}x = \delta(x_0 - x_1).$$

- In terms of the position and momentum eigenstates, the identity operator is written

$$\mathrm{I} = \int_{-\infty}^{\infty} |p\rangle \langle p| \, \mathrm{d}p = \int_{-\infty}^{\infty} |x\rangle \langle x| \, \mathrm{d}x.$$

3 Examples

3.1 The Harmonic Oscillator and The Ladder Operators

Discuss the quantum harmonic oscillator. How are the ladder operators a and $a^\dagger$ defined, and how are the position, momentum and Hamiltonian operators written in terms of the ladder operators? Discuss the matrix forms of the ladder operators, position, momentum, and Hamiltonian.

- In one dimension, the quantum harmonic oscillator's Hamiltonian reads

$$H = \frac{p^2}{2m} + \frac{1}{2}kx^2 = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{1}{2}m\omega^2 x^2, \qquad \omega = \sqrt{\frac{k}{m}}.$$

- The dimensionless ladder operatrs a and $a^\dagger$ are defined as

$$
\begin{aligned}
a &= \tfrac{1}{\sqrt{2}}\left(\tfrac{x}{x_0} + x_0\tfrac{\mathrm{d}}{\mathrm{d}x}\right) &= \tfrac{1}{\sqrt{2}}\left(\tfrac{x}{x_0} + i\tfrac{p}{p_0}\right) \\
a^\dagger &= \tfrac{1}{\sqrt{2}}\left(\tfrac{x}{x_0} - x_0\tfrac{\mathrm{d}}{\mathrm{d}x}\right) &= \tfrac{1}{\sqrt{2}}\left(\tfrac{x}{x_0} - i\tfrac{p}{p_0}\right),
\end{aligned}
$$

where the characteristic length x_0 and momentum p_0 are defined as

$$x_0 = \sqrt{\frac{\hbar}{m\omega}} \qquad \text{and} \qquad x_0 p_0 = \hbar.$$

- The ladder operators obey the important commutation relation

$$\left[a, a^\dagger\right] = 1.$$

Derivation: We begin with the definitions of a and $a^\dagger$ in terms of x and p and apply $x_0 p_0 = \hbar$ followed by the canonical commutation relation $[x, p] = i\hbar$. This reads:

$$
\begin{aligned}
\left[a, a^\dagger\right] = \left(aa^\dagger - a^\dagger a\right) &= \frac{1}{2}\left[\frac{x^2}{x_0} - \frac{i}{x_0 p_0}xp + \frac{i}{x_0 p_0}px + \frac{p^2}{p_0^2}\right] \\
&\quad - \frac{1}{2}\left[\frac{x^2}{x_0} + \frac{i}{x_0 p_0}xp - \frac{i}{x_0 p_0}px + \frac{p^2}{p_0^2}\right] \\
&= -\frac{i}{\hbar}(xp - px) = -\frac{i}{\hbar}[x, p] = -\frac{i}{\hbar}(i\hbar) = 1.
\end{aligned}
$$

- In terms of the ladder operators, the position and momentum operators read

$$x = \frac{x_0}{\sqrt{2}}\left(a + a^\dagger\right) \qquad \text{and} \qquad p = \frac{p_0}{\sqrt{2}i}\left(a - a^\dagger\right),$$

while the Hamiltonian operator reads

$$H = \frac{\hbar\omega}{2}\left(aa^\dagger + a^\dagger a\right) = \hbar\omega\left(a^\dagger a + \tfrac{1}{2}\right),$$

where the last line uses the commutation relation $\left[a, a^\dagger\right] = 1$.

3.1.1 Ladder Operators in Matrix Form

- In matrix form, written in the harmonic oscillator eigenbasis $\{|n\rangle\}$, the ladder operator $a^\dagger$ reads

$$a^\dagger = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix},$$

where the matrix elements are found according to

$$a^\dagger_{mn} = \langle m | a^\dagger | n \rangle = \langle n+1 | \sqrt{n+1} | n \rangle \, \delta_{m,n+1},$$

which is derived from the equation $a^\dagger |n\rangle = \sqrt{n+1}\,|n+1\rangle$, which is itself derived in the next question.

Note that $a^\dagger$ is asymmetric and thus non-Hermitian.

- Similarly, the expressions for x and p in the harmonic oscillator eigenbasis are

$$x = \sqrt{\frac{\hbar}{2m\omega}} \begin{pmatrix} 0 & 1 & 0 & \cdots \\ 1 & 0 & \sqrt{2} & \cdots \\ 0 & \sqrt{2} & 0 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix} \quad \text{and} \quad p = \sqrt{\frac{m\hbar\omega}{2}} \begin{pmatrix} 0 & i & 0 & \cdots \\ -i & 0 & i\sqrt{2} & \cdots \\ 0 & -i\sqrt{2} & 0 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

As expected, both x and p have Hermitian matrices.

3.2 The Harmonic Oscillator: Eigenvalues and Eigenstates

State the harmonic oscillator's eigenvalues and eigenstates, and explain the derivation process using the algebraic ladder operator method. Give special attention to the ground state eigenfunction and eigenvalue.

State and derive the action of the ladder operators on the harmonic oscillator's eigenstates.

Coordinate form of eigenfunction.

- The harmonic oscillator's eigenfunctions and eigenvalues are

$$H |n\rangle = E_n |n\rangle \qquad E_n = \left(n + \tfrac{1}{2}\right)\hbar\omega \qquad \langle m|n\rangle = \delta_{mn}.$$

The ground state is a Gaussian function, with energy $E_0 = \frac{1}{2}\hbar\omega$ and wavefunction

$$\phi_0(x) = \frac{1}{\sqrt{\sqrt{\pi}x_0}} e^{-\frac{1}{2}\frac{x^2}{x_0^2}}.$$

- The action of the creation operator $a^\dagger$ on the harmonic oscillator eigenfuncions is encoded by the recursion relation

$$|n\rangle = \frac{a^\dagger}{\sqrt{n}} |n-1\rangle = \frac{\left(a^\dagger\right)^n}{\sqrt{n!}} |0\rangle \,,$$

while the analogous action of the annihilation operator a is encoded by

$$|n\rangle = \frac{a}{\sqrt{n+1}} |n+1\rangle \qquad \text{and} \qquad |0\rangle = \frac{a^n}{\sqrt{n!}} |n\rangle \,.$$

3.2.1 Ground State Solution

- We begin with the expression for the harmonic oscillator's Hamiltonian, i.e.

$$H = \hbar\omega\left(a^\dagger a + \tfrac{1}{2}\right).$$

Since H is written only in terms of $a^\dagger a$, finding the harmonic oscillator's eigenfunctions and eigenvalues reduces to finding the eigenfunctions and eigenvalues of $a^\dagger a$.

For reasons that will soon become clear, we will call $a^\dagger a$ the counting operator and denote it $\hat{n}$. The eigenvalue equation for $\hat{n} \equiv a^\dagger a$ reads

$$\hat{n}\,|\phi_n\rangle = n\,|\phi_n\rangle,$$

where n is the index of the eigenfunction ϕ_n.

- First, using the inner product identity $\langle\psi|\psi\rangle \geq 0$, we show $n \geq 0$, which follows from

$$\langle\phi_n|\hat{n}|\phi_n\rangle = \langle\phi_n|a^\dagger a|\phi_n\rangle = \langle a\phi_n|a\phi_n\rangle = n\,\langle\phi_n|\phi_n\rangle \geq 0 \implies n \geq 0.$$

- Next, we confirm $n = 0$ is a valid solution of counting operator's eigenvalue equation. For $n = 0$, using $\hat{n} = a^\dagger a$, the eigenvalue equation reads

$$a^\dagger a\,|\phi_0\rangle = 0 \cdot |\phi_0\rangle \implies a^\dagger a\,|\phi_0\rangle = 0 \implies a\,|\phi_0\rangle = 0.$$

We then transition to the x coordinate representation and use the definition of a in terms of x and $\frac{\mathrm{d}}{\mathrm{d}x}$, in which case the eigenvalue equation $a\,|\varphi_0\rangle = 0$ becomes

$$\frac{1}{\sqrt{2}}\left(\frac{x}{x_0} + x_0\frac{\mathrm{d}}{\mathrm{d}x}\right)\phi_0(x) = 0 \implies x_0\frac{\mathrm{d}}{\mathrm{d}x}\phi_0(x) = -\frac{x}{x_0}\phi_0(x).$$

The solution is the Gaussian function

$$\phi_0(x) = \frac{1}{\sqrt{\sqrt{\pi}x_0}}e^{-\frac{1}{2}\frac{x^2}{x_0^2}} \equiv \langle x|\phi_0\rangle.$$

The state $|\phi_0\rangle$ is the oscillator's ground state, with energy $E_0 = \frac{1}{2}\hbar\omega$. We can find all other solutions from the ground state solution.

3.2.2 Action of the Creation Operator $a^\dagger$

- To find excited states from the known ground state solution $|\phi_0\rangle$, we first derive the commutator relation

$$\left[\hat{n}, a^\dagger\right] = \left[a^\dagger a, a^\dagger\right] = a^\dagger\left[a, a^\dagger\right] + \left[a^\dagger, a^\dagger\right]a = a^\dagger$$

We will now use this relationship to show that $a^\dagger$ acts on a state with eigenvalue n to create a state with eigenvalue $n + 1$. To show this, we first calculate

$$\hat{n}\left[a^\dagger\,|\phi_n\rangle\right] \equiv a^\dagger a a^\dagger\,|\phi_n\rangle = a^\dagger\left(a^\dagger a + 1\right)|\phi_n\rangle = \left(a^\dagger\hat{n} + a^\dagger\right)|\phi_n\rangle$$

$$= a^\dagger n\,|\phi_n\rangle + a^\dagger\,|\phi_n\rangle = (n+1)\left[a^\dagger\,|\phi_n\rangle\right],$$

where the square brackets are added to stress that we view $a^\dagger \ket{\phi_n}$ as one state. Because the counting operator $\hat{n}$ acts on the state $a^\dagger \ket{\phi_n}$ to produce an eigenvalue $(n+1)$, the operator $a^\dagger$ must have the effect of raising $\ket{\phi_n}$'s index by one.

In symbols, the action of $a^\dagger$ on $\ket{\phi_n}$ reads

$$a^\dagger \ket{\phi_n} = c_n \ket{\phi_{n+1}},$$

where the constant c_n is to be determined.

- We find the constant c_n by applying the normalization condition $\braket{\phi_n|\phi_n} = 1$ to the state $\ket{\phi_n}$. The calculation of c_n reads

$$\begin{aligned}
\braket{c_n^* \phi_{n+1} | c_n \phi_{n+1}} &= \braket{a^\dagger \phi_n | a^\dagger \phi_n} = \braket{\phi_n | a a^\dagger \phi_n} = \braket{\phi_n | (a^\dagger a + 1)\phi_n} \\
&= \braket{\phi_n | (n+1)\phi_n} = (n+1)\braket{\phi_n | \phi_n} \\
&= (n+1),
\end{aligned}$$

where we have used the eigenvalue relation $a^\dagger a \ket{\phi_n} = n \ket{\phi_n}$ at the end of the first line. Comparing the first and last equality implies $|c_n|^2 = (n+1)$ and thus $c_n = \sqrt{n+1}$. The action of $a^\dagger$ on $\ket{\psi_n}$ is then fully summarized with

$$a^\dagger \ket{\phi_n} = \sqrt{n+1} \ket{\phi_{n+1}} \qquad \text{or} \qquad \ket{\phi_{n+1}} = \frac{a^\dagger}{\sqrt{n+1}} \ket{\phi_n}.$$

If we start with $\ket{\phi_n} = \ket{\phi_0}$, the latter expression produces to the recursive relation

$$\ket{\phi_n} = \frac{a^\dagger}{\sqrt{n}} \ket{\phi_{n-1}} = \frac{\left(a^\dagger\right)^n}{\sqrt{n!}} \ket{\phi_0}.$$

3.2.3 Action of the Annihilation Operator a

- While the creation operator $a^\dagger$ raises the index of a harmonic oscillator's eigenstate, the annihilation operator a lowers a eigenstate's index. The derivation follows the same pattern as above for $a^\dagger$: we use the commutator relation

$$[\hat{n}, a] = a^\dagger [a, a] + [a^\dagger, a]a = -a$$

to show that

$$\hat{n}[a \ket{\phi_n}] = (a\hat{n} - a)\ket{\phi_n} = (n-1)[a \ket{\phi_n}].$$

Because the counting operator acts on the state $a \ket{\phi_n}$ to produce an eigenvalue $(n-1)$, the operator a must have the effect of lowering $\ket{\phi_n}$'s index by one, i.e.

$$a \ket{\phi_n} = d_n \ket{\phi_{n-1}}.$$

- Just like for $a^\dagger$ and c_n, we find the constant d_n under that assumption that the original state $\ket{\phi_n}$ is normalized, i.e. $\braket{\phi_n|\phi_n} = 1$. The relevant calculation reads

$$\begin{aligned}
\braket{d_n^* \phi_{n-1} | d_n \phi_{n-1}} &= \braket{a\phi_n | a\phi_n} = \braket{\phi_n | a^\dagger a \phi_n} = \braket{\phi_n | n\phi_n} \\
&= n\braket{\phi_n | \phi_n} \\
&= n,
\end{aligned}$$

which implies $|d_n|^2 = n$ and thus $d_n = \sqrt{n}$. With d_n known, the action of a on an eigenstate $|\phi_n\rangle$ is then fully summarized with

$$a\,|\phi_n\rangle = \sqrt{n}\,|\phi_{n-1}\rangle \qquad \text{or} \qquad |\phi_{n-1}\rangle = \frac{a}{\sqrt{n}}\,|\phi_n\rangle.$$

The latter expression results in the recursion relations

$$|\phi_n\rangle = \frac{a}{\sqrt{n+1}}\,|\phi_{n+1}\rangle \qquad \text{and} \qquad |\phi_0\rangle = \frac{a^n}{\sqrt{n!}}\,|\phi_n\rangle.$$

- In fact, the recursive relations between the ladder operators $a^\dagger$ and a and the eigenstates $|\phi_n\rangle$, together with the known ground state $|\phi_0\rangle$, fully solve the harmonic oscillator problem. The results are

$$H\,|\phi_n\rangle = E_n\,|\phi_n\rangle \qquad E_n = \left(n + \tfrac{1}{2}\right)\hbar\omega \qquad \langle\phi_m|\phi_n\rangle = \delta_{mn}.$$

In practice, the harmonic oscillator eigenstates are written with only their index, i.e.

$$H\,|n\rangle = E_n\,|n\rangle \qquad E_n = \left(n + \tfrac{1}{2}\right)\hbar\omega \qquad \langle m|n\rangle = \delta_{mn}.$$

Some Discussion of the Solution

- In one dimension, the harmonic oscillator's energy eigenvalues E_n are nondegenerate.

We prove nondegeneracy by contradiction: assume that in addition to $|\phi_n\rangle$ there exists another linearly independent eigenstates $|\tilde{\phi}_n\rangle$ with the same energy E_n. From the recursion relation

$$|\phi_n\rangle = \frac{a^n}{\sqrt{n!}}\,|\phi_0\rangle,$$

the state $|\tilde{\phi}_n\rangle$ must obey $a^n|\tilde{\phi}_n\rangle \propto |\tilde{\phi}_0\rangle$. However, the harmonic oscillator's ground state is non-degenerate, since the earlier ground state equation, i.e.

$$x_0\frac{\mathrm{d}}{\mathrm{d}x}\phi_0(x) = -\frac{x}{x_0}\phi_0(x),$$

has only one normalized solution:

$$\langle x|\phi_0\rangle = \frac{1}{\sqrt{\sqrt{\pi}x_0}}e^{-\frac{1}{2}\frac{x^2}{x_0^2}}.$$

Because the ground state is nondegenerate and all higher states are proportional to the ground state via $a^n|\phi_n\rangle \propto |\phi_0\rangle$, all higher states are also nondegenerate.

- The harmonic oscillator's energy eigenvalues have only integer indexes $n \in \mathbb{N}$.

We again prove this by contradiction: assume there exists an energy eigenstate $|\phi_\lambda\rangle$ with index $\lambda = n + \nu$ where $\nu \in (0,1)$. Applying the counting operator to $|\phi_\lambda\rangle$ produces

$$\hat{n}\,|\phi_\lambda\rangle = \lambda\,|\phi_\lambda\rangle = (n+\nu)\,|\phi_\lambda\rangle$$

Repeatedly applying the annihilation operator a to the state $|\phi_\lambda\rangle$ and using the recursion relation $|\phi_n\rangle = \frac{a^n}{\sqrt{n!}}\,|\phi_0\rangle$ would eventually lead to a state with the index $-1 < \lambda < 0$, i.e. a negative index. This contradicts the fact that harmonic oscillator's indexes are non-negative, i.e. $n \geq 0$.

3.2.4 Eigenfunctions in the Coordinate Representation

- In the coordinate representation, the harmonic oscillators eigenfunctions are found with the generating formula

$$\langle x | \phi_n \rangle = \phi_n(x) = \frac{1}{\sqrt{2^n n!}} \left(\frac{x}{x_0} - x_0 \frac{\mathrm{d}}{\mathrm{d}x} \right)^n \phi_0(x).$$

The ground state eigenfunction with $n = 0$ is even, and the excited state eigenfunctions with $n = 1, 2, \ldots$ alternate between even and odd according to the parity of the index n.

- Perhaps more intuitively, the eigenfunctions are just the product of a Hermite polynomial and the fundamental Gaussian solution $\phi_0(x)$. In this form, the eigenfunctions are written

$$\phi_n = C_n H_n \left(\frac{x}{x_0} \right) e^{-\frac{1}{2}\frac{x^2}{x_0^2}}, \qquad \text{where } C_n = \frac{1}{\sqrt{2^n n! x_0 \sqrt{n}}}$$

and H_n is the nth Hermite polynomial and

- The characteristic width of each eigenfunction increases with the index n; the width σ_x of the nth state obeys

$$\frac{\sigma_{x_n}^2}{x_0^2} = \tfrac{1}{2}\langle n | \left(a^\dagger \right)^2 + a^\dagger a + a a^\dagger + a^2 | n \rangle = n + \tfrac{1}{2}$$

In the p-space representation, the nth eigenfunction's characteristic width is

$$\frac{\sigma_{p_n}^2}{p_0^2} = -\tfrac{1}{2}\langle n | \left(a^\dagger \right)^2 - a^\dagger a - a a^\dagger + a^2 | n \rangle = n + \tfrac{1}{2}.$$

The product $\sigma_{x_n} \sigma_{p_n}$ is thus

$$\sigma_{x_n} \sigma_{p_n} = x_0 p_0 \left(n + \tfrac{1}{2} \right) = \hbar \left(n + \tfrac{1}{2} \right).$$

Note that the ground state obeys $\sigma_{x_n} \sigma_{p_n} = \hbar/2$, which brings to mind the Heisenberg uncertainty principle.

4 Symmetries

4.1 Translational Symmetry

How is the quantum mechanical transaltion operator defined in one and three dimensions? What is the operator's generator? Which problems have translational symmetry?

Note: In this question we consider only active translations, corresponding to a translation of a wavefunction, as opposed to a translation of the coordinate system or basis vectors.

- In one dimension, the operator for translating a wavefunction by the distance s is

$$U(s) = e^{-is\frac{p}{\hbar}}.$$

  The analogous three-dimensional operator for translating a wavefunction by the distance s in the direction of the unit vector $\hat{\mathbf{n}}$ is

$$U(s\,\hat{\mathbf{n}}) = e^{-is\frac{\hat{\mathbf{n}}\cdot\boldsymbol{p}}{\hbar}} \qquad \text{or} \qquad U(\boldsymbol{s}) = e^{-i\frac{\boldsymbol{s}\cdot\boldsymbol{p}}{\hbar}}.$$

  Derivation: In one dimension, a translation of a wavefunction ψ by s reads

$$\tilde{\psi}(x) = \psi(x - s),$$

  which we writte in terms of a yet-to-be-determined translation operator $U(s)$ as

$$\tilde{\psi}(x) = U(s)\psi(x) = \psi(x - s).$$

  We find the expression for $U(s)$ with a Taylor series expansion of $\psi(x-s)$:

$$\begin{aligned} U(s)\psi(x) = \psi(x - s) &= \psi(x) - s\frac{\partial\psi(x)}{\partial x} \pm \cdots + \frac{(-s)^n}{n!}\frac{\partial^n\psi(x)}{\partial x^n} + \cdots \\ &= e^{-s\frac{\partial}{\partial x}}\psi(x) \\ &= e^{-is\frac{p}{\hbar}}\psi(x). \end{aligned}$$

  The last line motivates the definition of the translation operator as

$$U(s) = e^{-is\frac{p}{\hbar}}.$$

- In one dimension, the translation's generator is the momentum p, and the corresponding generator for a three-dimensional translation $\boldsymbol{s} = s\,\hat{\mathbf{n}}$ is $\hat{\mathbf{n}} \cdot \boldsymbol{p}$, i.e. the projection of momentum in the direction $\hat{\mathbf{n}}$.

- Like in classical mechanics, translational symmetry corresponds to conservation of (translational) momentum.

  In free space (for a globally constant potential), momentum is conserved under the condition $[\boldsymbol{p}, H] = 0$, which occurs when the Hamiltonian is invariant under translation, i.e. when

$$[U(\boldsymbol{s}), H] = 0 \text{ for all } \boldsymbol{s} \in \mathbb{R}^3$$

- In the presence of a periodic potential with period $\boldsymbol{a}$, ie. $V(\boldsymbol{r}) = V(\boldsymbol{r} + n\boldsymbol{a})$ where $n \in \mathbb{Z}$ is an integer, translational invariance holds for translations of the form $\boldsymbol{s}_n = n\boldsymbol{a}$. In this case, the wavefunction takes the form

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}}u(\boldsymbol{r})$$

  where $u(\boldsymbol{r} + \boldsymbol{a}) = u(\boldsymbol{r})$ is a periodic function.

4.2 Rotation

State and derive the quantum-mechanical rotation operator. What is the operator's generator? Which problems have rotational symmetry?

Note: In this question we consider active rotations of a wavefunction ψ about an axis in the direction of the unit vector $\hat{\mathbf{n}}$.

- The rotation operator for an angle ϕ about the axis $\hat{\mathbf{n}}$ is thus

$$U(\phi\,\hat{\mathbf{n}}) = U(\boldsymbol{\phi}) = \exp\left(-\frac{i}{\hbar}(\boldsymbol{\phi}\cdot\boldsymbol{L})\right),$$

  where we have defined the "vector angle" $\boldsymbol{\phi} = \phi\,\hat{\mathbf{n}}$. The generator of the rotation operator is $\hat{\mathbf{n}}\cdot\boldsymbol{L}$, the component of angular momentum along the rotation axis $\hat{\mathbf{n}}$.

- Rotational symmetry corresponds to conservation of angular momentum. A system's angular momentum is conserved if the system's Hamiltonian commutes with the angular momentum operator, i.e. $[\boldsymbol{L}, H] = 0$, which occurs when the Hamiltonian is invariant under rotation, i.e.

$$[U(\boldsymbol{\phi}), H] = 0 \text{ for all rotations } \boldsymbol{\phi}$$

  This form of conservation occurs for spherically symmetric potentials of the form $V(\boldsymbol{r}) = V(r)$.

Derivation: The Rotation Operator

- We first consider rotations by an infinitesimal angle $\mathrm{d}\phi$, for which a rotated wavefunction $\tilde{\psi}$ reads

$$\tilde{\psi}(\boldsymbol{r}) = \psi(\boldsymbol{r} - \mathrm{d}\boldsymbol{r}) \text{ where } \mathrm{d}\boldsymbol{r} = \mathrm{d}\phi(\hat{\mathbf{n}}\times\boldsymbol{r}).$$

  We then find the expression for the infinitesimal rotation operator with a first-order Taylor expansion, which reads

$$\tilde{\psi}(\boldsymbol{r}) = \psi(\boldsymbol{r} - \mathrm{d}\boldsymbol{r}) = \psi(\boldsymbol{r}) - \frac{i}{\hbar}\big[(\hat{\mathbf{n}}\times\boldsymbol{r})\cdot\boldsymbol{p}\big]\psi(\boldsymbol{r})\,\mathrm{d}\phi + \mathcal{O}(\mathrm{d}\phi^2)$$

$$= \left[\mathbf{I} - \frac{i}{\hbar}\big[\hat{\mathbf{n}}\cdot(\boldsymbol{r}\times\boldsymbol{p})\big]\,\mathrm{d}\phi\right]\psi(\boldsymbol{r}) + \mathcal{O}(\mathrm{d}\phi^2)$$

$$= \left[\mathbf{I} - \frac{i}{\hbar}(\hat{\mathbf{n}}\cdot\boldsymbol{L})\,\mathrm{d}\phi\right]\psi(\boldsymbol{r}) + \mathcal{O}(\mathrm{d}\phi^2),$$

  where $\mathbf{I}$ and $\boldsymbol{L} = \boldsymbol{r}\times\boldsymbol{p}$ are the identity and angular momentum operators.

- We then construct a rotation by the macroscopic angle ϕ from a product of $N\to\infty$ infinitesimal rotations by $\mathrm{d}\phi = \frac{\phi}{N}$ according to

$$\tilde{\psi}(\boldsymbol{r}) = \lim_{N\to\infty}\left(\mathbf{I} - \frac{i}{\hbar}(\hat{\mathbf{n}}\cdot\boldsymbol{L})\frac{\phi}{N}\right)^N \psi(\boldsymbol{r}) \equiv \exp\left(-\frac{i}{\hbar}(\hat{\mathbf{n}}\cdot\boldsymbol{L})\phi\right)\psi(\boldsymbol{r})$$

$$= \exp\left(-\frac{i}{\hbar}(\boldsymbol{\phi}\cdot\boldsymbol{L})\right).$$

  The last equality motivates the definition of the rotation operator as

$$U(\phi\,\hat{\mathbf{n}}) = U(\boldsymbol{\phi}) = \exp\left(-\frac{i}{\hbar}(\boldsymbol{\phi}\cdot\boldsymbol{L})\right),$$

  where we have defined the "vector angle" $\boldsymbol{\phi} = \phi\,\hat{\mathbf{n}}$.

4.3 Parity

Discuss the quantum-mechanical parity operator and give a physical interpretation of parity transformation. State and derive some of the parity operator's important quantities. Discuss the relationship of the parity operator to problems with even potentials.

- Parity transformation corresponds to space inversion, and is encoded by the parity operator $\mathcal{P}$, which maps $\boldsymbol{r}$ to $-\boldsymbol{r}$ in the form $\mathcal{P} : \psi(\boldsymbol{r}) \mapsto \psi(-\boldsymbol{r})$.

- The parity operator is Hermitian, which we prove with

$$\langle \phi(\boldsymbol{r})|\mathcal{P}|\psi(\boldsymbol{r})\rangle = \langle \phi(\boldsymbol{r})|\psi(-\boldsymbol{r})\rangle = \langle \phi(-\boldsymbol{r})|\psi(\boldsymbol{r})\rangle = \langle \mathcal{P}\phi(\boldsymbol{r})|\psi(\boldsymbol{r})\rangle .$$

  The parity operator is also unitary, i.e. $\mathcal{P}\mathcal{P} = \mathrm{I} \implies \mathcal{P} = \mathcal{P}^{-1}$.

- The parity operator changes the sign of the gradient (or derivative) operator, i.e.

$$\mathcal{P}\nabla\psi = -\nabla\mathcal{P}\psi \implies \mathcal{P}\nabla = -\nabla\mathcal{P}.$$

  The relationship $\mathcal{P}\nabla = -\nabla\mathcal{P}$ implies

$$\mathcal{P}\nabla^n = (-1)^n \nabla^n \mathcal{P} \qquad \text{and} \qquad \mathcal{P}\frac{\mathrm{d}^2}{\mathrm{d}x^2} = \frac{\mathrm{d}^2}{\mathrm{d}x^2}\mathcal{P},$$

  and the last two identities lead to

$$\mathcal{P}\boldsymbol{p} = -\boldsymbol{p}\mathcal{P} \qquad \text{and} \qquad \mathcal{P}(\boldsymbol{r} \times \boldsymbol{p}) = \mathcal{P}\boldsymbol{L} = \boldsymbol{L}\mathcal{P}.$$

- For problems with an even potential, is always possible to create an even or odd stationary state eigenfunction for each energy eigenvalue E.

Derivation: Parity Operator and an Even Potential

- For an even potential, i.e. $V(\boldsymbol{r}) = V(-\boldsymbol{r})$, the parity operator acts on V according to $\mathcal{P}V(\boldsymbol{r}) = V(-\boldsymbol{r})\mathcal{P} = V(\boldsymbol{r})\mathcal{P}$, in which case $\mathcal{P}$ and H commute, which follows from

$$\mathcal{P}H\psi(\boldsymbol{r}) = H\mathcal{P}\psi(\boldsymbol{r}) \implies [\mathcal{P}, H] = 0.$$

  If $\mathcal{P}$ and H commute, and if $|\psi(\boldsymbol{r})\rangle$ is a stationary state of the Hamiltonian and obeys the stationary Schrödinger equation

$$H |\psi(\boldsymbol{r})\rangle = E |\psi(\boldsymbol{r})\rangle ,$$

  then $|\psi(-\boldsymbol{r})\rangle$ is also a stationary state with the same energy E, i.e.

$$H |\psi(-\boldsymbol{r})\rangle = E |\psi(-\boldsymbol{r})\rangle$$

- We can then combine the stationary state solutions $|\psi(\boldsymbol{r})\rangle$ and $|\psi(-\boldsymbol{r})\rangle$ to create the odd and even functions $|\psi_+(\boldsymbol{r})\rangle$ and $|\psi_-(\boldsymbol{r})\rangle$ according to

$$\psi_\pm(\boldsymbol{r}) = \frac{1}{\sqrt{2}} \left(\psi(\boldsymbol{r}) \pm \psi(-\boldsymbol{r}) \right).$$

  In other words, for an even potential, we can always create an even or odd stationary state eigenfunction for each energy eigenvalue E (assuming E is nondegenerate).

  Note also that both $|\psi_+(\boldsymbol{r})\rangle$ and $|\psi_-(\boldsymbol{r})\rangle$ are eigenfunctions of the parity operator with eigenvalues ± 1, i.e.

$$\mathcal{P}\psi_+(\boldsymbol{r}) = \psi_+(\boldsymbol{r}) \qquad \text{and} \qquad \mathcal{P}\psi_-(\boldsymbol{r}) = -1 \cdot \psi_-(\boldsymbol{r})$$

4.4 Time Reversal

How is the time-reversal operator defined? Discuss the time reversal in the context of problems with time-independent potentials.
Discuss the basic properties of time reversal symmetry in quantum mechanics.

- The time reversal operator T maps time t to $-t$ in the form $T : \Psi(\boldsymbol{r}, t) \mapsto \Psi(\boldsymbol{r}, -t)$.

  The modified time reversal operator $\mathcal{T}$ is defined as

  $$\mathcal{T} = KT,$$

  where $K : \psi \mapsto \psi^*$ is the complex conjugation operator. The complex conjugation operator obeys $Kz = z^*K$ for all $z \in \mathbb{C}$ and equals its inverse, ie. $K = K^{-1}$.

- If $\Psi(\boldsymbol{r}, t)$ solves the Schrödinger equation for a time-independent potential $V = V(\boldsymbol{r})$ and Hamiltonian $H \neq H(t)$, then the transformed wavefunction $T\Psi(\boldsymbol{r}, t) = \Psi(\boldsymbol{r}, -t)$ solves the same Schrödinger equation with a transformed Hamiltonian $H \to -H$.

  Derivation: The Schrödinger equation for $\Psi(\boldsymbol{r}, t)$ reads

  $$i\hbar\frac{\partial \Psi(\boldsymbol{r}, t)}{\partial t} = H\Psi(\boldsymbol{r}, t).$$

  We then act on the equation with the time reversal operator T to get

  $$T\left(i\hbar\frac{\partial \Psi(\boldsymbol{r}, t)}{\partial t}\right) = i\hbar\frac{\partial \Psi(\boldsymbol{r}, -t)}{\partial(-t)} = H\Psi(\boldsymbol{r}, -t)$$

  $$\implies i\hbar\frac{\partial \Psi(\boldsymbol{r}, -t)}{\partial t} = -H\Psi(\boldsymbol{r}, -t).$$

  In other words, the transformed wavefunction $T\Psi(\boldsymbol{r}, t) = \Psi(\boldsymbol{r}, -t)$ solves the same Schrödinger equation with a transformed Hamiltonian $H \to -H$.

- If, as before, $\Psi(\boldsymbol{r}, t)$ solves the Schrödinger equation for a time-independent Hamiltonian $H \neq H(t)$, then the transformed wavefunction $\mathcal{T}\Psi(\boldsymbol{r}, t) = \Psi^*(\boldsymbol{r}, -t)$ solves the same Schrödinger equation with the same Hamiltonian H.

  Derivation: As before, the Schrödinger equation for $\Psi(\boldsymbol{r}, t)$ reads

  $$i\hbar\frac{\partial \Psi(\boldsymbol{r}, t)}{\partial t} = H\Psi(\boldsymbol{r}, t).$$

  We then act on the equation with the time reversal operator $\mathcal{T}$ to get

  $$\mathcal{T}\left(i\hbar\frac{\partial \Psi^*(\boldsymbol{r}, t)}{\partial t}\right) = -i\hbar\frac{\partial \Psi^*(\boldsymbol{r}, -t)}{\partial(-t)} = H\Psi^*(\boldsymbol{r}, -t)$$

  $$\implies i\hbar\frac{\partial \Psi^*(\boldsymbol{r}, -t)}{\partial t} = H\Psi^*(\boldsymbol{r}, -t)$$

  In this case, the transformed wavefunction $\mathcal{T}\Psi(\boldsymbol{r}, t) = \Psi^*(\boldsymbol{r}, -t)$ solves the same Schrödinger equation with the same Hamiltonian H.

- The operator $\mathcal{T}$ acts on stationary states of the form $\Psi(\boldsymbol{r}, t) = \psi(\boldsymbol{r})e^{-i\frac{E}{\hbar}t}$ as follows:
$$\Psi^*(\boldsymbol{r}, -t) = \psi^*(\boldsymbol{r})e^{-(-i)\frac{E}{\hbar}(-t)} = \psi^*(\boldsymbol{r})e^{-i\frac{E}{\hbar}t}$$

  In other words, $\mathcal{T}$ affects only the position-dependent term $\psi(\boldsymbol{r})$, which it conjugates. Thus, $\mathcal{T}$ acts on the stationary Schrödinger equation $H\psi(\boldsymbol{r}) = E\psi(\boldsymbol{r})$ to produce
$$H\psi^*(\boldsymbol{r}) = E\psi^*(\boldsymbol{r}),$$

  which implies that both ψ and ψ^* solve the stationary Schrödinger equation for a given energy eigenvalue E. For a non-degenerate spectrum, the conjugated function can be written $\psi^* = e^{i\phi}\psi(\boldsymbol{r})$. Since ψ and ψ^* differ only by a constant phase term $e^{i\delta}$ of magnitude 1, they correspond to physically identical wavefunction.

- The time reversal operator $\mathcal{T}$ acts on the momentum operator $\boldsymbol{p}$, angular momentum operator $\boldsymbol{L}$, and Hamiltonian H (assuming H is time-independent and real) as
$$\mathcal{T}\boldsymbol{p} = -\boldsymbol{p}\mathcal{T} \qquad \mathcal{T}\boldsymbol{L} = -\boldsymbol{L}\mathcal{T} \qquad \mathcal{T}H = H\mathcal{T}$$

- Finally, we note that position doesn't change sign under $\mathcal{T}$, i.e. $\mathcal{T}x = x\mathcal{T}$, as opposed to momentum, which obeys $\mathcal{T}p = -p\mathcal{T}$. These two identities imply
$$\mathcal{T}[x, p] = -[x, p]\mathcal{T}$$

  For the fundamental commutator relationship $[x, p] = i\hbar$ to remain invariant under $\mathcal{T}$ reversal, $\mathcal{T}$ must obey $\mathcal{T}i = i\mathcal{T}$, i.e. $\mathcal{T}$ must be an anti-unitary operator.

4.5 Invariance Under Phase Shift

Discuss a wavefunction's invariance under phase change in the context of gauge transformations. How do a wavefunction and basis change under a global phase or potential energy shift?

- Multiplying a wavefunction by a phase factor $e^{i\delta}$ of magnitude one, which is physically interpreted as a phase shift, has no physically observable effect on the wavefunction.

  Multiplication by $e^{i\delta}$ is a case of a so-called global gauge transformation, which is a unitary transformation of the form
$$U(\delta)\ket{\psi} = e^{i\delta}\ket{\psi} \equiv \ket{\tilde{\psi}},$$

  where we have defined the phase shift operator $U(\delta) = e^{i\delta}$.

- If we apply the phase shift transformation $U(\delta) = e^{i\delta}$ to all basis functions $\{\ket{n}\}$ spanning the Hilbert space of wavefunctions, then all matrix elements of an arbitrary operator $\mathcal{O}$ remain unchanged, i.e.
$$\braket{U(\delta)\phi | \mathcal{O} | U(\delta)\psi} = \braket{\tilde{\phi} | \mathcal{O} | \tilde{\psi}} = \braket{\phi | \mathcal{O} | \psi}$$

  Even more, we can multiply each basis vector $\ket{n}$ by an individual factor $e^{i\delta_n}$, and all physical observable remain unchanged.

- A potential energy transformation $V(\boldsymbol{r}) \to V(\boldsymbol{r}) + V_0$ shifts a system's energy eigenvalues by V_0, i.e. $E_n \to E_n + V_0$ and transforms the time evolution operator to

$$e^{-i\frac{E}{\hbar}t} \left|\psi\right\rangle \to e^{-i\frac{E+V_0}{\hbar}t} \left|\psi\right\rangle = e^{-i\frac{E}{\hbar}t}e^{-i\frac{V_0}{\hbar}t} \left|\psi\right\rangle.$$

The change in energy creates a phase shift, which is encoded by a time-dependent global gauge transformation of the form

$$U(\delta(t)) \equiv e^{-i\frac{V_0}{\hbar}t} \qquad \text{where } \delta(t) = -\frac{V_0}{\hbar}t.$$

- We can also define a so-called local gauge transformation of the form

$$U(\delta(\boldsymbol{r}, t)) \left|\Psi(\boldsymbol{r}, t)\right\rangle = e^{i\delta(\boldsymbol{r}, t)} \left|\Psi(\boldsymbol{r}, t)\right\rangle \equiv \left|\tilde{\Psi}(\boldsymbol{r}, t)\right\rangle,$$

which preserves probability density, i.e.

$$\left|\tilde{\Psi}(\boldsymbol{r}, t)\right|^2 = \left|\Psi(\boldsymbol{r}, t)\right|^2.$$

We will return to local gauge transformations when discussing a particle in an electromagnetic field.

5 Angular Momentum

5.1 Important Properties of Angular Momentum

Discuss the definition and basic properties of angular momentum in quantum mechancis. Be sure to include the important angular momentum commutation relations. How is the squared angular momentum operator defined?

- Angular momentum is assigned to the Hermitian angular momentum operator $\boldsymbol{L}$, which is defined as
$$\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}.$$

- The operator $\boldsymbol{L}$ obeys the relationship $\boldsymbol{L} = -\boldsymbol{p} \times \boldsymbol{r}$.

  Derivation of $\boldsymbol{L} = -\boldsymbol{p} \times \boldsymbol{r}$: Using the definition $\boldsymbol{L} = \boldsymbol{r} \times \psi$ and using only ∇ instead of ψ (since the the constant factor $-i\hbar$ doesn't affect the end result), we have

$$(\nabla \times \boldsymbol{r})\psi = \nabla \times (\boldsymbol{r}\psi) = (\nabla\psi) \times \boldsymbol{r} + \psi(\nabla \times \boldsymbol{r}) = (\nabla\psi) \times \boldsymbol{r} + 0$$
$$= -\boldsymbol{r} \times (\nabla\psi) = -(\boldsymbol{r} \times \nabla)\psi$$
$$\implies \boldsymbol{r} \times \boldsymbol{p} = -\psi \times \boldsymbol{p}.$$

  We could also prove the relationship by components, e.g.

$$L_z\psi = -i\hbar(yp_x - xp_y)\psi = i\hbar(p_yx - p_xy)\psi.$$

  We would proceed analogously for $L_x = yp_z - zp_y$ and $L_y = zp_x - xp_z$.

- The most important angular momentum commutation relation is

$$\left[L_\alpha, L_\beta\right] = i\hbar\epsilon_{\alpha\beta\gamma}L_\gamma \qquad \text{or, in vector form,} \qquad \boldsymbol{L} \times \boldsymbol{L} = i\hbar\boldsymbol{L}.$$

  The components of angular momentum L_x, L_y and L_z obey analogous commutator relations to the Poisson bracket relations in classical mechanics, for example

$$[L_x, L_y] = i\hbar(xp_y - yp_x) = i\hbar L_z \qquad \text{and} \qquad [L^2, L_\alpha] = 0, \quad \alpha = x, y, z.$$

  More generally, angular momentum obeys the commutator relation

$$[L_\alpha, \mathcal{O}_\beta] = i\hbar\epsilon_{\alpha\beta\gamma}\mathcal{O}_\gamma,$$

  where the operator $\mathcal{O}$ can be any of $\boldsymbol{r}$, $\boldsymbol{p}$ or $\boldsymbol{L}$.

  Derivation of $[L_x, L_y] = i\hbar L_y$: We begin with the definitions of L_x and L_y:

$$[L_x, L_y] = [yp_z - zp_y, zp_x - xp_z]$$
$$= [yp_z, zp_x] - [zp_y, zp_z] - [yp_z, xp_z] + [zp_y, xp_z]$$

  - The middle two commutators are zero, since both sides contain identical terms z and p_z, respectively.
  - We expand the remaining commutators using

$$[AB, C] = A[B, C] + [A, C]B \qquad \text{and} \qquad [A, BC] = B[A, C] + [A, B]C.$$

– Finally, we apply the canonical commutator relations $[r_\alpha, r_\beta] = [p_\alpha, p_\beta] = 0$ and $[r_\alpha, p_\beta] = i\hbar\delta_{\alpha\beta}$ to get

$$[L_x, L_y] = yp_x[p_z, z] + xp_y[z, p_z] = i\hbar(yp_x + xp_y) = i\hbar L_z$$

- In general, any operator $\mathcal{O}$ that is invariant under rotation (i.e. for which $U(\phi)\mathcal{O} = \mathcal{O}U(\phi)$) commutes with the angular momentum operator. In symbols,

$$U(\phi)\mathcal{O} = \mathcal{O}U(\phi) \implies [\boldsymbol{L}, \mathcal{O}] = 0$$

where $U(\phi)$ is the operator encoding rotation about the axis $\hat{\mathbf{n}}$ by the angle ϕ. Relevant operators invariant under rotation include $\boldsymbol{r}\cdot\boldsymbol{p}$, $\boldsymbol{p}^2$, $\boldsymbol{L}^2$ and all rotationally invariant potentials of the form $V = V(|\boldsymbol{r}|)$.

- We often work in terms of the squared angular momentum $\boldsymbol{L}^2$, which we can write in any of the equivalent forms

$$\boldsymbol{L}^2 = \boldsymbol{L} \cdot \boldsymbol{L} = L^2 = \sum_\alpha L_\alpha^2$$

The square of angular momentu L^2 is invariant under rotations, which is summarized by the commutation relation

$$[L_\alpha, L^2] = 0 \quad \text{for } \alpha \in \{x, y, z\}.$$

Because L_α and L^2 commute, squared angular momentum L^2 and its components L_α can share the same eigenvectors and basis.

5.2   The Ladder Operators

Define the ladder operators for angular momentum, and discuss some of their important properties.

- The angular momentum ladder operators L_+ and L_- are defined as

$$L_+ \equiv L_x + iL_y \qquad \text{and} \qquad L_- \equiv L_x - iL_y.$$

The ladder operators obey $L_\pm = L_\mp^\dagger$, i.e. they are each other's Hermitian conjugates.

- To recover the operators L_x and L_y, we invert the above relationships to get

$$L_x = \frac{L_+ + L_-}{2} \qquad \text{and} \qquad L_y = \frac{L_+ - L_-}{2i}.$$

- The ladder operators commute with the squared angular momentum operator, i.e.

$$[L^2, L_+] = [L^2, L_-] = 0.$$

- The ladder operators and the component L_z obey the commutation relation

$$[L_z, L_\pm] = \pm\hbar L_\pm.$$

Derivation: We begin with the defintions of $L_{\pm}$ in terms of L_x and L_y, and then apply the identity $[L_\alpha, L_\beta] = i\hbar\epsilon_{\alpha\beta\gamma}L_\gamma$. The calculation reads

$$[L_z, L_\pm] \equiv [L_z, L_x \pm iL_y] = [L_z, L_x] \pm i[L_z, L_y] = i\hbar L_y \pm i(-i\hbar L_x)$$
$$= \hbar[\pm L_x + iL_y] = \pm\hbar L_\pm.$$

- The ladder operators the equally important commutation relation

$$[L_+, L_-] = 2\hbar L_z.$$

Derivation: We begin with the auxiliary calculation

$$L_\pm L_\mp = (L_x \pm iL_y)(L_x \mp iL_y) = L_x^2 + L_y^2 \pm iL_yL_x \mp iL_xL_y$$
$$= L^2 - L_z^2 \pm iL_yL_x \mp iL_xL_y.$$

Next, we use $[L_\alpha, L_\beta] = i\hbar\epsilon_{\alpha\beta\gamma}L_\gamma$ to show

$$\pm iL_yL_x \mp iL_xL_y = \pm\hbar L_z,$$

which we substitute into the expresion for $L_\pm L_\mp$ to get

$$L_\pm L_\mp = L^2 - L_z^2 \pm \hbar L_z,$$

which is really two equations of the form

$$L_+L_- = L^2 - L_z^2 + \hbar L_z \qquad \text{and} \qquad L_-L_+ = L^2 - L_z^2 - \hbar L_z.$$

Finally, we subtract the two equations two get the desired identity

$$L_+L_- - L_-L_+ \equiv [L_+, L_-] = 2\hbar L_z$$

- Analogy to the quantum harmonic oscillator: the ladder operators L_+ and L_- are analogous to the creation and annihilation operators $a^\dagger$ and a, i.e. they "raise" and "lower" the indexes of angular momentum basis states, just like $a^\dagger$ and a raise and lower the indexes of the harmonic oscillator's Hamiltonian's basis states.

  The operator L_z is analogous to the counting operator $\hat{n} = a^\dagger a$, in that it counts the number of angular momentum quanta in an angular momentum basis state.

5.3 Angular Momentum Eigenvalues and Eigenfunction

State and derive the angular momentum eigenfunctions, eigenvalues and eigenbasis. Discuss the degeneracy of the L^2 eigenvalue spectrum.

How is the angular momentum operator written in matrix form in its eigenbasis?

- The eigenvalue relations for L_z and L^2 read

$$L_z |lm\rangle = m\hbar |lm\rangle \qquad \text{and} \qquad L^2 |lm\rangle = \hbar^2 l(l+1) |lm\rangle,$$

where, mathematically, the quantum numbers l and m can take the values

$$l \in \{0, \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, \ldots\} \qquad \text{and} \qquad m \in \{-l, -l+1, \ldots, l-1, l\}.$$

However, only integer values of m satisfy the Schrödinger equation and correspond to physical eigenstates of L_z. When allowing for only integer values of m, the possible values of l and m are

$$l \in \{0, 1, 2, \ldots\} \qquad \text{and} \qquad m \in \{-l, -l+1, \ldots, l-1, l\}.$$

Half-integer values of the angular quantum number correspond to spin, which is discussed in [Subsection 8.1](#).

- Since m can take on $2l+1$ values at a given l, the L^2 eigenvalue spectrum has degeneracy $2l+1$, since at a given l there are $2l+1$ linearly independent eigenstates $|lm\rangle$ with the same eigenvalue $\lambda = l(l+1)\hbar^2$.

- The eigenfunctions of L_z are oscillating wavefunctions of the form

$$\psi_m(\phi) = Ce^{im\phi},$$

where ϕ is the azimuthal angle in spherical coordinates and m is the eigenfunction's quantum number.

- The eigenfunctions of L^2 are the spherical harmonics,

$$\langle \boldsymbol{r}|lm\rangle = Y_l^m(\theta, \phi),$$

which we often write in quantum mechanics for $m \geq 0$ as

$$Y_l^m(\theta, \phi) = (-1)^m \sqrt{\frac{(2l+1)}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta)e^{im\phi},$$

where P_l^m are the associated Legendre polynomials and θ and ϕ are the polar and azimuthal angle in spherical coordinates.

5.3.1 Derivation: Eigenvalues of L_z and L^2

- Let $|m\rangle$ be a hypothetical eigenstate of L_z with eigenvalue $m\hbar$, where m is yet to be determined. In terms of $|m\rangle$ and $m\hbar$, the L_z eigenvalue equation reads

$$L_z |m\rangle = m\hbar |m\rangle.$$

- Next, we use $[L_z, L_\pm] = \pm\hbar L_\pm$ to show that the operator $L_z L_\pm$ acts on $|m\rangle$ as

$$L_z\big[L_\pm |m\rangle\big] = (m \pm 1)\hbar\big[L_\pm |m\rangle\big].$$

Derivation: We rewrite $[L_z, L_\pm] = \pm\hbar L_\pm$ in the form $L_z L_\pm = L_\pm L_z \pm \hbar L_\pm$, distribute, use the eigenvalue relation $L_z |m\rangle = m\hbar |m\rangle$ and factor to get

$$\begin{aligned}
L_z L_\pm |m\rangle = (L_\pm L_z \pm \hbar L_\pm)|m\rangle &= L_\pm L_z |m\rangle \pm \hbar L_\pm |m\rangle \\
&= L_\pm m\hbar |m\rangle \pm \hbar L_\pm |m\rangle \\
&= (m \pm 1)\hbar L_\pm |m\rangle.
\end{aligned}$$

Since L_z acts on the state $L_\pm \left| m \right\rangle$, to produce $(m \pm 1)\hbar L_\pm \left| m \right\rangle$, i.e. the same state with eigenvalue $(m \pm 1)\hbar$, it follows that L_+ and L_- raise and lower the index m of the state $\left| m \right\rangle$ by one,[1] i.e.

$$L_\pm \left| m \right\rangle \propto \left| m + 1 \right\rangle.$$

- Next, because L_z and L^2 commute, L_z's eigenstate $\left| m \right\rangle$ is also an eigenstate of $\left| L^2 \right\rangle$. We write the eigenvalue relation for L^2 in the form

$$L^2 \left| m \right\rangle = \lambda \left| m \right\rangle,$$

where λ is a to-be-determined eigenvalue. First, we show that the eigenvalues of L^2 cannot be negative, i.e.

$$\lambda \geq 0.$$

Derivation: We multiply $L^2 \left| m \right\rangle = \lambda \left| m \right\rangle$ through by $\left\langle m \right|$ and apply the positive-definite inner product identity $\left\langle m | m \right\rangle \geq 0$ to get

$$\langle m | L^2 | m \rangle = \sum_\alpha \left\langle L_\alpha m | L_\alpha m \right\rangle = \lambda \left\langle m | m \right\rangle \geq 0 \implies \lambda \geq 0.$$

- Next, we use the commutator relation $[L^2, L_\pm] = 0$ to show that

$$L^2 \big[L_\pm \left| m \right\rangle \big] = \lambda \big[L_\pm \left| m \right\rangle \big].$$

Derivation: We rewrite $[L^2, L_\pm] = 0$ to get $L^2 L_\pm = L_\pm L^2$ and apply the eigenvalue relation $L^2 \left| m \right\rangle = \lambda \left| m \right\rangle$ to get

$$L^2 L_\pm \left| m \right\rangle = L_\pm L^2 \left| m \right\rangle = L_\pm \lambda \left| m \right\rangle = \lambda L_\pm \left| m \right\rangle.$$

In other words, the state $L_\pm \left| m \right\rangle$ is also an eigenstate of L^2 with the eigenvalue λ. Since the state $L_\pm \left| m \right\rangle$ obeys $L_\pm \left| m \right\rangle \propto \left| m + 1 \right\rangle$, i.e. has quantum number one higher or one lower than $\left| m \right\rangle$, and both $\left| m \right\rangle$ and $\left| m \pm 1 \right\rangle$ have the same eigenvalue λ, it follows that under the action of L^2 states $\left| m \pm 1 \right\rangle$ have the same eigenvalue as $\left| m \right\rangle$.

- Next, we parameterize the L^2 eigenvalue in the form

$$\lambda = l(l+1)\hbar^2,$$

where we have introduced the orbital quantum number $l > 0$. As an auxiliary calculation, we use the identities $L_\pm^\dagger = L_\mp$ and $L_\pm L_\mp = L^2 - L_z^2 \pm \hbar L_z$ to show

$$\left\langle L_\pm m | L_\pm m \right\rangle = \left\langle m | L_\mp L_\pm | m \right\rangle = \left\langle m | (L^2 - L_z^2 \mp \hbar L_z) | m \right\rangle,$$

into which we then substitute the eigenvalue relations for L^2 and L_z to get

$$\left\langle L_\pm m | L_\pm m \right\rangle = \left\langle m \big| \big[l(l+1) - m(m \pm 1) \big] \hbar^2 \big| m \right\rangle$$
$$\equiv (C_{l,m\pm 1})^2 \left\langle m | m \right\rangle,$$

where we have defined the constant

$$C_{l,m\pm 1} = \hbar \sqrt{l(l+1) - m(m \pm 1)} \in \mathbb{R}.$$

Since both $\left\langle L_\pm m | L_\pm m \right\rangle$ and $\left\langle m | m \right\rangle$ are non-negative, it follows that $(C_{l,m\pm 1})^2 \geq 0$, which is why $C_{l,m\pm 1}$ is real.

---

[1] See also the analogous analysis of the harmonic oscillator ladder operators in <u>Subsection 3.2</u>.

- The identity $C_{l,m\pm1} \in \mathbb{R}$ leads to the desired L_z eigenvalue relation

$$|m| \leq l.$$

Here's why: The fact that $C_{l,m,\pm1}$ is real means that for a state with a given orbital quantum number l, we can raise or lower states with $L_\pm$ only as long as $C_{l,m\pm1}$ remains real. The constant $C_{l,m\pm1}$ is real as long as the square root in its definition remains non-negative, which implies the condition

$$l(l+1) \geq m(m \pm 1) \implies |m| \leq l.$$

We will complete the derivation of $m \in \{-l, -l+1, \ldots, l-1, l\}$ below.

- Next, we consider a generic L^2 eigenstate $|lm\rangle$ indexed by both m and l. We start with the maximum permitted value of m, i.e. $m = l$, and act on the state $|ll\rangle$ with the operator L_- until we reach the minimum possible value $m = -l$. This reads

$$
\begin{aligned}
L_- |ll\rangle &= C_{l,l-1} |l, l-1\rangle \\
L_-^2 |ll\rangle &= C_{l,l-2} |l, l-2\rangle \\
&\vdots \\
L_-^k |ll\rangle &= C_{l,l-k} |l, l-k\rangle = C_{l,l-k} |l, -l\rangle.
\end{aligned}
$$

Since we reached the state with $m = -l$ after $k \in \mathbb{N}$ integer steps, we have

$$l - k = -l \implies 2l = k \implies 2l \in \mathbb{N}.$$

The possible values of l, (accounting for $l \geq 0$), are thus

$$
l = \begin{cases} 0, 1, 2, \ldots & k \text{ even} \\ \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots & k \text{ odd.} \end{cases}
$$

The unit increments of l mean that $|m| \leq l$ can be written in the desired form

$$m \in \{-l, -l+1, \ldots, l-1, l\},$$

which proves the L_z eigenvalue relation quoted ad the beginning of this question.

- The recursive action of L_- on $|lm\rangle$ also reveals the relationship

$$L_\pm |lm\rangle = C_{l,m\pm1} |l, m \pm 1\rangle = \hbar\sqrt{l(l+1) - m(m \pm 1)} \, |l, m \pm 1\rangle$$

Earlier, we had determined this relationship only to $L_\pm |m\rangle \propto |m \pm 1\rangle$.

5.3.2   Derivation: Eigenfunctions of L_z

- In spherical coordinates, the coordinate representation of the operator L_z reads

$$L_z = -i\hbar \frac{\partial}{\partial \phi}$$

where ϕ is the azimuthal angle. For each m, the eigenvalues equation

$$L_z \psi_m = \left(-i\hbar \frac{\partial}{\partial \phi} \right) \psi_m = m\hbar \psi_m$$

has the unique solution

$$\psi_m = C e^{im\phi}$$

- We consider only ψ_m solving the Schrödinger equation, which must be continuous. To satisfy continuity, the ψ_m must be periodic over $\phi \in [0, 2\pi]$, i.e.

$$\psi_m(\phi) = \psi_m(\phi + 2\pi) \iff 1 = e^{2\pi i m} \implies m \in \mathbb{Z}$$

In other words, only integer values of m satisfy the Schrödinger equation and correspond to physical eigenstates of L_z, which leads to the eigenvalue relationship

$$l \in \{0, 1, 2, \ldots\} \qquad \text{and} \qquad m \in \{-l, -l+1, \ldots, l-1, l\},$$

which was quoted at the begining of this question.

5.3.3 Discussion: Eigenfunctions of L^2

- Without derivation, the eigenfunctions of the angular momentum operator L^2, in the coordinate representation, are the spherical harmonics, i.e.

$$\langle \boldsymbol{r} | lm \rangle = Y_l^m(\theta, \phi)$$

The spherical harmonics arise in the angular solution of the Laplace equation $\nabla^2 u(\boldsymbol{r}) = 0$, i.e. if we separate $u(\boldsymbol{r})$ into radial and angular component, the solution is

$$\nabla^2 u(\boldsymbol{r}) = f(r) Y_l^m(\theta, \phi) = 0$$

where $Y_l^m(\theta, \phi)$ are the spherical harmonics.

- In quantum mechanics for $m \geq 0$ we often use the definition

$$Y_l^m(\theta, \phi) = (-1)^m \sqrt{\frac{(2l+1)}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) e^{im\phi}$$

where P_l^m are the associated Legendre polynomials.

- The spherical harmonics obey

$$Y_l^{-m} = (-1)^m Y_l^{m*}$$

- As a concrete example, the first few spherical harmonics for for $l = 0, 1, 2$ are

$$\begin{aligned}
Y_0^0 &= \tfrac{1}{\sqrt{4\pi}} \\
Y_1^0 &= \sqrt{\tfrac{3}{4\pi}} \cos\theta & Y_l^{\pm 1} &= \mp\sqrt{\tfrac{3}{8\pi}} \sin\theta e^{\pm i\phi} \\
Y_2^0 &= \sqrt{\tfrac{5}{16\pi}}(3\cos^2\theta - 1) & Y_2^{\pm 1} &= \mp\sqrt{\tfrac{15}{8\pi}} \sin\theta\cos\theta e^{\pm i\phi} \\
Y_2^{\pm 2} &= \sqrt{\tfrac{5}{32\pi}} \sin^2\theta e^{\pm 2i\phi}
\end{aligned}$$

5.3.4 Matrix Representation of Angular Momentum

- We write a generic state $|\psi\rangle$ in the angular momentum eigenbasis $\{|lm\rangle\}$ as

$$|\psi\rangle = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{lm} |lm\rangle$$

Note that the presence of two quantum numbers l and m introduces a double sum.

- We write a generic operator $\mathcal{O}$ in the $|lm\rangle$ basis as

$$\mathcal{O} = \sum_{l'lm'm} |l'm'\rangle \, \mathcal{O}_{l'lm'm} \, \langle lm|$$

- Finally, as a concrete example, for $l = 1$ the matrices for $L_{x,y,z}$ and L^2 read

$$L_x = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad L_y = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}$$

$$L_z = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \qquad L^2 = 2\hbar^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

As would be expected, L^2 and L_z are diagonal in the $|lm\rangle$ basis, since $|lm\rangle$ are the eigenstates of L^2 and L_z.

6   Central Potenial

6.1   The Quantum-Mechanical Central Potential Problem

State the basis problem of a particle in a central potential. Explain the general solution procedure, and state and derive the radial eigenvalue equation.

- The quantum central potential problem involves solving a Hamiltonian of the form

$$H = \frac{p^2}{2m} + V(r) = -\frac{\hbar^2}{2m}\nabla^2 + V(r),$$

  where $V = V(r)$ is a time-independent central potential.

- A quantum central potential problem is characterized by the following relations:

  - $[\boldsymbol{L}, H] = [L^2, H] = 0$, implying conservation of momentum $\boldsymbol{L}$ and magntiude of magnitude of angular momentum L^2.
  - $\boldsymbol{r} \cdot \boldsymbol{L} = 0$ implying that the particle's motion is confined to a plane.
  - $\boldsymbol{p} \cdot \boldsymbol{L} = 0$, implying that the particle's velocity is confined to the same plane.

  The last two equations are the quantum mechanical analog of a particle's motion and velocity lying in a two-dimensional plane in central force motion.

- We solve the central potential eigenvalue problem $H\Psi(\boldsymbol{r}) = E\Psi(\boldsymbol{r})$ using the ansatzes

$$\Psi(\boldsymbol{r}) = \psi(r)Y_l^m(\theta, \phi) \qquad \text{and} \qquad \psi(r) = \frac{u}{r},$$

  where we have separated $\Psi(\boldsymbol{r})$ into a radial and angular component and defined a further ansatz for the radial component $\psi(r)$. Substituting these ansatzes into the central potential Schrödinger equation leads to the radial eigenvalue equation

$$-\frac{\hbar^2}{2m}u''(r) + V_{\text{eff}}(r)u(r) = Eu(r), \qquad V_{\text{eff}}(r) = V(r) + \frac{l(l+1)\hbar^2}{2mr^2}.$$

  The derivation of the radial eigenvalue equation follows below.

6.1.1   Rewriting the Laplacian

- Since the central potential problem is spherically symmetric, we analyze the problem in spherical coordinates, where the Laplace operator reads

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\sin\theta\frac{\partial}{\partial \theta} + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2}{\partial \phi^2}.$$

- Without proof, the Laplace operator's angular component is related to angular momentum L^2 via

$$\frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\sin\theta\frac{\partial}{\partial \theta} + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2}{\partial \phi^2} = -\frac{L^2}{\hbar^2 r^2},$$

  and the Laplacian can thus be written

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} - \frac{L^2}{\hbar^2 r^2},$$

where the angular component $\frac{L^2}{\hbar^2 r^2}$ corresponds to rotational kinetic energy.

We then use the above form of ∇^2 to decompose the Hamiltonian into a radial and angular component:

$$H = -\frac{\hbar^2}{2m}\nabla^2 + V(r) = -\frac{h^2}{2m}\left(\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}\right) + \frac{L^2}{2mr^2} + V(r).$$

6.1.2   Initial Form of the Radial Equation

- Our next step is to solve the stationary Schrödinger equation

$$H\Psi(\boldsymbol{r}) = E\Psi(\boldsymbol{r}).$$

To do this, we use the ansatz

$$\Psi(\boldsymbol{r}) = \psi(r)Y_l^m(\theta,\phi),$$

where we have separated $\Psi(\boldsymbol{r})$ into a radial and angular component. The spherical harmonics $Y_l^m(\theta,\phi)$ are a natural choice for the angular component because they are the eigenfunctions of the angular momentum operator L^2.

- We substitute the ansatz $\Psi(\boldsymbol{r}) = \psi(r)Y_l^m(\theta,\phi)$ into the stationary Schrödinger equation $H\Psi = E\Psi$, which produces

$$\left[-\frac{\hbar^2}{2m}\left(\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}\right) + \frac{L^2}{2mr^2} + V(r)\right]\psi(r)Y_l^m(\theta,\phi) = E\psi(r)Y_l^m(\theta,\phi).$$

We then apply the angular momentum eigenvalue relation

$$L^2 Y_l^m = l(l+1)\hbar^2 Y_l^m$$

and cancel Y_l^m from both sides of the equation to produce the purely radial problem

$$-\frac{h^2}{2m}\left(\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}\right)\psi(r) + \left(V(r) + \frac{l(l+1)\hbar^2}{2mr^2}\right)\psi(r) = E\psi(r).$$

6.1.3   Simplifying the Radial Equation

- We solve for the radial eigenfunction $\psi(r)$ with ansatz

$$\psi(r) = \frac{u(r)}{r},$$

which we substitute into the radial Schrödinger equation to get

$$-\frac{h^2}{2m}\left(\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}\right)\left(\frac{u}{r}\right) + \left(V(r) + \frac{l(l+1)\hbar^2}{2mr^2}\right)\left(\frac{u}{r}\right) = E\frac{u}{r}.$$

- We then make the auxiliary calculation

$$\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}\left(\frac{u}{r}\right) = \frac{1}{r^2}\frac{\partial}{\partial r}r^2\left(\frac{u'}{r} - \frac{u}{r^2}\right) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(ru' - u\right) = \frac{u''}{r},$$

substitute the result into the radial eigenvalue equation in terms of u, and multiply through by r, which simplifies things considerably to

$$-\frac{\hbar^2}{2m}u''(r) + \left[V(r) + \frac{l(l+1)\hbar^2}{2mr^2}\right]u(r) = Eu(r).$$

- Finally, we define an effective potential

$$V_{\text{eff}}(r) = V(r) + \frac{l(l+1)\hbar^2}{2mr^2}$$

which includes the potential $V(r)$ in additional to the "centrifugal" term $\frac{l(l+1)\hbar^2}{2mr^2}$. In terms of V_{eff}, the stationary Schrödinger equation for u reads

$$-\frac{\hbar^2}{2m}u''(r) + V_{\text{eff}}(r)u(r) = Eu(r),$$

which is the radial eigenvalue equation quoted at the beginning of the question.

Note that we have reduced originally three-dimensional problem, involving the complete position vector $\boldsymbol{r} = (r, \phi, \theta)$, to a one-dimensional problem, involving only r.

6.2   The Radial Solution for $r \to 0$

State and derive the solution to the radial eigenvalue equation in the limit $r \to 0$. You may restrict your analysis to potentials obeying the limit $\lim_{r \to 0} r^2 V(r) = 0$.

- The solution to the radial eigenvalue equation in the limit $r \to 0$ is

$$u(r) = C_l r^{l+1} \qquad \text{or} \qquad \psi(r) = C_l r^l,$$

where l is the orbital quantum number.

Derivation: Solution to Radial Equation for $r \to 0$

- For review, the radial eigenvalue equation, with V_{eff} written out, reads

$$-\frac{\hbar^2}{2m}u''(r) + \left(V(r) + \frac{l(l+1)\hbar^2}{2mr^2}\right)u(r) = Eu(r).$$

For potentials $V(r)$ of the form

$$\lim_{r \to 0} r^2 V(r) = 0,$$

we can neglect the terms $V(r)$ and E. This because the centrifugal component $\frac{l(l+1)\hbar^2}{2mr^2}$ scales as $\sim r^{-2}$, and as $r \to 0$, the centrifugal term dominates $V(r)$ and E. We thus neglect $V(r)$ and E, in which case the eigenvalue equation simplifies to

$$-\frac{\hbar^2}{2m}u''(r) + \frac{l(l+1)\hbar^2}{2mr^2}u(r) = 0 \implies u''(r) = \frac{l(l+1)}{r^2}u(r).$$

- We solve the resulting equation with the ansatz $u(r) = Cr^\lambda$, which produces

$$\lambda(\lambda - 1)Cr^{\lambda-2} = \frac{l(l+1)}{r^2}Cr^\lambda \implies \lambda(\lambda - 1) = l(l+1).$$

We solve the polynomial equation $\lambda(\lambda - 1) = l(l+1)$ with the quadratic formula:

$$\lambda_\pm = \frac{1}{2} \pm \frac{1}{2}\sqrt{1 + 4l(l+1)} = \frac{1}{2} \pm \frac{1}{2}\sqrt{(2l+1)^2} = \frac{1}{2} \pm \left(l + \frac{1}{2}\right).$$

The two possible values of λ are thus

$$\lambda_+ = l + 1 \qquad \text{and} \qquad \lambda_- = -l,$$

and the general solution to the second-order linear eigenvalue equation is the linear combination

$$u(r) = C_+ r^{\lambda_+} + C_- r^{\lambda_-} = C_l r^{l+1} + \frac{D_l}{r^l}.$$

- Our next step is to determine the constants C_l and D_l from boundary and normalization conditions. We start all the way back at the normalization condition on $\Psi(\boldsymbol{r})$, which, when integrating in spherical coordinates, reads

$$1 \equiv \langle \Psi | \Psi \rangle = \int_{r=0}^{\infty} |\psi(r)|^2 r^2 \,\mathrm{d}r \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} |Y_l^m(\theta, \phi)|^2 \sin\theta \,\mathrm{d}\theta \,\mathrm{d}\phi.$$

Since the spherical harmonics are normalized, the integral's angular component evaluates to one, which implies

$$\int_0^{\infty} |\psi(r)|^2 r^2 \,\mathrm{d}r = \int_0^{\infty} |u(r)|^2 \,\mathrm{d}r \equiv 1.$$

This normalization condition on u requires $D_l = 0$ for $l > 0$, since the integral of $|u(r)|^2$ would otherwise diverge at 0.

- We now separately consider the case $l = 0$, in which case the radial equations for u and $\psi = u/r$ are

$$u(r) = C_0 r + D_0 \qquad \text{and} \qquad \psi(r) = C_0 + \frac{D_0}{r}.$$

It turns out that $D_0 = 0$, leaving us with

$$u(r) = C_0 r \qquad \text{and} \qquad \psi(r) = C_0.$$

We will explain why $D_0 = 0$ using an analogy from electrostatics. We consider the Poisson equation for the electrostatic potential ϕ, which reads

$$\nabla^2 \phi(\boldsymbol{r}) = -\frac{\rho(\boldsymbol{r})}{\epsilon_0}.$$

For a point charge with charge density $\rho(\boldsymbol{r}) = q\delta^3(\boldsymbol{r})$, taking the result from electrostatics and leaving out the derivation, the solution to the Poisson equation is

$$\phi(\boldsymbol{r}) = \frac{q}{4\pi\epsilon_0 r}.$$

If we substitute in the expressions for ρ and ϕ for a point charge into the Poisson equation and cancel common terms, the result is

$$\nabla^2 \frac{q}{4\pi\epsilon_0 r} = -\frac{q\delta^3(\boldsymbol{r})}{\epsilon_0} \implies \nabla^2 \frac{1}{4\pi r} = -\delta^3(\boldsymbol{r}).$$

We then rewrite the last equality in the more general form

$$\nabla^2 \psi(r) = -4\pi D_0 \delta^3(\boldsymbol{r}),$$

where we have defined $\psi = D_0/r$. In the case of a point charge we have $D_0 = 1$, but we have included D_0 to match the term D_0/r in the $l = 0$ case of the radial equation for $r \to 0$. The conclusion so far is: a wavefunction of the form $\psi(r) = D_0/r$ with $D_0 \neq 0$ corresponds to an equation fo the form

$$\nabla^2 \psi(\boldsymbol{r}) = -4\pi D_0 \delta^3(\boldsymbol{r}),$$

where the key point is that the above equation for ψ contains a delta function $\delta^3(\boldsymbol{r})$ at the origin. However, ψ is supposed to solve the Schrödinger equation

$$H\psi = \left(-\frac{\hbar^2}{2m}\nabla^2 + V(r)\right)\psi = E\psi,$$

which certainly does not contain a delta function at the origin. The conclusion is that ψ thus cannot contain terms of the form D_0/r, which requires $D_0 = 0$ for $l = 0$, and in fact $D_l = 0$ for all l. With $D_l = 0$ for all l, the solution to the radial equation in the limit $r \to 0$ simplifies to

$$u(r) = C_l r^{l+1} \qquad \text{and} \qquad \psi(r) = C_l r^l,$$

which is the result quoted at the begining of the chapter.

6.3   The Radial Solution for $r \to \infty$

State, derive and interpret the solution to the radial eigenvalue equation in the limit $r \to \infty$. You may restrict your analysis to potentials vanishing at infinity, but be sure to formally discuss the validity of the results.

- In the limit $r \to \infty$, assuming a vanishing potential, the solutions to the radial equaiton are

$$u(r) = \begin{cases} C_- e^{-ikr} + C_+ e^{ikr}, & k = \sqrt{\frac{2mE}{\hbar^2}}; & E > 0 \\ D_- e^{-\kappa r} + D_+ e^{\kappa r}, & \kappa = \sqrt{\frac{2m|E|}{\hbar^2}}; & E < 0. \end{cases}$$

Interpreted physically, the solutions with $E > 0$ represent free, oscillating plane wave states, while the solutions with $E < 0$ represent bound states.

- For precisely defined energies $E = E_n$, the bound state solutions with $E < 0$ will satisfy $D_+ = 0$, corresponding to bound eigenstates of the form

$$u_n(r) = D e^{-\kappa r}.$$

6.3.1   Initial Derivation: Free and Bound States for $r \to \infty$

- For review, the radial eigenvalue equation, with V_{eff} written out in full, is

$$-\frac{\hbar^2}{2m}u''(r) + \left(V(r) + \frac{l(l+1)\hbar^2}{2mr^2}\right)u(r) = Eu(r).$$

For potentials $V(r)$ obeying the limit

$$\lim_{r \to \infty} V(r) = 0,$$

both the potential $V(r)$ and the centrifugal term $\frac{l(l+1)\hbar^2}{2mr^2}$ grow neglible for large r, and the radial eigenvalue equation reduces to

$$-\frac{\hbar^2}{2m}u''(r) = Eu(r).$$

- If $E > 0$, the equation is solved by the oscillating solution

$$u(r) = C_- e^{-ikr} + C_+ e^{ikr}, \qquad k = \sqrt{\frac{2mE}{\hbar^2}}.$$

These solutions represent free scattering states and have a continuous spectrum of energy eigenvalues. Note that each energy has degeneracy two, since there exist two linearly independent eigenfunctions $u(r)$ at a given value of k.

- If $E < 0$, the equation in the limit of large r is solved by the exponential solution

$$u(r) = D_- e^{-\kappa r} + D_+ e^{\kappa r}, \qquad \kappa = \sqrt{\frac{2m|E|}{\hbar^2}}.$$

At precisely defined energy eigenvalues $E = E_n$ we have $D_+ = 0$, and the corresponding solutions

$$u_n(r) = D_n e^{-\kappa r}$$

represent bound states. The discrete eigenvaleus E_n are degenerate, since the corresponding eigenfunction $u_n(r)$ obeys the non-degeneracy theorem.

6.3.2 Formal Analysis: Validity of the $r \to \infty$ Solutions

- We now consider more formally how fast the potential $V(r)$ must fall as r approaches infinity to justify the free and bound state general solutions

$$u_{\text{free}}(r) = C_- e^{-ikr} + C_+ e^{ikr} \qquad \text{and} \qquad u_{\text{bound}}(r) = D_- e^{-\kappa r} + D_+ e^{\kappa r}.$$

Bound States with $E < 0$

- We first consider the bound state with $E < 0$ and write the solution in the form

$$u(r) = v(r)e^{\pm \kappa r}.$$

We substitute this expression for $u(r)$ into the full radial eigenvalue equation to get

$$-\frac{\hbar^2}{2m}\left[v''(r) \pm 2\kappa v'(r) + \kappa^2 v(r)\right]e^{\pm \kappa r} + V_{\text{eff}}(r)v(r)e^{\pm \kappa r} = Ev(r)e^{\pm \kappa r}.$$

We then cancel $e^{\pm i\kappa r}$ from the equation, multiply through by $\frac{2m}{\hbar^2}$, and recognize that $\kappa^2 = \sqrt{\frac{-2mE}{\hbar^2}}$ (recall $E < 0$) cancels with $\frac{2mE}{\hbar^2}$ to get

$$v''(r) \pm \kappa v'(r) - \frac{2m}{\hbar^2}V_{\text{eff}}(r)v(r) = 0.$$

- Since $v(r)$ is just a correction to $e^{\pm \kappa r}$, we assume $v(r)$ changes slowly with r and neglect the second derivative $v''(r)$. We're left with

$$\pm \kappa v'(r) = \frac{2m}{\hbar^2} V_{\text{eff}}(r) v(r) \qquad \text{or} \qquad \frac{v'(r)}{v(r)} = \pm \frac{m}{\kappa \hbar^2} V_{\text{eff}}(r).$$

This is a first-order equation with separable variables, which we can integrate, as in

$$\int \frac{\mathrm{d}v}{v} = \pm \frac{m}{\kappa \hbar^2} \int V_{\text{eff}}(r) \, \mathrm{d}r,$$

to get

$$v(r) = v(r_0) \exp\left(\pm \frac{m}{\kappa \hbar^2} \int_{r_0}^{r} V_{\text{eff}}(\tilde{r}) \, \mathrm{d}\tilde{r} \right),$$

where r_0 is a sufficiently large value of r that V_{eff} decays slowly.

- From the above expression for $v(r)$, we see that the bound state ansatz

$$u_{\text{bound}}(r) = D_- e^{-\kappa r} + D_+ e^{\kappa r} = v(r) e^{\pm \kappa r}$$

is valid as long as $v(r)$ converges to a constant value as r approaches infinity. This holds when the limit

$$\lim_{r \to \infty} \int_{r_0}^{r} V(\tilde{r}) \, \mathrm{d}\tilde{r}$$

converges, which occurs when

$$\lim_{r \to \infty} r V(r) = 0$$

Note that the centrifugal component of V_{eff} falls with r^{-2} and is not problematic. To summarize, the bound state ansatz $u_{\text{bound}}(r) = D_- e^{-\kappa r} + D_+ e^{\kappa r}$ is valid for potentials for which $rV(r)$ vanishes at infinity.

- The limiting case at which the bound state condition

$$\lim_{r \to \infty} r V(r) = 0$$

no long holds is potentials of the form $V(r) = -\lambda/r$, for which we have

$$\lim_{r \to \infty} \int_{r_0}^{r} V(\tilde{r}) \, \mathrm{d}\tilde{r} = \lim_{r \to \infty} \left(-\lambda \ln \frac{r}{r_0} \right) \to \infty.$$

The corresponding solution for $u(r)$ in this limiting case is

$$u(r) \to v(r) e^{\pm \kappa r} = e^{\pm \kappa r} \exp\left(\mp \nu \ln \frac{r}{r_0} \right), \qquad \nu = \frac{m\lambda}{\kappa \hbar^2}.$$

Canceling the exponent and logarithm shows the bound state solutions fall as

$$u(r) \sim r^{\nu} e^{-\kappa r}.$$

Free Scattering States with $E > 0$

- We now consider the free scattering states with positive energy, for which we assumed the general solution

$$u_{\text{free}}(r) = C_- e^{-ikr} + C_+ e^{ikr}, \qquad k = \sqrt{\frac{2mE}{\hbar^2}}.$$

- Because the general solutions for the free and bound states are so similar, differing only by the presence of the imaginary unit i in the exponent and the replacement of κ with k, we would follow an analogous procedure to the above analysis of the bound states. To avoid repeating the same procedure, we simply quote the result:

  As before, to justify the exponential free state ansatz, the potential $V(r)$ must obey

  $$\lim_{r \to \infty} rV(r) = 0.$$

  The corresponding free state solutions are

  $$u(r) \to e^{\pm ikr} \exp\left(\mp i\nu \ln \frac{r}{r_0}\right), \qquad \nu = \frac{m\lambda}{k\hbar^2},$$

  and decay asymptotically as

  $$u(r) \sim r^\nu e^{-ikr}.$$

6.4 The Coulomb Potential

Discuss the general solution procedure for the quantum-mechanical problem of a bound electron in a Coulomb potenial. State and derive the relevant radial eigenvalue equation

State and derive the electron's energy eigenvalues.

Discuss energy degeneracy, the energy eigenfunctions, and the problem's classical limit.

- For an electron in a Coulomb potential, the radial eigenvalue equation reads

$$-\frac{\hbar^2}{2m_{\text{e}}} u''(r) + \left(\frac{l(l+1)\hbar^2}{2m_{\text{e}}r^2} - \frac{e_0^2}{4\pi\epsilon_0 r}\right) u(r) = Eu(r)$$

  This problem is physically signficant as the foundation for the quantum analysis of the hydrogen atom problem.

- The problem of an electron in a Coulomb potential is solved with the quantities

$$\kappa = \sqrt{\frac{2m_{\text{e}}E}{\hbar^2}} \qquad \rho = \kappa r \qquad \rho_0 = \frac{m_{\text{e}}e_0^2}{2\pi\epsilon_0 \kappa \hbar^2}$$

  and the ansatz

$$u(\rho) = \rho^{l+1} v(\rho) e^{-\rho}.$$

  In terms of these quantities, the Coulomb potential radial equation is

$$\rho v'' + 2(l+1-\rho)v' + \big[\rho_0 - 2(l+1)\big]v = 0.$$

- The energy eigenvalues of an electron in a Coulomb potential are

$$E_n = -\frac{m_e}{2\hbar^2}\left(\frac{e_0^2}{4\pi\epsilon_0}\right)^2\frac{1}{n^2} \equiv -\frac{\text{Ry}}{n^2}, \quad n = 1, 2, 3, \ldots,$$

where we have defined the Rydberg energy unit

$$1\,\text{Ry} = \frac{m_e}{2\hbar^2}\left(\frac{e_0^2}{4\pi\epsilon_0}\right)^2 = |E_1| = 13.6\,\text{eV}.$$

- Neglecting spin, the energy degeneracy N_degen of an energy level E_n is

$$N_\text{degen} = \sum_{l=0}^{n-1}(2l+1) = n^2,$$

where n and l are the principle and orbitral quantum numbers, respectively.

- The energy eigenfunctions for an electron in a Coulomb potential are

$$\Psi_{nlm}(\boldsymbol{r}) = \psi_{nl}(r)Y_l^m(\theta, \phi),$$

where Y_l^m are the spherical harmonics and the radial component is

$$\psi_{nl} = \sqrt{\left(\frac{2}{na_0}\right)^3\frac{(n-l-1)!}{2n(n+l)!}} \cdot \left(\frac{2r}{na_0}\right)^l \cdot L_{n-l-1}^{2l+1}\left(\frac{2r}{na_0}\right) \cdot e^{-\frac{r}{na_0}},$$

where we have introduced the Bohr radius

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e_0^2} = 0.053\,\text{nm}.$$

- The semi-classical limit for the energy eigenvalues is found with the Wilson-Sommerfeld quantization condition

$$\frac{1}{2\pi}\oint p\,\mathrm{d}q = n\hbar, \qquad n \in \mathbb{Z},$$

where p and q are a system's momentum and coordinates. Applying this quantization condition to the electron results in quantized values of angular momentum L_z, and combining the result with Newton's law reproduces the Bohr energy formula

$$E_n = -\frac{m_e}{2\hbar^2}\left(\frac{e_0^2}{4\pi\epsilon_0}\right)^2\frac{1}{n^2} \equiv -\frac{\text{Ry}}{n^2}, \quad n = 1, 2, 3, \ldots.$$

This is derived at the end of this subsection.

- In the classical limit, the electron orbits the center of Coulomb attraction (e.g. the proton in a hydrogen atom) in a circular orbit. This behaviour agrees with the quantum-mechanical solution in the limit of large angular momentum, where

$$n \gtrsim l \gg 1 \qquad \text{and} \qquad l \gtrsim m \gg 1.$$

In this limit, the relative uncertainty in orbital radius r is

$$\lim_{n\to\infty}\frac{\Delta r}{\langle r\rangle} = \lim_{n\to\infty}\frac{1}{\sqrt{2n+1}} = 0,$$

which approaches a perfectly spherical orbit of radius $\langle r\rangle$.

6.4.1   Derivation: The Radial Equation For a Coulomb Potential

- We aim to find energy levels of an electron with charge e_0 and mass m_e in a Coulomb potential, for which the radial eigenvalue equation reads

$$-\frac{\hbar^2}{2m_e}u''(r) + \left(\frac{l(l+1)\hbar^2}{2m_e r^2} - \frac{e_0^2}{4\pi\epsilon_0 r}\right)u(r) = Eu(r)$$

An electron in a Coulomb potential is the basis for solving the problem of the hydrogen atom.

- We first introduce the dimensionless coordinate $\rho = \kappa r$, where $\kappa = \sqrt{\frac{2m_e E}{\hbar^2}}$. In this case the equation simplifies to

$$\left(-\frac{d^2}{d\rho^2} + \frac{l(l+1)}{\rho^2} - \frac{m_e e_0^2}{2\pi\epsilon_0 \kappa \hbar^2}\frac{1}{\rho}\right)u(\rho) = -u(\rho)$$

Finally, in terms of ρ_0, we have

$$u'' - \frac{l(l+1)}{\rho^2}u + \frac{\rho_0}{\rho}u - u = 0, \qquad \rho_0 = \frac{m_e e_0^2}{2\pi\epsilon_0 \kappa \hbar^2}$$

- We proceed with the ansatz

$$u(\rho) = \rho^{l+1}v(\rho)e^{-\rho},$$

which is intended to model the localized behavior of bound states for large ρ.

- As an intermediate step, the ansatz's first and second derivatives are

$$u' = \rho^l e^{-\rho}\big[(l+1-\rho)v + \rho v'\big]$$
$$u'' = \rho^l e^{-\rho}\left\{\left[-2l - 2 + \rho + \frac{l(l+1)}{\rho}\right]v + 2(l+1-\rho)v' + \rho v''\right\}.$$

We substitute then substitute u and u'' into the dimensionless radial eigenvalue equation. After some tedious but straightforward algebra involving combining like terms and dividing through by $\rho^l e^{-\rho}$ we get the equation

$$\rho v'' + 2(l+1-\rho)v' + \big[\rho_0 - 2(l+1)\big]v = 0.$$

Note that this equation contains only $v(\rho)$.

6.4.2   Derivation: Energy Eigenvalues in a Coulomb Potential

- We solve the radial equation for $v(\rho)$ with the Frobenius method. This involves writing $v(\rho)$ as a power series, i.e.

$$v(\rho) = \sum_{k=0}^{\infty} c_k \rho^k.$$

The plan is to find v's first two derivatives, substitute the power series ansatz into the equation for v, cancel like terms, and find a recursion relation for the coefficients. The first two derivatives are

$$v' = \sum_{k=0}^{\infty} k c_k \rho^{k-1} \stackrel{k \to k+1}{=} \sum_{k=0}^{\infty} (k+1) c_{k+1} \rho^k$$

$$v'' = \sum_{k=0}^{\infty} k(k+1) c_{k+1} \rho^{k-1}.$$

Note that we have shifted the index from k to $k+1$, which is shown explicitly for the first derivative v' and left implicit for v''.

- We then substitute the power series ansatz expressions into the equation for v to get

$$\rho \sum_{k=0}^{\infty} k(k+1) c_k \rho^{k-1} + 2(l+1-\rho) \sum_{k=0}^{\infty} (k+1) c_{k+1} \rho^k$$

$$+ \big[\rho_0 - 2(l+1)\big] \sum_{k}^{\infty} c_k \rho^k = 0.$$

Next, we distribute coefficients and re-index the $2(l+1)$ term from $k+1$ to k to get

$$\sum_{k=0}^{\infty} k(k+1) c_k \rho^k + 2(l+1) \sum_{k=0}^{\infty} (k+1) c_{k+1} \rho^k - 2 \sum_{k=1}^{\infty} k c_k \rho^k$$

$$+ \big[\rho_0 - 2(l+1)\big] \sum_{k=0}^{\infty} c_k \rho^k = 0.$$

For the equation to hold, the coefficients of ρ^k at a given k must be equal, implying

$$\big[k(k+1) + 2(l+1)(k+1)\big] c_{k+1} = \big[2k - (\rho_0 - 2(l+1))\big] c_k.$$

We then rearrange the above equation to get the recursive relation

$$c_{k+1} = \frac{2(k+l+1) - \rho_0}{(k+1)(k+2l+2)} c_k.$$

- For large k, i.e. $k \gg l, \rho_0$, the above recursion relation reduces to the power series coefficient relationship for the exponential function, i.e.

$$\frac{c_k}{c_{k-1}} \to \frac{2}{k} \qquad \text{or} \qquad c_k = \frac{2^k}{k!} c_0,$$

where we have re-indexed the first term by one. With the coefficients c_k known (at least for large k) the solution for $v(\rho)$ is

$$v(\rho) = \sum_{k=0}^{\infty} c_k \rho^k = c_0 \sum_{k=0}^{\infty} \frac{1}{k!} (2\rho)^k = c_0 e^{2\rho}.$$

- In terms of the large k solution $v(\rho) \sim e^{2\rho}$, which corresponds to the asymptotic behavior of $v(\rho)$ for large ρ, the solution for $u(\rho)$ is

$$u(\rho) = \rho^{l+1} v(\rho) e^{-\rho} = \rho^{l+1} e^{2\rho} e^{-\rho} = \rho^{l+1} e^{\rho}.$$

The relationship $u(\rho) \sim \rho^{l+1} e^{\rho}$ does not in general converge for large ρ. In fact, $u(\rho)$ is convergent only if the series ansatz for $v(\rho)$, i.e.

$$v(\rho) = \sum_{k=0}^{\infty} c_k \rho^k,$$

truncates at a finite $k_{\max} \geq 0$. Truncating at $k_{\max}$ implies $c_k = 0$ for $k > k_{\max}$. If we return to the recursive coefficient relation, i.e.

$$c_{k+1} = \frac{2(k+l+1) - \rho_0}{(k+1)(k+2l+2)} c_k,$$

the condition $c_k = 0$ for $k > k_{\max}$ implies

$$2(k_{\max} + l + 1) - \rho_0 = 0 \implies \rho_0 = 2(k_{\max} + l + 1) \in \mathbb{N}.$$

In other words, ρ_0 must be integer-valued to satisfy the convergence of $v(\rho)$ and thus $u(\rho)$ for large ρ.

- With this integer restriction on ρ_0 in mind, we define the principle quantum number

$$n \equiv k_{\max} + l + 1,$$

which implies $\rho_0 = 2n$. In terms of $\rho_0 = 2n$ and the earlier equations

$$\rho_0 = \frac{m_e e_0^2}{2\pi\epsilon_0 \kappa \hbar^2} \qquad \text{and} \qquad \kappa = \sqrt{\frac{2m_e E}{\hbar^2}},$$

the energy eigenvalues of an electron in a Coulomb potential are thus

$$E_n = -\frac{m_e}{2\hbar^2} \left(\frac{e_0^2}{4\pi\epsilon_0}\right)^2 \frac{1}{n^2} \equiv -\frac{\text{Ry}}{n^2}, \quad n = 1, 2, 3, \ldots,$$

where we have defined the Rydberg energy unit

$$1\,\text{Ry} = \frac{m_e}{2\hbar^2} \left(\frac{e_0^2}{4\pi\epsilon_0}\right)^2 = |E_1| = 13.6\,\text{eV}.$$

6.4.3   Discussion: Eigenfunctions and Degeneracy in a Coulomb Potential

- First, we consider energy eigenvalue degeneracy. For each value of E_n, which depends only on the principle quantum number n, their exist n values of the orbital quantum number $l = 0, 1, \ldots, n - 1$. The complete wavefunction

$$\Psi(\boldsymbol{r}) = \psi(r) Y_l^m(\theta, \phi)$$

is thus n-times degenerate with respect to the radial component $\psi(r)$, since n values of l correspond to the same energy E_n.

Additionally, the energy eigenvalues are degenerate with respect to the quantum number m, which corresponds to the projection of angular momentum onto the z axis. Since E_n does not depend on m and m can assume $2l + 1$ values from $-l$ to l, there are $2l + 1$ states proportional to Y_l^m with energy E_n at a given l.

Considering both the degeneracy with respect to both l and m, the total degeneracy of a given energy level E_n is

$$\sum_{l=0}^{n-1}(2l+1) = n^2.$$

In other words, the energy level E_n has degeneracy n^2.

- Next, we return to the series for $v(\rho)$, i.e.

$$v(\rho) = \sum_{k=0}^{k_{\max}} c_k \rho^k,$$

where we have made the truncation at $k_{\max}$ explicit. Since $k_{\max} = n - l - 1$, the function $v(\rho)$ is a polynomial of order $n - l - 1$. Without derivation, it turns out that $v(\rho)$ takes the form of an associated Laguerre polynomial, i.e.

$$v(\rho) \propto L_{k_{\max}}^{2l+1}(2\rho) = L_{n-l-1}^{2l+1}.$$

Note that $v(\rho)$ is indexed by both l and n.

- The complete wavefunction is thus thus

$$\Psi_{nlm}(\boldsymbol{r}) = \psi_{nl}(r)Y_l^m(\theta, \phi),$$

where the radial component is

$$\psi_{nl} = \sqrt{\left(\frac{2}{na_0}\right)^3 \frac{(n-l-1)!}{2n(n+l)!}} \cdot \left(\frac{2r}{na_0}\right)^l \cdot L_{n-l-1}^{2l+1}\left(\frac{2r}{na_0}\right) \cdot e^{-\frac{r}{na_0}},$$

where we have introduced the Bohr radius

$$a_0 = \frac{4\pi\epsilon_0 \hbar^2}{m_e e_0^2} = 0.053\,\text{nm}.$$

- The wavefunctions Ψ_{nlm} are conveniently orthonormal, i.e.

$$\int \Psi_{n'l'm'}^*(\boldsymbol{r})\Psi_{nlm}(\boldsymbol{r})\,\mathrm{d}^3\boldsymbol{r} = \langle n'l'm' | nlm \rangle = \delta_{n'n}\delta_{l'l}\delta_{m'm}.$$

6.4.4   Semi-Classical and Classical Limits

The Energy Eigenvalues

- The angular momentum of the electron around the hydrogen nucleus is quantized according to the Wilson-Sommerfeld quantization condition

$$\frac{1}{2\pi} \oint p\,\mathrm{d}q = n\hbar, \qquad n \in \mathbb{Z},$$

where p and q are a system's momentum and coordinates.

- For an electron on a hypothetical circular orbit of radius r at speed v about the nucleus, the integral reads

$$n\hbar \equiv \frac{1}{2\pi} \oint (m_{\mathrm{e}} v)\, \mathrm{d}q = \frac{1}{2\pi} m_{\mathrm{e}} v (2\pi r) = m_{\mathrm{e}} r v = L_z,$$

which produces the angular momentum quantization condition $L_z = n\hbar$.

- Combining the quantization $n\hbar = L_z = m_{\mathrm{e}} r v$ with Newton's law

$$F = m_{\mathrm{e}} a = \frac{m_{\mathrm{e}} v^2}{r} = \frac{e_0^2}{4\pi\epsilon_0 r^2},$$

reproduces the Bohr energy formula

$$E_n = -\frac{m_{\mathrm{e}}}{2\hbar^2} \left(\frac{e_0^2}{4\pi\epsilon_0} \right)^2 \frac{1}{n^2} \equiv -\frac{\mathrm{Ry}}{n^2}, \quad n = 1, 2, 3, \ldots$$

The Eigenfunctions and Shape of the Orbit

- Next, we consider the eigenfunctions $|nlm\rangle$, which do not in general correspond to a uniform circular orbit of the electron about the nucleus.

  Circular orbits correspond only to solutions with large angular momentum, for which the quantum numbers obey

$$n \gtrsim l \gg 1 \qquad \text{and} \qquad l \gtrsim m \gg 1.$$

  For maximum possible angular momentum, i.e. $l = n - 1$, the associated Laguerre polynomial is $L_{n-l-1}^{2l+1} = L_0^{2l+1} = 1$ and the corresponding radial eigenfunction is

$$\psi_{n,l}(r) = \psi_{n,n-1}(r) = 2^n \left[n^4 (2n-1)! a_0^3 \right]^{-1/2} \left(\frac{r}{na_0} \right)^{n-1} e^{-\frac{r}{na_0}}.$$

- The expectation values of r and r^2 for the above radial function are

$$\langle r \rangle = \int_0^\infty r \psi_{n,n-1}^2(r) r^2 \, \mathrm{d}r = n \left(n + \frac{1}{2} \right) a_0$$

$$\langle r^2 \rangle = \cdots = n^2 (n+1) \left(n + \frac{1}{2} \right) a_0^2,$$

and the corresponding uncertainty in r is

$$\Delta r = \sqrt{\langle r^2 \rangle - \langle r \rangle^2} = \frac{n}{2} \sqrt{2n+1}\, a_0.$$

In the limit of large n, the relative uncertainty in radius is

$$\lim_{n \to \infty} \frac{\Delta r}{\langle r \rangle} = \lim_{n \to \infty} \frac{1}{\sqrt{2n+1}} = 0.$$

In other words, the orbit approaches a spherical shell with radius $\langle r \rangle$ for large n, in agreement with the classical limit of a circularly orbiting electron.

7 Charged Particle in EM Field

7.1 The Schrödinger Equation for a Charged Particle in an EM Field

State and derive the Schrödinger equation for a charged particle in an electromagnetic field. Be sure to consider the simplification resulting from using the Coulomb gauge.

- The Schrödinger equation for a particle of charge q in electromagnetic field with electrostatic potential ϕ and magnetic vector potential $\boldsymbol{A}$ is

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + i\frac{\hbar q}{m}\boldsymbol{A}\cdot\nabla\Psi + \left(i\frac{\hbar q}{2m}(\nabla\cdot\boldsymbol{A}) + \frac{q^2}{2m}\boldsymbol{A}^2 + q\phi\right)\Psi.$$

- In the Coulomb gauge $\nabla\cdot\boldsymbol{A} = 0$, the Schrödinger equation reduces to

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + i\frac{\hbar q}{m}\boldsymbol{A}\cdot\nabla\Psi + \left(\frac{q^2}{2m}\boldsymbol{A}^2 + q\phi\right)\Psi$$

Derivation: The Schrödinger Equation for a Particle in an EM Field

- We analyze a particle of charge q and mass m in an electric field in terms of the electric potential ϕ and magnetic potential $\boldsymbol{A}$:

$$\boldsymbol{B} = \nabla\times\boldsymbol{A} \qquad \text{and} \qquad \boldsymbol{E} = -\nabla\phi - \frac{\partial\boldsymbol{A}}{\partial t}$$

Without derivation, the particle's Hamiltonian is

$$H = \frac{(\boldsymbol{p} - q\boldsymbol{A})^2}{2m} + q\phi.$$

- In terms of this Hamiltonian, the Schrödinger equation reads

$$i\hbar\frac{\partial\Psi}{\partial t} = \frac{(\boldsymbol{p} - q\boldsymbol{A})^2}{2m}\Psi + q\phi\Psi,$$

which we multiply out to get

$$i\hbar\frac{\partial\Psi}{\partial t} = \frac{1}{2m}\left(-i\hbar\nabla - q\boldsymbol{A}\right)^2\Psi + q\phi\Psi$$
$$= -\frac{\hbar^2}{2m}\nabla^2\Psi + \frac{i\hbar q}{2m}(\nabla\cdot\boldsymbol{A} + \boldsymbol{A}\cdot\nabla)\Psi + \frac{q^2}{2m}\boldsymbol{A}^2\Psi + q\phi\Psi.$$

- We then make the intermediate calculation

$$(\nabla\cdot\boldsymbol{A})\Psi + \boldsymbol{A}\cdot\nabla\Psi = \nabla\cdot(\Psi\boldsymbol{A}) + \boldsymbol{A}\cdot\nabla\Psi = \boldsymbol{A}\cdot\nabla\Psi + \Psi\nabla\cdot\boldsymbol{A} + \boldsymbol{A}\cdot\nabla\Psi$$
$$= 2\boldsymbol{A}\cdot\nabla\Psi + \Psi\nabla\cdot\boldsymbol{A},$$

in terms of which the Schrödinger equation simplifies to the desired expression

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + i\frac{\hbar q}{m}\boldsymbol{A}\cdot\nabla\Psi + \left(i\frac{\hbar q}{2m}(\nabla\cdot\boldsymbol{A}) + \frac{q^2}{2m}\boldsymbol{A}^2 + q\phi\right)\Psi.$$

- In the Coulomb gauge $\nabla\cdot\boldsymbol{A} = 0$, the Schrödinger equation reduces further to

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + i\frac{\hbar q}{m}\boldsymbol{A}\cdot\nabla\Psi + \left(\frac{q^2}{2m}\boldsymbol{A}^2 + q\phi\right)\Psi.$$

7.2 The Normal Zeeman Effect

Define and interpret the magnetic dipole moment operator and the Bohr magneton. Discuss the coupling of magnetic moment to a homogeneous external magnetic field, defined the Zeeman coupling term, and show that in a homogeneous field the quadratic coupling term is usually negligible in comparison to the Zeeman term. Explain and discuss the normal Zeeman effect.

- The magnetic dipole moment operator $\boldsymbol{\mu}$ and Bohr magnetic μ_B are defined as

$$\boldsymbol{\mu} = \frac{q}{2m}\boldsymbol{L} \qquad \text{and} \qquad \mu_B = \frac{e_0 \hbar}{2m_{\mathrm{e}}}.$$

  The Bohr magneton is important: it is the quantum of magnetic dipole moment, just like e_0 is the quantum of electric charge.

- The dominant coupling interaction between the particle and the magnetic field is encoded by the Schrödinger equation term

$$i\frac{\hbar q}{m}\boldsymbol{A}\cdot\nabla\Psi.$$

  For a homogeneous magnetic field, the linear term can be written

$$i\frac{\hbar q}{m}\boldsymbol{A}\cdot\nabla = -\frac{q}{2m}\boldsymbol{L}\cdot\boldsymbol{B} = -\boldsymbol{\mu}\cdot\boldsymbol{B} \equiv H_{\mathrm{Zeeman}},$$

  and is often called the Zeeman coupling term.

- For a homogeneous magnetic field, the quadratic coupling term $\frac{q^2}{2m}\boldsymbol{A}^2$ reads

$$\frac{q^2}{2m}\boldsymbol{A}^2 = \frac{q^2 B^2}{8m}(x^2 + y^2).$$

  For an electron with magnetic moment μ_B, charge $q = e_0$ and characteristic distance scale $x^2 + y^2 \sim a_0$ in a magnetic field $B = 1\,\mathrm{T}$, we have

$$\frac{e_0 B^2}{8m_{\mathrm{e}}}a_0^2 \Big/ (\mu_B B) \sim 10^{-6}.$$

  In other words, the quadratic coupling term is completely negligible relative to the Zeeman term $\boldsymbol{\mu}\cdot\boldsymbol{B}$.

- The normal Zeeman effect refers to the splitting of the electron's energy levels in a hydrogen atom exposed to a weak, homogeneous magnetic field according to

$$H\ket{nlm} = \left(-\frac{\mathrm{Ry}}{n^2} + m\mu_B B\right)\ket{nlm}.$$

  Note that each energy level now depends on m as well as n, which reduces (but does not fully remove) energy level degeneracy, which still exists with respect to l.

Derivation: Zeeman Coupling

- Recall the Schrödinger equation for a charged particle in an EM field reads

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + i\frac{\hbar q}{m}\boldsymbol{A}\cdot\nabla\Psi + \left(\frac{q^2}{2m}A^2 + q\phi\right)\Psi.$$

- A homogeneous magnetic field $\boldsymbol{B} = (0, 0, B)$, corresponds to the vector potential

$$A = -\frac{1}{2} \boldsymbol{r} \times \boldsymbol{B}.$$

  For the above choice of vector potential, the dominant magnetic coupling term reads

$$
\begin{aligned}
i\frac{\hbar q}{m} \boldsymbol{A} \cdot \nabla \Psi &= -i\frac{\hbar q}{2m}(\boldsymbol{r} \times \boldsymbol{B}) \cdot \nabla \Psi = i\frac{\hbar q}{2m}(\boldsymbol{r} \times \nabla) \cdot \boldsymbol{B}\Psi \\
&= \frac{q}{2m}\big[\boldsymbol{r} \times (i\hbar\nabla)\big] \cdot \boldsymbol{B}\Psi = -\frac{q}{2m}(\boldsymbol{r} \times \boldsymbol{p}) \cdot \boldsymbol{B}\Psi \\
&= -\frac{q}{2m}\boldsymbol{L} \cdot \boldsymbol{B}\Psi.
\end{aligned}
$$

- In terms of magnetic dipole moment $\boldsymbol{\mu} = \frac{q}{2m}\boldsymbol{L}$, the above coupling term between a particle and an external magnetic field reads

$$i\frac{\hbar q}{m} \boldsymbol{A} \cdot \nabla = -\frac{q}{2m}\boldsymbol{L} \cdot \boldsymbol{B} = -\boldsymbol{\mu} \cdot \boldsymbol{B}$$

  This coupling term is called the normal Zeeman coupling term, defined as

$$H_{\text{Zeeman}} = -\boldsymbol{\mu} \cdot \boldsymbol{B}.$$

The Quadratic Coupling Term

- The quadratic coupling term $\frac{q^2}{2m}\boldsymbol{A}^2$, again using the vector potential $\boldsymbol{A} = -\frac{1}{2}\boldsymbol{r} \times \boldsymbol{B}$ and homogeneous magnetic field $\boldsymbol{B} = (0, 0, B)$, reads

$$
\begin{aligned}
\frac{q^2}{2m}\boldsymbol{A} \cdot \boldsymbol{A} &= \frac{q^2}{2m}\left[\left(-\frac{1}{2}\boldsymbol{r} \times \boldsymbol{B}\right) \cdot \left(-\frac{1}{2}\boldsymbol{r} \times \boldsymbol{B}\right)\right] = \frac{q^2}{8m}\big[B^2 r^2 - (\boldsymbol{B} \cdot \boldsymbol{r})^2\big] \\
&= \frac{q^2}{8m}\Big\{B^2 r^2 - \big[(0, 0, B) \cdot (x, y, z)\big]^2\Big\} = \frac{q^2}{8m}\big(B^2 r^2 - B^2 z^2\big) \\
&= \frac{q^2 B^2}{8m}(x^2 + y^2)
\end{aligned}
$$

  Note that the value of this quadratic coupling term depends on the choice of the gauge for $\boldsymbol{A}$, and the above result holds only for the Coulomb gauge.

- In terms of the above quadratic coupling term and the linear Zeeman coupling term, the Schrödinger equation for a particle in a homogeneous magnetic field reads

$$H = \frac{p^2}{2m} - \boldsymbol{\mu} \cdot \boldsymbol{B} + \frac{q^2 B^2}{8m}(x^2 + y^2) + q\phi.$$

Discussion: The Normal Zeeman Effect

- We begin with the Hamiltonian for a charged particle in a magnetic field, and, assuming a weak magnetic field, neglect the quadratic coupling term to get

$$H \approx \frac{p^2}{2m} - \boldsymbol{\mu} \cdot \boldsymbol{B} + q\phi.$$

We can find the wavefunction reusing the results from the central potential chapter, where the general wavefunction reads

$$\Psi_{nlm}(\boldsymbol{r}) = \psi_{nl}(r)Y_l^m(\theta, \varphi),$$

and Y_l^m and ψ_{nl} are the spherical harmonics and radial eigenfunctions, respectively. Although this eigenfunction was originally derived in the context of angular momentum $\boldsymbol{L}$, since $\boldsymbol{\mu}$ and $\boldsymbol{L}$ commute (recall the relationship $\boldsymbol{\mu} = \frac{q}{2m}\boldsymbol{L}$), eigenfunctions of $\boldsymbol{L}$ are also eigenfunctions of $\boldsymbol{\mu}$.

- Recall that in the absence of a magnetic field, considering only the electrostatic Coulomb potential $\phi(r)$, the hydrogen atom's energy eigenfunctions obey

$$H\left|nlm\right\rangle = -\frac{\mathrm{Ry}}{n^2}\left|nlm\right\rangle.$$

In the presence of a homogeneous magnetic field and the Zeeman coupling term $-\boldsymbol{\mu} \cdot \boldsymbol{B}$, the energy levels split according

$$H\left|nlm\right\rangle = \left(-\frac{\mathrm{Ry}}{n^2} + m\mu_B B\right)\left|nlm\right\rangle,$$

which is the equation quoted at the begining of this subsection.

7.3 Landau Levels

Explain Landau quantization and Landau levels and give a physical interpretation. Define the Landau gauge potential. Restrict your analysis to the xy plane.

- The following is from Wikipedia, which does a good job and I see no reason to change:

  > Landau quantization is the quantization of the cyclotron orbits of charged particles in magnetic fields. Because of this quantization, charged particles can only occupy orbits with discrete energy values, called Landau levels.
  >
  > The Landau levels are degenerate, with the number of electrons per level directly proportional to the strength of the applied magnetic field.

- The Landau gauge potential is used generate a homogeneous magnetic field of the form $\boldsymbol{B} = B\,\hat{\mathbf{z}}$. The two possible gauge potentials generating this magnetic field are

$$\boldsymbol{A} = xB\,\hat{\mathbf{y}} \qquad \text{and} \qquad \boldsymbol{A} = -yB\,\hat{\mathbf{x}}.$$

Both recover $\boldsymbol{B} = B\,\hat{\mathbf{z}}$ via $\boldsymbol{B} = \nabla \times \boldsymbol{A}$; we will work with the latter, i.e. $\boldsymbol{A} = -yB\,\hat{\mathbf{x}}$.

- The cyclotron orbit problem is solved with the ansatz

$$\Psi(\boldsymbol{r}) = \exp\left[i\left(\frac{p_x}{\hbar}x + \frac{p_z}{\hbar}z\right)\right]\chi(y),$$

and the cyclotron frequency ω, magnetic length ξ, wave vector $k = p_x/\hbar$ and displacement y_k, which are defined as

$$\omega = \frac{qB}{m} \qquad \xi = \sqrt{\frac{\hbar}{qB}} \qquad y_k = -\xi^2 k.$$

- For a particle confined to the xy plane in the absence of an electric potential, the solutions for χ are simply the eigenstates of a displaced harmonic oscillator, i.e.

$$\chi_{nk}(y) = \psi_n(y - y_k),$$

and the complete solution Ψ for a charged particle in a cyclotron orbit is thus

$$\Psi_{nk}(x, y) = e^{i\frac{p_x}{\hbar}x}\chi_{nk}(y) = \frac{1}{\sqrt{2\pi}}e^{ikx}\psi_n(y - y_k).$$

- The corresponding energy eigenvalues

$$E_n = \left(n + \tfrac{1}{2}\right)\hbar\omega, \qquad \omega = \frac{qB}{m},$$

are called Landau levels. These energy eigenvalues are highly degenerate, since a continuum of linearly independent eigenstates Ψ_{nk}, which can exist for any $k \in \mathbb{R}$, will have the same energy eigenvalue E_n.

- The Landau level eigenstates maintain their initial state after time evolution and the wavefunction's probability density $\rho = |\psi|^2$ does not change with time, i.e.

$$\rho(x, t) = \rho(x, 0).$$

This trivial time evolution is a consequence of the Landau levels E_n being independent of $k = p_x/\hbar$.

7.3.1 Derivation: Landau Levels

- We begin with the stationary Schrödinger equation for a charged particle in an electromagnetic field. In terms of the gauge potential $\boldsymbol{A} = -yB\,\hat{\mathbf{x}}$, this reads

$$H\Psi = \frac{(\boldsymbol{p} - q\boldsymbol{A})^2}{2m} + q\phi = \frac{1}{2m}\left(-i\hbar\nabla + qyB\right)^2\Psi + q\phi\Psi$$

$$= \frac{1}{2m}\left[\left(-i\hbar\frac{\partial}{\partial x} + qBy\right)^2 - \hbar^2\frac{\partial^2}{\partial y^2} - \hbar^2\frac{\partial^2}{\partial z^2}\right]\Psi + q\phi\Psi = E\Psi.$$

Next, we assume the electric potential depends only on the single coordinate y, ie. $\phi(\boldsymbol{r}) = \phi(y)$, and solve the equation with the ansatz

$$\Psi(\boldsymbol{r}) = \exp\left[i\left(\frac{p_x}{\hbar}x + \frac{p_z}{\hbar}z\right)\right]\chi(y).$$

The exponential term represents plane waves in the x and z directions.

- Next, as an intermediate step, we note that applying a function of the operator $\hat{p}_x$ to the operator's plane wave eigenfunction is the same as multiplying the plane wave by the momentum operator's eigenvalue p_x. In equation form, this reads.

$$\hat{f}(\hat{p}_x)e^{i\frac{p_x}{\hbar}x} \equiv \hat{f}\left(-i\hbar\frac{\partial}{\partial x}\right)e^{i\frac{p_x}{\hbar}x} = f(p_x)e^{i\frac{p_x}{\hbar}x}.$$

In our case, we use this identity to show

$$\left(-i\hbar\frac{\partial}{\partial x} + qBy\right)^2\exp\left[i\left(\frac{p_x}{\hbar}x + \frac{p_z}{\hbar}z\right)\right]\chi(y) = (p_x + qBy)^2\exp\left[i\left(\frac{p_x}{\hbar}x + \frac{p_z}{\hbar}z\right)\right]\chi(y),$$

where p_x on the right hand side is the eigenvalue of the operator $\hat{p}_x \to -i\hbar\frac{\partial}{\partial x}$ on the left hand side, and the corresponding function is $f(x) = x^2$. When applied to the Schrödinger equation using the above ansatz for Ψ, this identity produces

$$\left(-i\hbar\frac{\partial}{\partial x} + qBy\right)^2 \Psi(\boldsymbol{r}) \to (p_x + qBy)^2 \Psi(\boldsymbol{r}).$$

- We then substitute Ψ into the Schrödinger equation, apply the above identity, note that $\phi = \phi(y)$ acts only on the $\chi(y)$ factor of Ψ, evaluate the relevant momentum eigenvalue relations for x and z, and cancel out the common factor $\exp\left[i\left(\frac{p_x}{\hbar}x + \frac{p_z}{\hbar}z\right)\right]$, which considerably simplifies the Schrödinger equation to

$$\frac{1}{2m}(p_x + qBy)^2\chi(y) - \frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}y^2}\chi(y) + \frac{p_z^2}{2m}\chi(y) + q\phi(y)\chi(y) = E\chi(y).$$

Note that p_x and p_z are scalar eigenvalues and not operators.

- Next, since we consider only a particle restricted to the xy plane, we have $p_z = 0$, and the Schrödinger equation simplifies further to

$$\frac{1}{2m}\left[(p_x + qBy)^2 - \hbar\frac{\mathrm{d}^2}{\mathrm{d}y^2}\right]\chi(y) + V(y)\chi(y) = E\chi(y),$$

where we have defined $V(y) = q\phi(y)$ and used $p_z = 0$. We have thus reduced the problem to a one-dimensional problem involving only the coordinate y.

- Next, we define the cyclotron frequency ω and characteristic magnetic length ξ as

$$\omega = \frac{qB}{m} \qquad \text{and} \qquad \xi = \sqrt{\frac{\hbar}{qB}},$$

and introduce a wave vector k, allowing use to write $p_x = \hbar k$, and define the displacement

$$y_k = -\xi^2 k.$$

In terms of these new quantities, we can write the particle's Hamiltonian in the same form as a displaced harmonic oscillator with an additional potential $V(y)$, i.e.

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\chi(y)}{\mathrm{d}y^2} + \frac{m\omega^2}{2}(y - y_k)^2\chi(y) + V(y)\chi(y) = E\chi(y).$$

- In absence of an electric potential, and thus $\phi = V = 0$, the solutions of the above equation are simply the eigenstates of a displaced harmonic oscillator, i.e.

$$\chi_{nk}(y) = \psi_n(y - y_k).$$

Using the just-derived wave function χ_{nk}, the complete solution Ψ for a charged particle in a cyclotron orbit in the xy plane is then

$$\Psi_{nk}(x, y) = e^{i\frac{p_x}{\hbar}x}\chi_{nk}(y) = \frac{1}{\sqrt{2\pi}}e^{ikx}\psi_n(y - y_k),$$

where the factor $\sqrt{2\pi}$ is included for normalization. The corresponding eigenvalues

$$E_n = \left(n + \tfrac{1}{2}\right)\hbar\omega$$

are the Landau levels.

7.3.2 Derivation: Time Evolution of Landau Level Eigenstates

- Since the energy eigenvalue E_n is independent of k, the time evolution of a wavefunction expanded in the $\{\Psi_{nk}\}$ basis is trivial, since, because of the high degeneracy with respect to k, any linear combination of the stationary states is still a stationary state with the same energy.

- As an example, consider an arbitrary wavefunction ψ initially expanded in the plane wave basis, i.e.

$$\psi(x,0) = \frac{1}{\sqrt{2\pi}} \int \widehat{\psi}(k) e^{ikx} \, \mathrm{d}k.$$

The wavefunction maintains its shape for $t > 0$, since the time evolution reads

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} \int \widehat{\psi}(k) e^{ikx} e^{-i\frac{E}{\hbar}t} \, \mathrm{d}k = e^{-i\frac{E}{\hbar}t} \frac{1}{\sqrt{2\pi}} \int \widehat{\psi}(k) e^{ikx} \, \mathrm{d}k$$
$$= e^{-i\frac{E}{\hbar}t} \psi(x,0),$$

where we can move the time-dependent factor out of the integral since E does not depend of k. Thus, the wavefunction's probability density $\rho = |\psi|^2$ does not change with time, i.e.

$$\rho(x,t) = \rho(x,0),$$

since the factor $e^{-i\frac{E}{\hbar}t}$ vanishes when taking ψ's squared absolute value.

7.4 Gauge Transformations

Discuss gauge transformations in the context of a particle in an electromagnetic field.

- A so-called local gauge transformation, encoding a time-dependent phase shift, reads

$$\Psi'(\boldsymbol{r},t) = e^{i\delta(\boldsymbol{r},t)}\Psi(\boldsymbol{r},t).$$

- For particles in in an electromagnetic field, we define the transformed potentials as

$$\boldsymbol{A}' = \boldsymbol{A} + \nabla\Lambda \qquad \text{and} \qquad \phi' = \phi - \frac{\partial\Lambda}{\partial t},$$

which are chosen to preserve the magnetic and electric field $\boldsymbol{B}' = \boldsymbol{B}$ and $\boldsymbol{\mathcal{E}}' = \boldsymbol{\mathcal{E}}$.
We also define the phase term $\delta(\boldsymbol{r},t)$ in the local gauge transform to be

$$\delta(\boldsymbol{r},t) = \frac{q}{\hbar}\Lambda(\boldsymbol{r},t) \implies \Psi'(\boldsymbol{r},t) = e^{i\frac{q}{\hbar}\Lambda(\boldsymbol{r},t)}\Psi(\boldsymbol{r},t)$$

- Under the local, time-dependent phase shift transformation, the Schrödinger equation for a charged particle in an electromagnetic field, in terms of both the transformed wavefunction Ψ' and the original wavefunction Ψ read

$$i\hbar\frac{\partial}{\partial t}\Psi' = \frac{(-i\hbar\nabla - q\boldsymbol{A}')^2}{2m}\Psi' + q\phi'\Psi',$$

and, for the original wavefunction Ψ,

$$i\hbar\frac{\partial}{\partial t}\Psi = \frac{1}{2m}\sum_{\alpha}\left(-i\hbar\frac{\partial}{\partial x_\alpha} - qA'_\alpha + \hbar\frac{\partial\delta}{\partial x_\alpha}\right)^2\Psi + \left(q\phi' + \hbar\frac{\partial\delta}{\partial t}\right)\Psi.$$

- Matrix elements of observable quantities, which correspond to Hermitian operators, are gauge-invariant, i.e.

$$\rho'(\boldsymbol{r}, t) = \left| e^{i\delta} \Psi(\boldsymbol{r}, t) \right|^2 = \rho(\boldsymbol{r}, t)$$

  As an example, we consider the kinetic moment operator $\boldsymbol{\pi} = \boldsymbol{p} - q\boldsymbol{A}$:

$$\begin{aligned}
\left\langle \Psi_1' \middle| \boldsymbol{\pi}' \middle| \Psi_2' \right\rangle &= \left\langle \Psi_1' \middle| \boldsymbol{p} - q\boldsymbol{A}' \middle| \Psi_2' \right\rangle = \left\langle e^{i\frac{q}{\hbar}\Lambda} \Psi_1 \middle| \boldsymbol{p} - q(\boldsymbol{A} + \nabla\Lambda) \middle| e^{i\frac{q}{\hbar}\Lambda} \Psi_2 \right\rangle \\
&= \left\langle e^{i\frac{q}{\hbar}\Lambda} \Psi_1 \middle| e^{i\frac{q}{\hbar}\Lambda} (\boldsymbol{p} - q\boldsymbol{A}) \middle| \Psi_2 \right\rangle = \left\langle \Psi_1 \middle| (\boldsymbol{p} - q\boldsymbol{A}) \middle| \Psi_2 \right\rangle \\
&= \left\langle \Psi_1 \middle| \boldsymbol{\pi} \middle| \Psi_2 \right\rangle.
\end{aligned}$$

Derivation: Schrödinger Equation After a Local Gauge Transformation

- From above, the local, time-dependent gauge transformation reads

$$\Psi'(\boldsymbol{r}, t) = e^{i\delta(\boldsymbol{r}, t)} \Psi(\boldsymbol{r}, t).$$

  We substitute this transformation into the Schrödinger equation for a particle in an electromagnetic field, i.e.

$$i\hbar \frac{\partial}{\partial t} \Psi = \frac{(\boldsymbol{p} - q\boldsymbol{A})^2}{2m} \Psi + q\phi\Psi = \frac{(-i\hbar\nabla - q\boldsymbol{A})^2}{2m} \Psi + q\phi\Psi,$$

  which results in

$$i\hbar \frac{\partial}{\partial t} \Psi' = \frac{(-i\hbar\nabla - q\boldsymbol{A}')^2}{2m} \Psi' + q\phi'\Psi',$$

  which is the Schrödinger equation for the transformed wavefunction Ψ.

- To derive the transformed equation in terms of Ψ, we first make an auxiliary calculation that will help us simplify the expression $(-i\hbar\nabla - q\boldsymbol{A}')^2$. We begin by deriving the general identity

$$\begin{aligned}
\left(i\frac{\partial}{\partial x} + f \right) e^{i\delta} \Psi &= -\frac{\partial\delta}{\partial x} e^{i\delta} \Psi + i e^{i\delta} \frac{\partial\Psi}{\partial x} + f e^{i\delta} \Psi \\
&= e^{i\delta} \left(i\frac{\partial}{\partial x} + f - \frac{\partial\delta}{\partial x} \right) \Psi,
\end{aligned}$$

  which we repeat recursively to produce

$$\left(i\frac{\partial}{\partial x} + f \right)^2 e^{i\delta} \Psi = e^{i\delta} \left(i\frac{\partial}{\partial x} + f - \frac{\partial\delta}{\partial x} \right)^2 \Psi.$$

- Next, as quoted above, we define the transformed gauge potentials

$$\boldsymbol{A}' = \boldsymbol{A} + \nabla\Lambda \qquad \text{and} \qquad \phi' = \phi - \frac{\partial\Lambda}{\partial t}$$

  and the phase term $\delta(\boldsymbol{r}, t)$:

$$\delta(\boldsymbol{r}, t) = \frac{q}{\hbar} \Lambda(\boldsymbol{r}, t) \implies \Psi'(\boldsymbol{r}, t) = e^{i\frac{q}{\hbar}\Lambda(\boldsymbol{r}, t)} \Psi(\boldsymbol{r}, t)$$

  Finally, to make the next steps more clear, we use $\Psi'(\boldsymbol{r}, t) = e^{i\delta(\boldsymbol{r}, t)} \Psi(\boldsymbol{r}, t)$ to write the transformed Schrödinger equation in the form

$$i\hbar \frac{\partial}{\partial t} e^{i\delta(\boldsymbol{r}, t)} \Psi(\boldsymbol{r}, t) = \frac{(-i\hbar\nabla - q\boldsymbol{A}')^2}{2m} e^{i\delta(\boldsymbol{r}, t)} \Psi(\boldsymbol{r}, t) + i\phi' e^{i\delta(\boldsymbol{r}, t)} \Psi(\boldsymbol{r}, t).$$

- In terms of the above gauge transforms and the earlier mathematical identity with $f \to q\boldsymbol{A}'$, the Schrödinger equation for the original wavefunction Ψ reads

$$i\hbar\frac{\partial}{\partial t}\Psi = \frac{1}{2m}\sum_\alpha\left(-i\hbar\frac{\partial}{\partial x_\alpha} - qA'_\alpha + \hbar\frac{\partial\delta}{\partial x_\alpha}\right)^2\Psi + \left(q\phi' + \hbar\frac{\partial\delta}{\partial t}\right)\Psi,$$

  where we have substituted in the identities

$$-qA_\alpha = -qA'_\alpha + \hbar\frac{\partial\delta}{\partial x_\alpha} \qquad \text{and} \qquad q\phi = q\phi' + \hbar\frac{\partial\delta}{\partial t}$$

  These follow from our choice of gauge potential and the earlier mathematical identity, but we state them here without proof.

7.5   Aharonov-Bohm Effect

Qualitatively describe the Aharonov-Bohm effect, and then derive the formalism needed to explain it quantitatively. Be sure to discuss the role of gauge transformations.

- Qualitatively, the Aharonov-Bohm effect describes a phenomenon, without a classical analog, in which electron current is affected by a magnetic field that occurs in a region that the electrons don't actually travel through.

- The quantitative analysis of the Aharonov-Bohm effect rests on the equation

$$\Psi_{\mathrm{A}}(\boldsymbol{r}) = \exp\left(i\frac{q}{\hbar}\int_{\boldsymbol{r}_0}^{\boldsymbol{r}}\boldsymbol{A}(\tilde{\boldsymbol{r}})\cdot\mathrm{d}\tilde{\boldsymbol{r}}\right)\Psi_0(\boldsymbol{r}),$$

  where Ψ_{A} denotes the wavefunction for a particle of charge q, evaluated in the presence of a non-zero vector potential, and Ψ_0 denotes the wavefunction for the same particle, but evaluated in the absence of the vector potential, i.e. with $\boldsymbol{A} = 0$.

- The necessary analytical formalism, followed by the actual Aharonov-Bohm experiment, is presented in the following to subsubsections, respectively.

7.5.1   Aharonov-Bohm Effect: Necessary Formalism

- We begin by considering a region of space without a magnetic field, which implies

$$\boldsymbol{B} = \nabla \times \boldsymbol{A} = 0.$$

  Since $\nabla \times \boldsymbol{A} = 0$, we can write $\boldsymbol{A}$ as the gradient of a scalar field $\boldsymbol{A} = -\nabla\Lambda$ to get

$$\Lambda(\boldsymbol{r}) = \Lambda(\boldsymbol{r}_0) + \int_{\boldsymbol{r}_0}^{\boldsymbol{r}}\boldsymbol{A}(\tilde{\boldsymbol{r}})\cdot\mathrm{d}\tilde{\boldsymbol{r}}.$$

- We then choose the origin $\boldsymbol{r}_0$ to occur in the region without magnetic field, i.e. $\boldsymbol{B}(\boldsymbol{r}_0) = 0$. The constant $\Lambda(\boldsymbol{r}_0)$ then represents a global phase shift, which we can freely set to zero without affecting any observable properties of the wavefunction.

- The Schrödinger equation for a particle of charge q reads

$$i\hbar\frac{\partial}{\partial t}\Psi = \frac{(\boldsymbol{p} - q\boldsymbol{A})^2}{2m}\Psi + V\Psi.$$

However, we can also work with the transformed Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\Psi' = \frac{(\boldsymbol{p} - q\boldsymbol{A}')^2}{2m}\Psi' + V\Psi' = \frac{p^2}{2m}\Psi' + V\Psi',$$

where the vector potential $\boldsymbol{A}'$ (but not necessarily $\boldsymbol{A}$) vanishes in the absence of a magnetic field via

$$\boldsymbol{A}' = \boldsymbol{A} + \nabla\Lambda = \boldsymbol{A} - \boldsymbol{A} = 0,$$

where we have substituted in $\boldsymbol{A} = -\nabla\Lambda$.

- In the absence of a magnetic field, both wavefunctions Ψ and Ψ' describe the same particle, and the two wavefunctions are related by the local gauge transformation

$$\Psi'(\boldsymbol{r}, t) = e^{i\frac{q}{\hbar}\Lambda}\Psi(\boldsymbol{r}, t) \qquad \text{where} \qquad i\frac{q}{\hbar}\Lambda(\boldsymbol{r}) = i\frac{q}{\hbar}\int_{\boldsymbol{r}_0}^{\boldsymbol{r}}\boldsymbol{A}(\tilde{\boldsymbol{r}})\cdot\mathrm{d}\tilde{\boldsymbol{r}}.$$

The above expression motivates the definition of Ψ_{A} as

$$\Psi_{\mathrm{A}}(\boldsymbol{r}) = \exp\left(i\frac{q}{\hbar}\int_{\boldsymbol{r}_0}^{\boldsymbol{r}}\boldsymbol{A}(\tilde{\boldsymbol{r}})\cdot\mathrm{d}\tilde{\boldsymbol{r}}\right)\Psi_0(\boldsymbol{r}),$$

where Ψ_{A} denotes the wavefunction evaluated in the presence of a non-zero vector potential and Ψ_0 denotes the wavefunction evaluated in the absence of a vector potential.

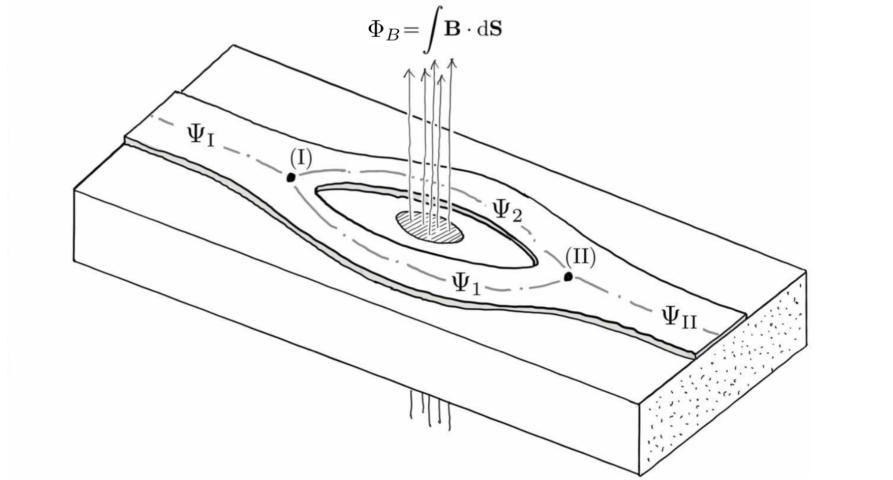Figure 1: Schematic for understanding the Aharonov-Bohm experiment; image credit is due to professor Ramšak.

7.5.2 The Aharonov-Bohm Experiment

- The experiment is based on the two-dimensional system shown in Figure 1, in which an electron can move along a straight quantum wire, which briefly splits into two branches at junction (I), which then join back into a single wire at junction (II).

- Assume, in the absence of both a magnetic field and vector potential (i.e. $\boldsymbol{B} = 0$ and $\boldsymbol{A} = 0$), that an electron wave packet Ψ travels along the wire, splits at the first junction (numbered with Roman numeral I) into two equal parts $\Psi_{\mathrm{I}} = \Psi_1 + \Psi_2$, where Ψ_1 and Ψ_2 encode the probability for taking either branch 1 and branch 2.

  Both Ψ_1 and Ψ_2 propagate along their respective branches and joint back at the second junction (numbered with Roman numeral II) into the wavefunction Ψ_{II}.

  The total electric current through the wire after the junction is proportional to $|\Psi_{\mathrm{ii}}|^2$, the probability of finding the particle at the second junction.

- Next, assume the region between the two branches is permeated with a magnetic field $\boldsymbol{B}$, and that $\boldsymbol{B} = 0$ everywhere else, including along the wires carrying the electron wave packet. Although $\boldsymbol{B} = 0$ along the wires, the vector potential, defined via $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ "spills out" from the region between the branches, and is non-zero along the wires.

- Next, using the earlier result

$$\Psi_{\mathrm{A}}(\boldsymbol{r}) = \exp\left(i\frac{q}{\hbar}\int_{\boldsymbol{r}_0}^{\boldsymbol{r}} \boldsymbol{A}(\tilde{\boldsymbol{r}}) \cdot \mathrm{d}\tilde{\boldsymbol{r}}\right)\Psi_0(\boldsymbol{r}),$$

  the wavefunction encoding the electron at the second junction is

$$\Psi_{\mathrm{II}} = e^{i\delta_1}\Psi_{0_1} + e^{i\delta_2}\Psi_{0_2},$$

  where Ψ_{0_1} and Ψ_{0_2} are the wavefunctions that would occur in branches 1 and 2 in the absence of $\boldsymbol{A}$, while the phase shifts in each branch are defined via

$$\delta_1 = \frac{q}{\hbar}\int_{\mathrm{branch}\ 1} \boldsymbol{A}(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{r} \qquad \text{and} \qquad \delta_2 = \frac{q}{\hbar}\int_{\mathrm{branch}\ 2} \boldsymbol{A}(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{r}.$$

- Assuming the probabilities of taking either branch are equal, we have $\Psi_{0_1} = \Psi_{0_2} \equiv \Psi_0$ and the expression for the wavefunction at the second junction simplifies to

$$\Psi_{\mathrm{II}} = e^{i\delta_2}\left(1 + e^{i(\delta_1 - \delta_2)}\right)\Psi_0.$$

  Since the phases δ_1 and δ_2 are defined by path integrals over branches 1 and 2, respectively, the total phase difference $\delta_1 - \delta_2$ corresponds to a closed line integral over the loop formed by the two branches, and reads

$$\delta_1 - \delta_2 = \frac{q}{\hbar}\oint_{\mathrm{loop}} \boldsymbol{A}(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{r}.$$

  We then rewrite the closed line integral with Stokes' theorem to get

$$\delta_1 - \delta_2 = \frac{q}{\hbar}\oiint_S \nabla \times \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{S} = \frac{q}{\hbar}\oiint_S \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S} = \frac{q}{\hbar}\Phi_{\mathrm{M}},$$

  where S is the surface bounded by the loop and Φ_{M} is the magnetic flux through the surface. In other words, the phase difference $\delta_1 - \delta_2$ is proportional to the magnetic flux through the region between the two branches.

  The wavefunction encoding the electron at the second junction is thus

$$\Psi_{\mathrm{II}} = e^{i\delta_2}\left(1 + e^{i\frac{q}{\hbar}\Phi_{\mathrm{M}}}\right)\Psi_0.$$

- Recall, as stated earlier, that electric current at the second junction is proportional to Ψ_{II}. The ratio of electric current I_{M} in the presence of a magnetic field in the intra-branch region, with non-zero Φ_{M}, to the electric current I_0 in the absence of a magnetic field, with $\Phi_{\mathrm{M}} = 0$, is thus

$$\frac{I_{\mathrm{M}}}{I_0} = \frac{|\Psi_{\mathrm{II}_{\mathrm{M}}}|^2}{|\Psi_{\mathrm{II}}|^2} = \frac{\left| e^{i\delta_2} \left(1 + e^{i\frac{q}{\hbar}\Phi_{\mathrm{M}}} \right) \Psi_0 \right|^2}{|e^{i\delta_2}(1+1)\Psi_0|^2} = \frac{1}{4} \left| 1 + e^{i\frac{q}{\hbar}\Phi_{\mathrm{M}}} \right|^2 = \cos^2 \left(\frac{q}{2\hbar}\Phi_{\mathrm{M}} \right).$$

In other words, the current through the second branch changes with the magnetic field in the region between the two conducting branches, even though the magnetic field is zero along the branches themselves.

Interpretation: The electron current I_{B} oscillates in an interference pattern that depends on a magnetic field that occurs in a region the electrons don't actually travel through!

8 Spin

8.1 Basics of Spin

Define spin, explain its relationship to angular momentum, and state some of the important spin properties.

- We will denote the spin operator by $\boldsymbol{S} = (S_x, S_y, S_z)$. The spin operator, like the angular momentum operator, obeys the fundamental commutation relations

$$[S_\alpha, S_\beta] = i\hbar\epsilon_{\alpha\beta\gamma}S_\gamma \qquad \text{or, in vector form,} \qquad \boldsymbol{S} \times \boldsymbol{S} = i\hbar\boldsymbol{S}.$$

In general, spin can be analyzed with the same formalism as angular momentum $\boldsymbol{L}$.

- Spin corresponds to angular momentum with half-integer eigenvalues $s = \frac{1}{2}, \frac{3}{2}, \ldots$, which do not lead to continuous solutions of the Schrödinger equation. As a result, the spin eigenvectors do not have a coordinate representation of the form $\langle \boldsymbol{r}|sm_s \rangle$.

  Instead, we represent the spin eigenstates with *spinors*, which are $(2s + 1)$-tuples in the complex vector space $\mathbb{C}^{2s+1}$.

- The spin states $|sm_s\rangle$ are eigenstates of both the S^2, the squared magnitude of spin, and of S_z, the projection of spin onto the z axis. The eigenvalue relations are

$$S_z |sm_s\rangle = m_s\hbar |sm_s\rangle \qquad \text{and} \qquad S^2 |sm_s\rangle = s(s+1)\hbar^2 |sm_s\rangle$$

Unlike, l, the spin quantum nubmer s takes on half-integer values, while the relationship between m_l and m_s is analogous. In equation form:

$$s \in \{\tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \ldots\} \qquad \text{and} \qquad m_s \in \{-s, -s+1, \ldots, s-1, s\}.$$

- The spin ladder operators S_+ and S_- are defined as

$$S_+ = S_x + iS_y \qquad \text{and} \qquad S_- = S_x - iS_y,$$

and obey the same relationships as the angular momentum operators L_+ and L_-. For example:

  - The ladder operators are not Hermitian, and instead obey $S_\pm = S_\mp^\dagger$.
  - When applied to the angular momentum basis states $|sm_s\rangle$, the spin ladder operators produce

$$S_{\pm_s} |sm_s\rangle = \hbar\sqrt{s(s+1) - m_s(m_s \pm 1)} |s, m_s \pm 1\rangle$$

  - We recover S_x and S_y from the ladder operators via

$$S_x = \frac{1}{2}(S_+ + S_-) \qquad \text{and} \qquad S_y = \frac{1}{2i}(S_+ - S_-)$$

8.2 Spin 1/2

Discuss the basis properties of a particle with spin $s = 1/2$, explain why such particles are physically important, and define the up and down arrow notation $|\uparrow\rangle$ and $|\downarrow\rangle$.

State and derive the matrix representations of the common spin operators in the $|sm\rangle$ basis for a particle with spin $s = 1/2$.

- We give states with spin $s = 1/2$ special attention—they are particularly important because many of the fundamental particles, such as the electron, proton and neutron, as well as many quarks, all have spin $s = 1/2$.

- We write these states with spin $s = 1/2$ in the form

$$|sm\rangle = \left|\tfrac{1}{2}m\right\rangle \equiv |\chi_m\rangle \, ,$$

  where we have written $m_s \equiv m$ for shorthand. We will write m_l and m_s explicitly when the context is unclear.

  When $s = 1/2$, the quantum number m can be either $-1/2$ or $1/2$, and we often abbreviate the two possible $|sm\rangle$ states with arrows, as in

$$\left|\tfrac{1}{2}\tfrac{1}{2}\right\rangle \equiv |\uparrow\rangle \qquad \text{and} \qquad \left|\tfrac{1}{2} - \tfrac{1}{2}\right\rangle \equiv |\downarrow\rangle \, ,$$

  to indicate "spin up" or "spin down". These states may also be written as the spinors

$$|\uparrow\rangle \equiv \chi_\uparrow = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad |\downarrow\rangle \equiv \chi_\downarrow = \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

  The states $|\uparrow\rangle$ and $|\downarrow\rangle$ form an orthonormal basis spanning the space of eigenstates for particles with spin $s = 1/2$.

- The spin ladder operators act on the states $|\uparrow\rangle$ and $|\downarrow\rangle$ according to

$$S_+ |\downarrow\rangle = \hbar |\uparrow\rangle \qquad \text{and} \qquad S_- |\uparrow\rangle = \hbar |\downarrow\rangle \, .$$

  In matrix form, in the $\{|\uparrow\rangle, |\downarrow\rangle\}$ basis, the ladder operators thus read

$$S_+ = \begin{pmatrix} \langle\uparrow|S_+|\uparrow\rangle & \langle\uparrow|S_+|\downarrow\rangle \\ \langle\downarrow|S_+|\uparrow\rangle & \langle\downarrow|S_+|\downarrow\rangle \end{pmatrix} = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$S_- = \begin{pmatrix} \langle\uparrow|S_-|\uparrow\rangle & \langle\uparrow|S_-|\downarrow\rangle \\ \langle\downarrow|S_-|\uparrow\rangle & \langle\downarrow|S_-|\downarrow\rangle \end{pmatrix} = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} .$$

  Note that the matrices preserve the relationship $S_\pm = S_\mp^\dagger$.

- In matrix form, the spin components S_x and S_y, which we can construct directly from the S_+ and S_- matrices, are

$$S_x = \frac{1}{2}(S_+ + S_-) = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$S_y = \frac{1}{2i}(S_+ - S_-) = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} .$$

We find the z component S_z with direct calculation:

$$S_z = \begin{pmatrix} \langle\uparrow|S_z|\uparrow\rangle & \langle\uparrow|S_z|\downarrow\rangle \\ \langle\downarrow|S_z|\uparrow\rangle & \langle\downarrow|S_z|\downarrow\rangle \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} \langle\uparrow|\uparrow\rangle & -\langle\uparrow|\downarrow\rangle \\ \langle\downarrow|\uparrow\rangle & -\langle\downarrow|\downarrow\rangle \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The corresponding eigenvectors (spinors) for S_x, S_y and S_z are

$$\chi_x = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \pm 1 \end{pmatrix} \qquad \chi_y = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \pm i \end{pmatrix} \qquad \chi_z = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

- For $\alpha \in \{x, y, z\}$, the squared components S_α^2 read

$$S_\alpha^2 = \frac{\hbar^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The squared components S_α^2 obey the commutation relations

$$\left[S_\alpha, S_\beta^2\right] = 0 \qquad \text{an} \qquad \left[S_\alpha^2, S_\beta^2\right] = 0.$$

Finally, the squared spin operator S^2 acts on the states $|\chi_m\rangle \equiv \left|\frac{1}{2}m\right\rangle$ according to

$$S^2|\chi_m\rangle = \sum_\alpha S_\alpha^2|\chi_m\rangle = \frac{3}{4}\hbar^2|\chi_m\rangle.$$

- As a final note, for larger spins $s = \frac{3}{2}, \frac{5}{2}, \ldots$, we write the spin wavefunction $|\psi\rangle$ by expanding $|\psi\rangle$ in the $\{|sm\rangle\}$ basis:

$$|\psi\rangle = \sum_{s=1/2}^{\infty}\sum_{m=-s}^{s} c_{sm}|sm\rangle, \qquad c_{sm} = \langle sm|\psi\rangle.$$

8.3   The Pauli Spin Matrices

Define the Pauli spin matrices and discuss some of their important properties. State how the common spin operators are written in terms of the Pauli spin matrices.

- The Pauli spin matrices, denoted by σ_x, σ_y and σ_z are defined as

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Together with the 2×2 identity matrix I, the Pauli matrices provide a convenient basis in which to expand an arbitrary 2×2 matrix.

- In terms of the Pauli spin matrices, the spin operator reads

$$\boldsymbol{S} = \frac{\hbar}{2}\boldsymbol{\sigma} \qquad \text{where} \qquad \boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z).$$

- The Pauli spin matrices are Hermitian, and obey the following properties:

$$\sigma_\alpha = \sigma_\alpha^\dagger \qquad \sigma_\alpha^2 = \text{I} \qquad \det \sigma_\alpha = -1 \qquad \text{tr}\,\sigma_\alpha = 0.$$

Each of the spin matrices has eigenvalues $\lambda_\pm = \pm 1$—note that the eigenvalues must be equal and opposite to satisfy $\text{tr}\,\sigma_\alpha = 0$.

- The product of two spin matrices obeys the general formula

$$\sigma_\alpha \sigma_\beta = \delta_{\alpha\beta} \, \mathrm{I} + i\epsilon_{\alpha\beta\gamma}\sigma_\gamma.$$

- The commutation relations between the spin matrices are analogous to the spin operator commutation relations, and read

$$[\sigma_\alpha, \sigma_\beta] = 2i\epsilon_{\alpha\beta\gamma}\sigma_\gamma.$$

More so, we can use matrix multiplication to derive the anti-commutator relation

$$\{\sigma_\alpha, \sigma_\beta\} = \sigma_\alpha\sigma_\beta + \sigma_\beta\sigma_\alpha = 2\delta_{\alpha\beta} \, \mathrm{I} \,.$$

- Finally, in terms of the spin matrices, an arbitrary vector $\boldsymbol{a} \in \mathbb{R}^3$ can be written

$$\boldsymbol{a} \cdot \boldsymbol{\sigma} = \sum_\alpha a_\alpha \sigma_\alpha = (\boldsymbol{a} \cdot \boldsymbol{\sigma})^\dagger.$$

Note that σ_α is a 2×2 matrix, while a_α is a scalar.

Meanwhile, for two vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^2$ we have

$$(\boldsymbol{a} \cdot \boldsymbol{\sigma})(\boldsymbol{b} \cdot \boldsymbol{\sigma}) = \sum_{\alpha,\beta} a_\alpha \sigma_\alpha b_\beta \sigma_\beta = \boldsymbol{a} \cdot \boldsymbol{b} \, \mathrm{I} + i \sum_{\alpha,\beta} \epsilon_{\alpha\beta\gamma} a_\alpha b_\beta \sigma_\gamma$$

$$= (\boldsymbol{a} \cdot \boldsymbol{b}) \, \mathrm{I} + i(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{\sigma}.$$

Again, we stress that σ_α and I are 2×2 matrices, while a_α and b_α are scalars.

- The expectation values of the Pauli spin matrices in a given spin state $|\chi\rangle$ are

$$\langle \sigma_x \rangle = \langle \chi | \sigma_x | \chi \rangle = (a^* \ b^*) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = (a^* \ b^*) \cdot \begin{pmatrix} b \\ a \end{pmatrix} = 2\,\mathrm{Re}\,[a^*b]$$

$$\langle \sigma_y \rangle = \langle \chi | \sigma_y | \chi \rangle = (a^* \ b^*) \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = (a^* \ b^*) \cdot \begin{pmatrix} -ib \\ ia \end{pmatrix} = 2\,\mathrm{Im}\,[a^*b]$$

$$\langle \sigma_z \rangle = \langle \chi | \sigma_z | \chi \rangle = (a^* \ b^*) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = (a^* \ b^*) \cdot \begin{pmatrix} a \\ -b \end{pmatrix} = |a|^2 - |b|^2.$$

8.4 Spinors and Spinor Transformations

Discuss spinors in the context of particles with spin $s = 1/2$, and discuss how spinors are transformed by rotation and time reversal. Describe the process of changing the axis of quantization from the conventional z axis to an arbitrary direction in space.

- For particles with spin $s = 1/2$, spinors, which describe general spin states, are determined by $2s + 1 = 2$ coordinates, e.g. $a, b \in \mathbb{C}$, that satisfy the normalization condition $|a|^2 + |b|^2 = 1$. In various notations, spin $s = 1/2$ spinors are written

$$\chi = \begin{pmatrix} a \\ b \end{pmatrix} = a\chi_\uparrow + b\chi_\downarrow \qquad |\chi\rangle = a\,|\!\uparrow\rangle + b\,|\!\downarrow\rangle \,.$$

The product of two spinor states reads

$$\langle \chi_1 | \chi_2 \rangle = \begin{pmatrix} a_1^* & b_1^* \end{pmatrix} \cdot \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = a_1^* a_2 + b_1^* b_2.$$

8.4.1   Rotation of Spinor States

- In terms of spin, the rotation operator reads by an angle ϕ about the axis $\hat{\mathbf{n}}$ reads

$$U(\phi\,\hat{\mathbf{n}}) = \exp\left(-i\phi\frac{\hat{\mathbf{n}}\cdot\boldsymbol{S}}{\hbar}\right) = \cos\left(\frac{\phi}{2}\right)\mathrm{I} - i\sin\left(\frac{\phi}{2}\right)\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}.$$

Note that rotating a spinor by an angle $\phi = 2\pi$ corresponds to multiplication by $-\mathrm{I}$, and not simply the identity I. We must rotate a spinor around "twice", i.e. by an angle of 4π, to recover its original orientation.

- In spherical coordinates, an arbitrary spinor can be parameterized by the angles θ and ϕ via

$$|\chi(\theta,\phi)\rangle = \cos\left(\frac{\theta}{2}\right)|\uparrow\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|\downarrow\rangle,$$

where θ encodes a rotation of the basis state $|\uparrow\rangle$ about the y axis, which is followed by a rotation of $|\uparrow\rangle$ by the angle ϕ about the z axis.

- If a spin operator $\hat{\mathbf{n}}(\theta,\phi)\cdot\boldsymbol{S}$ acts on an arbitrary spin state $|\chi(\theta,\phi)\rangle$ to produce the eigenvalue $\hbar/2$, as in the eigenvalue equation,

$$\left[\hat{\mathbf{n}}(\theta,\phi)\cdot\boldsymbol{S}\right]|\chi(\theta,\phi)\rangle = \frac{\hbar}{2}|\chi(\theta,\phi)\rangle,$$

then the spin operator $\boldsymbol{S}$ is parallel to the unit vector $\hat{\mathbf{n}}(\theta,\phi)$, where

$$\hat{\mathbf{n}}(\theta,\phi) = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta).$$

We can recover the direction $\hat{\mathbf{n}}$ from the spinor $\chi(\theta,\phi)$ via

$$\langle\chi(\theta,\phi)|\boldsymbol{\sigma}|\chi(\theta,\phi)\rangle = \hat{\mathbf{n}}(\theta,\phi).$$

Derivation: The Rotation Operator in Terms of Spin

- We rotate a spinor by the angle ϕ about the axis $\hat{\mathbf{n}}$ with the unitary rotation operator

$$U(\phi\,\hat{\mathbf{n}})\,|\chi\rangle = \exp\left(-i\phi\frac{\hat{\mathbf{n}}\cdot\boldsymbol{S}}{\hbar}\right)|\chi\rangle.$$

We begin by writing the rotation operator as a Taylor series in the spin matrix vector $\boldsymbol{\sigma}$, which gives

$$U(\phi\,\hat{\mathbf{n}}) = \exp\left(-i\phi\frac{\hat{\mathbf{n}}\cdot\boldsymbol{S}}{\hbar}\right) = \exp\left(-\frac{i\phi}{2}\,\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}\right) = \sum_{k=0}^{\infty}\frac{1}{k!}\left(-\frac{i\phi}{2}\,\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}\right)^k.$$

We then write the product $\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}$ in the form

$$(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})^k = \begin{cases} \mathrm{I} & k\text{ even} \\ \hat{\mathbf{n}}\cdot\boldsymbol{\sigma} & k\text{ odd,} \end{cases}$$

in terms of which the rotation operator reduces to the linear function

$$U(\phi\,\hat{\mathbf{n}}) = \sum_{k=0}^{\infty}\frac{1}{k!}\left(-\frac{i\phi}{2}\,\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}\right)^k = \left[1 - \frac{1}{2!}\left(\frac{\phi}{2}\right)^2 + \frac{1}{4!}\left(\frac{\phi}{2}\right)^4 \mp\cdots\right]\mathrm{I}$$
$$- i\left[\frac{\phi}{2} - \frac{1}{3!}\left(\frac{\phi}{2}\right)^3 + \frac{1}{5!}\left(\frac{\phi}{2}\right)^5 \mp\cdots\right]\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}.$$

Using the power series definitions of the sine and cosine function, this becomes

$$U(\phi\,\hat{\mathbf{n}}) = \cos\left(\frac{\phi}{2}\right) \mathrm{I} - i\sin\left(\frac{\phi}{2}\right) \hat{\mathbf{n}} \cdot \boldsymbol{\sigma},$$

as quoted at the begining of the subsection.

Parameterizing a Spinor

- We can parameterize an arbitrary spinor $|\chi\rangle$ with two angles. We begin by rotating a spin-up state $|\uparrow\rangle$ by an angle θ about the y axis, followed by a rotation by the angle ϕ about the z axis. The result is

$$|\chi\rangle = U(\phi\,\hat{\mathbf{e}}_z)U(\theta\,\hat{\mathbf{e}}_y)\,|\uparrow\rangle = e^{-i\frac{\phi}{2}\sigma_z}e^{-i\frac{\theta}{2}\sigma_y}\,|\uparrow\rangle$$
$$= e^{-i\frac{\phi}{2}}\left[\cos\left(\frac{\theta}{2}\right)|\uparrow\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|\downarrow\rangle\right].$$

Since a wavefunction is determined only up to a constant phase factor, we can neglect the coefficient $e^{-i\frac{\phi}{2}}$ go get

$$|\chi(\theta,\phi)\rangle = \cos\left(\frac{\theta}{2}\right)|\uparrow\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|\downarrow\rangle.$$

Example: Rotating Spinors

- As an example, the transformation encoding spinor rotation about only the y axis is

$$U(\theta\,\hat{\mathbf{e}}_y) = e^{-i\frac{\theta}{2}\sigma_y} = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}.$$

Note that if we rotate $|\uparrow\rangle \equiv |\chi_z\rangle$, which is an eigenstate of S_z, by an angle $\pi/2$ in the x direction, we end up with an eigenstate of the operator S_x:

$$U\left(\frac{\pi}{2}\,\hat{\mathbf{e}}_y\right)\begin{pmatrix}1\\0\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} = \chi_x,$$

since $S_x\,|\chi_x\rangle = \frac{\hbar}{2}\,|\chi_x\rangle$. We then repeat the procedure, rotating the S_z eigenstate $|\uparrow\rangle \equiv |\chi_z\rangle$ by an arbitrary angle θ about the y axis to get

$$U(\theta\,\hat{\mathbf{e}}_y)\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}\cos\frac{\theta}{2}\\\sin\frac{\theta}{2}\end{pmatrix} = \chi_z(\theta).$$

The resulting spinor $\chi_z(\theta)$ is an eigenstate of the projection of spin in the direction of the vector $\hat{\mathbf{n}}_0 = \cos\theta\,\hat{\mathbf{e}}_x + \sin\theta\,\hat{\mathbf{e}}_z$, which we confirm with the following calculation:

$$(\hat{\mathbf{n}}_0 \cdot \boldsymbol{S})\,|\chi_z(\theta)\rangle = (\sin\theta\sigma_x + \cos\theta\sigma_z)\,|\chi_z(\theta)\rangle = \frac{\hbar}{2}\begin{pmatrix}\cos\theta & \sin\theta \\ \sin\theta & -\cos\theta\end{pmatrix}|\chi_z(\theta)\rangle$$
$$= \frac{\hbar}{2}\,|\chi_z(\theta)\rangle.$$

- The above unit vector $\hat{\mathbf{n}}_0$ was constructed with $\phi = 0$. For a more general direction $\hat{\mathbf{n}} = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta)$, the eigenvalue equation reads

$$\left[\hat{\mathbf{n}}(\theta,\phi)\cdot\mathbf{S}\right]|\chi(\theta,\phi)\rangle = \frac{\hbar}{2}|\chi(\theta,\phi)\rangle.$$

Interpretation: If a spin operator $\hat{\mathbf{n}}(\theta,\phi)\cdot\mathbf{S}$ acts on an arbitrary spin state $|\chi(\theta,\phi)\rangle$ to produce the eigenvalue $\hbar/2$, as in the above eigenvalue equation, then the spin operator $\mathbf{S}$ is parallel to the unit vector $\hat{\mathbf{n}}(\theta,\phi)$, where

$$\hat{\mathbf{n}}(\theta,\phi) = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta).$$

We can recover the direction $\hat{\mathbf{n}}$ from the spinor $\chi(\theta,\phi)$ via

$$\langle\chi(\theta,\phi)|\boldsymbol{\sigma}|\chi(\theta,\phi)\rangle = \hat{\mathbf{n}}(\theta,\phi).$$

We prove this by components. Starting with σ_x, we have

$$\langle\chi(\theta,\phi)|\boldsymbol{\sigma}|\chi(\theta,\phi)\rangle = \left(\cos\frac{\theta}{2},\; e^{-i\phi}\sin\frac{\theta}{2}\right)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \cos\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2} \end{pmatrix}$$

$$= \left(\cos\frac{\theta}{2},\; e^{-i\phi}\sin\frac{\theta}{2}\right)\begin{pmatrix} e^{i\phi}\sin\frac{\theta}{2} \\ \cos\frac{\theta}{2} \end{pmatrix}$$

$$= \sin\theta\cos\phi = \hat{\mathbf{n}}\cdot\hat{\mathbf{e}}_x,$$

which is the projection of $\hat{\mathbf{n}}$ onto the x axis. The calculation for σ_y and σ_z is analogous.

8.4.2 Time Reversal of Spinors

- The time reversal operator for a particle with spin $s = 1/2$ is

$$\mathcal{T} = i\sigma_y KT \equiv \tau KT, \qquad \tau = i\sigma_y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

where K is the complex conjugation operator and T changes the sign of time.

- The operators $\boldsymbol{\sigma}$ and K obey the commutator and anti-commutator relations

$$\{\sigma_x, \tau\} = \{\sigma_z, \tau\} = 0 \qquad \text{and} \qquad [\sigma_y, \tau] = 0$$
$$[\sigma_x, K] = [\sigma_z, K] = 0 \qquad \text{and} \qquad \{\sigma_y, \tau\} = 0.$$

From these relations, it follows that $\mathcal{T}$ reverses spin, which we show with the components of $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$:

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} KT\sigma_\alpha = -\sigma_\alpha\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} KT \implies \mathcal{T}\mathbf{S} = -\mathbf{S}\mathcal{T}.$$

- Under time reversal, the matrix elements of $\boldsymbol{\sigma}$ obey

$$\langle\mathcal{T}\chi_1|\boldsymbol{\sigma}|\mathcal{T}\chi_2\rangle = \langle i\sigma_y KT\chi_1|\boldsymbol{\sigma}i\sigma_y K|T\chi_2\rangle = -\langle i\sigma_y KT\chi_1|i\sigma_y K\boldsymbol{\sigma}|T\chi_2\rangle$$
$$- \langle KT\chi_1|K\boldsymbol{\sigma}|T\chi_2\rangle = -\langle T\chi_1|\boldsymbol{\sigma}|T\chi_2\rangle^*$$
$$- \langle\chi_1(-t)|\boldsymbol{\sigma}|\chi_2(-t)\rangle^*.$$

The above result implies

$$\langle\mathcal{T}\chi|\boldsymbol{\sigma}|\mathcal{T}\chi\rangle = -\langle\chi(-t)|\boldsymbol{\sigma}|\chi(-t)\rangle.$$

- Time-dependent spinors $\chi(t)$ transform under time reversal according to

$$\mathcal{T}\chi(t) = \mathcal{T}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \begin{pmatrix} b^*(-t) \\ -a^*(-t) \end{pmatrix}.$$

- Finally, we write the time-reversed state as

$$\mathcal{T}\left|\chi(\theta,\phi)\right\rangle = \left|\chi(\theta+\pi,\phi)\right\rangle = \left|\tilde{\chi}(\theta,\phi)\right\rangle.$$

The time-transformed state $\mathcal{T}\left|\chi(\theta,\phi)\right\rangle$ is a second eigenstate of the spin eigenvalue equation, which for the original spin state $\left|\chi(\theta,\phi)\right\rangle$ read

$$\hat{\mathbf{n}}(\theta,\phi)\cdot\boldsymbol{S}\left|\chi(\theta,\phi)\right\rangle = \frac{\hbar}{2}\left|\chi(\theta,\phi)\right\rangle,$$

but with eigenvalue $-\hbar/2$. The relevant eigenvalue equation for $\mathcal{T}\left|\chi(\theta,\phi)\right\rangle$ reads

$$\hat{\mathbf{n}}(\theta,\phi)\cdot\boldsymbol{S}\left|\mathcal{T}\chi(\theta,\phi)\right\rangle = -\frac{\hbar}{2}\left|\mathcal{T}\chi(\theta,\phi)\right\rangle.$$

8.4.3   Changing the Axis of Quantization

- The basis states $\left|\uparrow\right\rangle$ and $\left|\downarrow\right\rangle$ correspond the z axis being the axis of quantization.

More generally, recall that the states $\left|\chi(\theta,\phi)\right\rangle$ and $\left|\mathcal{T}\chi(\theta,\phi)\right\rangle$ are eigenstates of the projection of spin $\boldsymbol{S}$ onto the direction

$$\hat{\mathbf{n}} = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta).$$

We can use the states $\left|\chi(\theta,\phi)\right\rangle$ and $\left|\chi(\theta,\phi)\right\rangle$ as basis vectors with respect to the new quantization axis $\hat{\mathbf{n}}$. In this cases the basis vectors are

$$\left|\chi_m(\hat{\mathbf{n}})\right\rangle = \begin{cases} \cos\theta\left|\uparrow\right\rangle + e^{i\phi}\sin\theta\left|\downarrow\right\rangle & m = \frac{1}{2} \\ \sin\theta\left|\uparrow\right\rangle - e^{i\phi}\cos\theta\left|\downarrow\right\rangle & m = -\frac{1}{2}. \end{cases}$$

- In terms of the basis vectors $\left|\chi_m(\hat{\mathbf{n}})\right\rangle$, we can write an arbitrary spin state $\left|\psi\right\rangle$ as

$$\left|\psi\right\rangle = \sum_m c_m \left|\chi_m(\hat{\mathbf{n}})\right\rangle \qquad \text{where} \qquad c_m = \left\langle\chi_m(\hat{\mathbf{n}})|\psi\right\rangle.$$

8.5   Spin-Orbit Coupling

What is the definition of spin magnetic moment? State the coupling term for spin-orbit coupling, and explain how the result is derived.

- The spin magnetic moment $\boldsymbol{\mu}_S$ of a charged particle with spin $s = 1/2$ is defined as

$$\boldsymbol{\mu}_S = g\frac{q}{2m}\boldsymbol{S}.$$

The factor g is called the gyromagnetic factor or simply g-factor, and takes on different values for different particles. For an electron, we have $g \approx 2$.

- The spin-orbit coupling term, which describes the coupling of a spin $s = 1/2$ particle with orbital angular momentum $\boldsymbol{L}$ orbiting a positively charged particle, is

$$H_{\text{LS}} = \frac{1}{2m^2c^2}\frac{1}{r}\frac{\partial V}{\partial r}\boldsymbol{L}\cdot\boldsymbol{S}.$$

Derivation: Spin-Orbit Coupling Term

- Next, we consider a semi-classical picture of an electron orbiting a hydrogen nucleus. The electron feels both an electric interaction due to the Coulomb force, and, in its own coordinate system, an "internal" magnetic field

$$\boldsymbol{B}_{\text{int}} = -\frac{1}{c^2}(\boldsymbol{v} \times \boldsymbol{\mathcal{E}}),$$

which arises because in the electron's coordinate system, the positively-charged proton orbits the electron at speed $\hat{\mathbf{v}}$, which leads to a magnetic force via the Lorentz interaction.

- In a spherically-symmetric potential $V = V(r)$ we have

$$-\nabla V \equiv \boldsymbol{F} = q\boldsymbol{\mathcal{E}} \implies \boldsymbol{\mathcal{E}} = -\frac{1}{q}\nabla V(r) = \frac{1}{q}\frac{\partial V}{\partial r}\frac{\boldsymbol{r}}{r}.$$

The corresponding magnetic field reads

$$\boldsymbol{B}_{\text{int}} = -\frac{1}{c^2}(\boldsymbol{v} \times \boldsymbol{\mathcal{E}}) = -\frac{1}{qc^2}\left(\boldsymbol{v} \times \frac{\partial V}{\partial r}\frac{\boldsymbol{r}}{r}\right) = \frac{1}{qmc^2}\frac{1}{r}\frac{\partial V}{\partial r}\boldsymbol{L}.$$

- In terms of the internal magnetic field $\boldsymbol{B}_{\text{int}}$ and the definition of spin magnetic moment, the coupling of a spin $s = 1/2$ with orbital angular momentum $\boldsymbol{L}$ to its internal magnetic field is

$$H_{\text{LS}} = \boldsymbol{\mu}_S \cdot \boldsymbol{B}_{\text{int}} = \frac{1}{m^2c^2}\frac{1}{r}\frac{\partial V}{\partial r}\boldsymbol{L} \cdot \boldsymbol{S}.$$

Note: The correct spin-orbit coupling term is actually

$$H_{\text{LS}} = \frac{1}{2m^2c^2}\frac{1}{r}\frac{\partial V}{\partial r}\boldsymbol{L} \cdot \boldsymbol{S},$$

and is smaller than our heuristically-derived result by a factor of two. This correct result is derived from the Dirac equation for relativistic particles, which is beyond the scope of this course.

8.6 Stern-Gerlach Experiment

Explain the outcome of the Stern-Gerlach experiment, including a formal derivation. Explain the implications of the experiment's results on quantization of magnetic moment and spin.

- Figure 2 and the associated caption summarize the Stern-Gerlach experiment. The discrete distribution of detected particles along the z axis, which differs from the classically expected result (i.e. a continuous distribution along the z axis) implies the atoms passing through the Stern-Gerlach magnet have quantized magnetic dipole moments. The measured quantum of magnetic dipole moment in the Stern-Gerlach experiment is the Bohr magneton

$$\mu_B = \frac{e_0\hbar}{2m_{\text{e}}},$$

in agreement with the quantization of magnetic moment discussed in Subsection 7.2

- Subsections 8.6.1 and 8.6.2 immediately below give a semi-classical and quantum-mechanical analysis, respectively, of the Stern-Gerlach experiment.
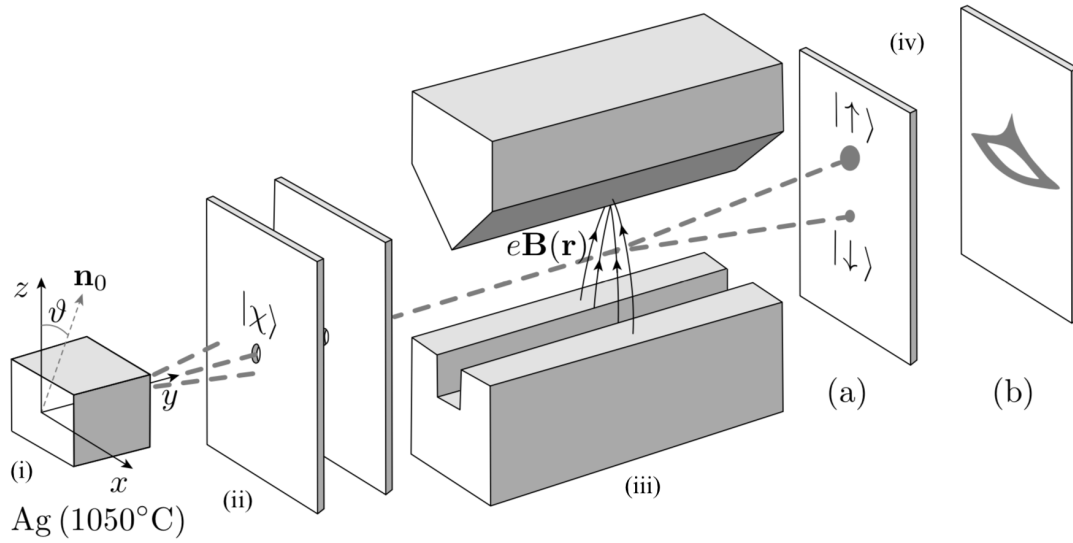
Figure 2: Schematic for understanding the Stern-Gerlach experiment; image credit is due to professor Ramšak. Silver atoms are ejected from a heated oven (i) in the y direction and pass through a collimator (ii) before entering a magnet with an nonhomogeneous magnetic field $\boldsymbol{B}(\boldsymbol{r})$, which exerts a magnetic torque on the atoms' magnetic dipole moments. The atoms correspondingly deflect in the z direction before hitting the detection plate (iv). Screen (a) shows the ideal result and screen (b) shows the actual result; in both cases the distribution along the z direction is discrete, demonstrating quantization of magnetic moment, which is an effect of the atoms' intrinsic spin.

8.6.1 Semi-Classical Analysis

Homogeneous Magnetic Field

- We first consider a simplified analysis with a homogeneous magnetic field. The actual experiment invovles a nonhomogeneous field, but this simplified analysis will serve as a good introduction to the experiment. Consider a silver atom of mass m exiting the Stern-Gerlach oven. The silver atom's shells are fully closed up to $4d^{10}$, with a single electron in the 5s shell. The electron in the 5s shell has magnetic dipole moment $\boldsymbol{\mu}$.

- The atom flies through a region with a homogeneous magnetic field $\boldsymbol{B}_0 = (0, 0, B_0)$, and the electron's magnetic moment $\boldsymbol{\mu}$ interacts with the magnetic field to produce a magnetic torque
$$\boldsymbol{M} = \boldsymbol{\mu} \times \boldsymbol{B}_0.$$

- We assume a simplified model in which the atom moves with uniform speed v_y in the y direction through the Stern-Gerlach apparatus in a precisely defined spin state
$$\chi(0) = \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix},$$
which is determined by the electron in the 5d orbital.

- The angle θ with respect to the z axis determines the spin's orientation in the xz plane. The associated spin orientation, in terms of the spin matrix operator $\boldsymbol{\sigma}$, is

$$\hat{\mathbf{n}}_0 = \langle \chi | \boldsymbol{\sigma} | \chi \rangle = (\sin \theta, 0, \cos \theta).$$

- In the presence of a homogeneous magnetic field $\boldsymbol{B}_0$, the atom would travel in an uninterrupted trajectory straight through the Stern-Gerlach apparatus. However, because of the torque $\boldsymbol{M} = \mu \times \boldsymbol{B}_0$ on the atom's magnetic dipole moment, the orientation of the atom's spin would rotate about the z axis by an angle $\delta\phi$ given by

$$\delta\phi = \omega_0 \tau = \omega_0 \frac{L}{v_y},$$

where L is then length of the Stern-Gerlach apparatus, τ is the (short) flight time through the apparatus, and $\omega_0 = \frac{\mu B_0}{\hbar}$ is the Larmor frequency.

After passing through the Stern-Gerlach apparatus, the silver atom continues along the y axis in the rotated spin state

$$\chi(\tau) = \begin{pmatrix} e^{\frac{i}{2}\delta\phi} \cos \frac{\theta}{2} \\ e^{-\frac{i}{2}\delta\phi} \sin \frac{\theta}{2} \end{pmatrix}.$$

The new spin orientation is given by

$$\hat{\mathbf{n}}(\tau) = \langle \chi(\tau) | \boldsymbol{\sigma} | \chi(\tau) \rangle = (\sin \theta \cos \delta\phi, -\sin \theta \sin \delta\phi, \cos \theta).$$

Nonhomogeneous Magnetic Field

- The true Stern-Gerlach apparatus contains a magnetic with a nonhomogeneous magnetic field of the form

$$\boldsymbol{B} = (B_x, 0, B_z).$$

In the presence of this magnetic field, a neutral particle with magnetic dipole moment $\boldsymbol{\mu}$ experiences a force

$$\boldsymbol{F} = (\boldsymbol{\mu} \cdot \nabla)\boldsymbol{B}(\boldsymbol{r}) = \nabla(\boldsymbol{\mu} \cdot \boldsymbol{B}(\boldsymbol{r})).$$

The expression for force rests on the assumption that $\boldsymbol{\mu}$ is constant and the condition $\nabla \times \boldsymbol{B} = 0$, and is derived via

$$\nabla(\boldsymbol{\mu} \cdot \boldsymbol{B}) = (\boldsymbol{\mu} \cdot \nabla)\boldsymbol{B} + (\boldsymbol{B} \cdot \nabla)\boldsymbol{\mu} + \boldsymbol{\mu} \times (\nabla \times \boldsymbol{B}) + \boldsymbol{B} \times (\nabla \times \boldsymbol{\mu}) = (\boldsymbol{\mu} \cdot \nabla)\boldsymbol{B}.$$

- Next, we expand the magnetic field to linear order in $\boldsymbol{B}$, which reads

$$\boldsymbol{B}(\boldsymbol{r}) = \boldsymbol{B}_0 + \sum_a \frac{\partial \boldsymbol{B}}{\partial x_\alpha} x_\alpha + \cdots,$$

Along the direction of travel along the y axis, where $x \sim 0$ and $z \sim 0$, we can neglect the force in the x direction, and the resulting force on the atom is then

$$\boldsymbol{F} = F_z \,\hat{\mathbf{e}}_z = \mu \frac{\partial B_z}{\partial z}\bigg|_{z=0} \hat{\mathbf{e}}_z.$$

- Assuming the atom's magntic moment $\boldsymbol{\mu}$ is parallel to the magnetic field in the z direction, then after passing through the Stern-Gerlach apparatus in time τ, the atom experiences an impulse $\pm F_z \tau$, and its speed changes according to

$$mv_z = \pm F_z \tau = \pm F_z \frac{L}{v_y},$$

where the sign depends on whether the magnetic moment is parallel or antiparallel to the magnetic field.

The particle is thus incident on the detection screen either above or below the plane $z = 0$, corresponding to either "up" or "down" states. By measuring the distribution of atoms on the detection screen, we can determine the probabilities $P_\uparrow$ and $P_\downarrow$ for an up or down outcome. In terms of these probabilities, the expected value of spin in the z direction, i.e. $\langle \sigma_z \rangle = \cos\theta$ is found with

$$\langle \sigma_z \rangle = \cos\theta = P_\uparrow - P_\downarrow.$$

Derivation: Finally, we derive the relationship $\langle \sigma_z \rangle = \cos\theta$. We begin by writing the atom's spin state $|\chi\rangle$ in the form

$$\chi = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{pmatrix}.$$

The probabilities for detecting the particle with spin up or spin down are

$$P_\uparrow = |\langle \uparrow | \chi \rangle|^2 = \cos^2\frac{\theta}{2} \qquad \text{and} \qquad P_\downarrow = |\langle \downarrow | \chi \rangle|^2 = \sin^2\frac{\theta}{2} = 1 - P_\uparrow,$$

while the expectation value of σ_z, using the definition $\langle \sigma_z \rangle = \langle \chi | \sigma_z | \chi \rangle$, is

$$\langle \sigma_z \rangle = \langle \chi | \sigma_z | \chi \rangle = \left(\cos\frac{\theta}{2}, \sin\frac{\theta}{2} \right) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{pmatrix}$$
$$= \cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2} \equiv P_\uparrow - P_\downarrow = 2\cos^2\frac{\theta}{2} - 1$$
$$= \cos\theta.$$

8.6.2 Quantum Stern-Gerlach Experiment

- The goal in this analysis is to find the time evolution $|\Psi(\boldsymbol{r}, t)\rangle$ of a wavefunction encoding a silver atom passing through the Stern-Gerlach apparatus, and then use the wavefunction to determine the spin expectation values $\langle \sigma_x \rangle$, $\langle \sigma_y \rangle$ and $\langle \sigma_z \rangle$.

- We begin by assigning the silver atom of mass m passing through the Stern-Gerlach apparatus a Hamiltonian of the form

$$H = -\frac{\hbar^2}{2m}\nabla^2 - \frac{q}{m_e}\boldsymbol{S} \cdot \boldsymbol{B} \equiv H_0 - \frac{q}{m_e}\boldsymbol{S} \cdot \boldsymbol{B},$$

where H_0 is the kinetic term and $-\frac{q}{m_e}\boldsymbol{S} \cdot \boldsymbol{B}$ represents the anomolous Zeeman coupling of the 5s electron to the external magnetic field.

- Both energy and force are constants of motion, and we define them as

$$2E_0 \equiv \hbar\omega_0 = \frac{q\hbar}{m_\mathrm{e}} B_0 \qquad \text{and} \qquad F = \frac{q\hbar}{2m_\mathrm{e}} \frac{\partial B_z}{\partial z}\bigg|_{z=0},$$

where we have assumed the magnetic field has only a z component. In terms of the energy and force terms, the silver atom's Hamiltonian reads

$$H = H_0 - \frac{1}{2}\hbar\omega_0\sigma_z - Fz\sigma_z = H_0 - E_0\sigma_z - Fz\sigma_z,$$

where the force term is constructed so as to satisfy $-\nabla Fz\sigma_z = F\sigma_z$.

- Just before the atom enters the Stern-Gerlach apparatus (before experiencing the magnetic field), which we define to occur at $t = 0$, we describe the atom with the wave packet

$$\Psi(\boldsymbol{r}, 0) = \psi_0(x, y, 0)\phi_0(z) \begin{pmatrix} c_1 \\ c_2 \end{pmatrix},$$

where the initial wavefunction $\phi(z, t = 0) \equiv \phi_0(z)$ is given in the momentum eigenbasis by

$$\phi_0(z) = \int_{-\infty}^{\infty} \widehat{\phi}(p) e^{i\frac{p}{\hbar}z} \, \mathrm{d}p.$$

The wavefunction $\phi_0(z)$ encodes motion along the z axis. The coefficients $c_1 = \cos\frac{\theta}{2}$ and $c_2 = \sin\frac{\theta}{2}$ are the spinor coefficients encoding the atom's initial spin state.

$$|\chi(0)\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{pmatrix} \equiv \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

The wavefunction ψ_0 corresponds to motion in the xy plane with speed

$$v_y = \frac{1}{m}\langle p_y\rangle.$$

We assume the speed v_y is large, so the initial wave packet doesn't widen significantly in the magnet. At later times, the wavefunction is determined by

$$\Psi(\boldsymbol{r}, t) = \psi_0(z, y, t) \begin{pmatrix} \phi_\uparrow(z, t) \\ \phi_\downarrow(z, t) \end{pmatrix}.$$

- The time evolution of ψ_0, which encodes motion in the xy plane, is simply uniform motion along the y axis and is unrelated to dynamics of spin in the experiment, which is confined to the z axis and represented by the wavefunction ϕ.

- The time evolution of the spin wavefunction is given by the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix} \phi_\uparrow(z, t) \\ \phi_\downarrow(z, t) \end{pmatrix} = -\left(\frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2} + \frac{1}{2}\hbar\omega_0\sigma_z + Fz\sigma_z\right)\begin{pmatrix} \sigma_\downarrow(z, t) \\ \sigma_\uparrow(z, t) \end{pmatrix}.$$

The initial wavefunction ϕ_0 has a unitary time evolution given by

$$e^{-i\frac{H}{\hbar}\tau}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} e^{i\frac{p}{\hbar}z} = \begin{pmatrix} e^{\frac{i}{2}\delta\varphi + i\frac{\delta p}{\hbar}z}c_1 \\ e^{-\frac{i}{2}\delta\varphi - i\frac{\delta p}{\hbar}z}c_2 \end{pmatrix} e^{i\frac{p}{\hbar}z - i\frac{E_p}{\hbar}\tau},$$

where $e^{i\frac{H}{\hbar}t}$ is the unitary time evolution operator, and we have defined

$$E_p = \frac{p^2}{2m} \qquad \tau \sim \frac{L}{v_y} \qquad \delta\varphi = \omega_0\tau \qquad \delta p = mv = F\tau.$$

- We denote the state immediately after exiting the magnet as

$$\Psi(\boldsymbol{r}, 0^+) \equiv \Psi(\boldsymbol{r}, \tau) = e^{i\frac{H}{\hbar}\tau}\Psi(\boldsymbol{r}, 0).$$

- The solution to the spin Schrödinger equation is

$$\begin{pmatrix} \phi_\uparrow(z, 0^+) \\ \phi_\downarrow(z, 0^+) \end{pmatrix} = \int_{-\infty}^{\infty} \begin{pmatrix} e^{\frac{i}{2}\delta\varphi + i\frac{p+\delta p}{\hbar}z}c_1 \\ e^{-\frac{i}{2}\delta\varphi - i\frac{p-\delta p}{\hbar}z}c_2 \end{pmatrix} \widehat{\phi}(p)\,\mathrm{d}p = \int_{-\infty}^{\infty} \begin{pmatrix} e^{\frac{i}{2}\delta\varphi}\widehat{\phi}(p-\delta p)c_1 \\ e^{-\frac{i}{2}\delta\varphi}\widehat{\phi}(p+\delta p)c_1 \end{pmatrix} e^{i\frac{p}{\hbar}z}\,\mathrm{d}p,$$

  where the spin wave packet has experienced an impulse $\pm\delta p$ and transformed as

$$\phi_0(z) \to \phi(z, 0^+) = e^{\pm i\frac{\delta p}{\hbar}z}\phi_0(z).$$

- Assuming the speed v_y through the Stern-Gerlach apparatus is large and the wavepacket does not appreciably spread out from its initial state, we can write the spin wavefunction at later times as

$$\phi(z, t) = e^{i\frac{(\delta p)^2}{2m\hbar}t}\phi(z - vt, 0^+) = \phi_0(z - vt)e^{i\frac{\delta p}{\hbar}\left(z - \frac{1}{2}vt\right)},$$

  and the solution for the complete wave packet's time evolution is

$$\Psi(\boldsymbol{r}, t) = \psi_0(x, y, t) \begin{pmatrix} e^{\frac{i}{2}\delta\varphi}\phi_0(z - vt)e^{i\frac{\delta p}{\hbar}\left(z - \frac{1}{2}vt\right)}c_1 \\ e^{-\frac{i}{2}\delta\varphi}\phi_0(z + vt)e^{i\frac{\delta p}{\hbar}\left(z + \frac{1}{2}vt\right)}c_2 \end{pmatrix}.$$

Expected Value of Spin

- We assume an initial spin state

$$\chi(0) = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{pmatrix},$$

  and write the time-dependent spin state in the form

$$|\chi(z, t)\rangle = \phi_\uparrow(z, t)\,|\uparrow\rangle + \phi_\downarrow(z, t)\,|\downarrow\rangle,$$

  where the coefficients $\phi_\uparrow$ and $\phi_\downarrow$ are time-dependent.

- The expected value of spin in the z direction is

$$\langle \sigma_z \rangle = \int_{-\infty}^{\infty} \left(|\phi_{ua}(z, t)|^2 - |\phi_{ua}(z, t)|^2 \right) \mathrm{d}z$$

$$= \int_{-\infty}^{\infty} \left(\cos^2\frac{\theta}{2}|\phi_0(z - vt)|^2 - \sin^2\frac{\theta}{2}|\phi_0(z + vt)|^2 \right) \mathrm{d}z$$

$$= P_\uparrow - P_\downarrow = \cos\theta,$$

  in agreement withthe semi-classical result.

- Next, we calculate spin expectation value in the x direction, which reads

$$\langle \sigma_x \rangle = 2 \operatorname{Re} \int_{-\infty}^{\infty} \phi_\uparrow^*(z,t)\phi_\downarrow(z,t)\, \mathrm{d}z$$

We begin by making the auxiliary calculation

$$I \equiv \int_{-\infty}^{\infty} \phi_0 \left(z - \frac{\delta p}{m} t \right) \phi_0 \left(z + \frac{\delta p}{m} t \right) e^{2i \frac{\delta p}{\hbar} z}\, \mathrm{d}z$$

$$= \exp \left[-\frac{(\delta p)^2}{2\sigma_p^2} \left(1 + \frac{4\sigma_p^4 t^2}{m^2 \hbar^2} \right) \right] = \exp \left[- \left(\frac{(\delta p)^2}{2\sigma_p^2} + \frac{2\sigma_p^2 (\delta p)^2}{m^2 \hbar^2} t^2 \right) \right]$$

$$\equiv \exp \left[- \left(\frac{(\delta p)^2}{2\sigma_p^2} + \lambda t^2 \right) \right],$$

where we have defined the constant

$$\lambda \equiv \frac{2\sigma_p^2 (\delta p)^2}{m^2 \hbar^2}.$$

For large times, the exponent and integral approach

$$\frac{(\delta p)^2}{2\sigma_p^2} + \lambda t^2 \to \lambda t^2 \implies I \to e^{-\lambda t^2}.$$

- Next, we combine the result of the integral $I \to e^{-\lambda t^2}$ with the expression for $\langle \sigma_x \rangle$. Without proof, the result is

$$\langle \sigma_x \rangle = e^{-\lambda t^2} \sin \theta \cos \delta \varphi.$$

Again without proof, a similar derivation for $\langle \sigma_y \rangle$ results in

$$\langle \sigma_y \rangle = -e^{-\lambda t^2} \sin \theta \sin \delta \varphi.$$

Note that both $\langle \sigma_x \rangle$ and $\langle \sigma_y \rangle$ approach zero for large times, in agreement with the idealized semi-classical result $\langle \sigma_x \rangle = \langle \sigma_y \rangle = 0$.

9 Addition of Angular Momentum

9.1 Formalism: Addition of Angular Momentum for Spin $1/2$ Particles

Explain the formalism for addition of angular momentum in quantum mechanics for a system of two particles with spin $s = 1/2$.

- Our concrete system: a hydrogen atom consisting of an electron in the ground state with angular momentum quantum number $l = 0$ and the proton in the nucleus. Both particles have spin $s_i = 1/2$.

- The spin operators and basis vectors for each particle are

$$S_i^2 = s_i(s_i + 1)\hbar^2 \left|s_i m_i\right\rangle \qquad \text{and} \qquad S_{z_i} \left|s_i m_i\right\rangle = m_i \hbar \left|s_i\right\rangle \left|m_i\right\rangle,$$

  where $i = 1$ corresponds to the electron and $i = 2$ to the proton.

- The electron's and proton's angular momentum operators commute with each other, which is summarized in the commutation relation

$$\left[S_{\alpha_i}, S_{\beta_j}\right] = i\hbar \delta_{ij} \epsilon_{\alpha\beta\gamma} S_{\gamma_j}, \qquad \alpha \in \{x, y, z\}.$$

  Both sets of spin operators can be written in terms of the Pauli spin matrices as

$$\boldsymbol{S}_i = \frac{\hbar}{2}\boldsymbol{\sigma}, \qquad \boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z).$$

- For a system of N particles with individual spin operators $\boldsymbol{S}_i$, the total spin $\boldsymbol{S}$ and total spin expectation value are defined as

$$\boldsymbol{S} = \sum_i^N \boldsymbol{S}_i \qquad \text{and} \qquad \langle\Psi|\boldsymbol{S}|\Psi\rangle = \sum_i \langle\chi_i|\boldsymbol{S}_i|\chi_i\rangle,$$

  where Ψ is the system's complete wavefunction and $|\chi_i\rangle$ are the spin states of the constituent particles.

- For the above definitions to hold, the operators $\boldsymbol{S}_i$ of each constituent particle must act only the spin basis states $|s_i m_i\rangle$ corresponding to that particular particle. With this requirement in mind, we define the total spin of a two-particle system as

$$\boldsymbol{S} = \boldsymbol{S}_1 \otimes \mathrm{I}_2 + \mathrm{I}_1 \otimes \boldsymbol{S}_2,$$

  where I_i is the identity operator for the spin subspace spanned by the ith particle's spin basis states $|\{s_i m_i\}\rangle$. Similarly, the total spin state $|s_1 s_2 m_1 m_2\rangle$ is written

$$|s_1 s_2 m_1 m_2\rangle = |s_1 m_1\rangle \otimes |s_2 m_2\rangle.$$

  Shorthand Notation: Since the spins s_1 and s_2 are both fixed at the constant values $s_{1,2} = 1/2$, we can leave the s_i terms implicit and write the spin state simply as

$$|s_1 s_2 m_1 m_2\rangle \equiv |m_1\rangle \otimes |m_2\rangle$$

- In the $|m_1 m_2\rangle$ notation, the total spin operator S_z acts on the basis state as

$$
\begin{aligned}
S_z \,|m_1 m_2\rangle &= (S_{z_1} \otimes \mathrm{I}_2 + \mathrm{I}_1 \otimes S_{z_2}) \,|m_1\rangle \otimes |m_2\rangle \\
&= m_1 \hbar \,|m_1\rangle \otimes |m_2\rangle + |m_1\rangle \otimes m_2 \hbar \,|m_2\rangle \\
&= (m_1 + m_2)\hbar \,|m_1 m_2\rangle .
\end{aligned}
$$

In other words, the total spin operator S_z has the eigenvalue $m_1 + m_2$ when acting on the state $|m_1 m_2\rangle$, as expected for the total spin operator.

- The total spin operator $\boldsymbol{S}$ obeys the usual angular momentum commutation relations

$$
[S_\alpha, S_\beta] = i\hbar \epsilon_{\alpha\beta\gamma} S_\gamma.
$$

Derivation: Using the fact that $[S_{\alpha_1} \otimes \mathrm{I}_2, \mathrm{I}_1 \otimes S_{\beta_2}] = 0$, which follows from the identity $[S_{\alpha_i}, S_{\beta_j}] = i\hbar \delta_{ij}\epsilon_{\alpha\beta\gamma} S_{\gamma_j}$ (or, more intuitively, from S_i and S_j acting on independent Hilbert spaces), we have

$$
\begin{aligned}
[S_\alpha, S_\beta] &= \left[S_{\alpha_1} \otimes \mathrm{I}_2 + I_1 \otimes S_{\alpha_2}, S_{\beta_1} \otimes \mathrm{I}_2 + I_1 \otimes S_{\beta_2} \right] \\
&= [S_{\alpha_1} \otimes \mathrm{I}_2, S_{\beta_1} \otimes \mathrm{I}_2] + [\mathrm{I}_2 \otimes S_{\alpha_2}, \mathrm{I}_1 \otimes S_{\beta_2}] \\
&= [S_{\alpha_1}, S_{\beta_1}] \otimes \mathrm{I}_2 + \mathrm{I}_1 \otimes [S_{\alpha_2}, S_{\beta_2}] \\
&= i\hbar \epsilon_{\alpha\beta\gamma} S_\gamma,
\end{aligned}
$$

where the last equality uses the identities $[S_{\alpha_i}, S_{\beta_i}] = i\hbar \epsilon_{\alpha\beta\gamma} S_{\gamma_i}$.

- The ladder operators for total spin operator $\boldsymbol{S}$ are defined as

$$
S_\pm = S_x \pm iS_y = S_{\pm_1} \otimes \mathrm{I}_2 + \mathrm{I}_1 \otimes S_{\pm_2}.
$$

Eigenstates

- The product states $\{|m_1 m_2\rangle\}$ form a basis spanning a two-particle system's total spin vector space.

The states $|m_1 m_2\rangle$ are eigenstates of the total angular momentum operator S_z, the individual operators S_{z_i} and individual operators S_i^2.

However, the states $|m_1 m_2\rangle$ are *not* eigenstates of the total angular momentum operator S^2—only the states $|sm\rangle$ are eigenstates of S^2.

The $|sm\rangle$ states obey the eigenvalue relations

$$
S^2 \,|sm\rangle = s(s+1)\hbar^2 \,|sm\rangle \qquad \text{and} \qquad S_z \,|sm\rangle = m\hbar \,|sm\rangle .
$$

The eigenstates $|sm\rangle$ of the total angular spin operator S^2 are called "good" spin states, or total angular momentum states, while $|m_1 m_2\rangle$ are called product states.

- Although the product states $|m_1 m_2\rangle$ are not directly eigenstates of S^2, we can expand the eigenstates $|sm\rangle$ in the $|m_1 m_2\rangle$ basis in the form

$$
|sm\rangle = \sum_{m_1 m_2} c_{m_1 m_2} \,|m_1 m_2\rangle ,
$$

where we sum over all $m_i = \pm 1/2$ obeying the condition $m = m_1 + m_2$.

9.2 The Singlet and Triplet States

State and derive the singlet and triplet states for a system of two particles with spin $s = 1/2$.

- The singlet and triplet states for a system of two particles with spin $s = 1/2$ are

$$|1m\rangle = \begin{cases} |\uparrow\uparrow\rangle & m = 1 \\ \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\right) & m = 0 \\ |\downarrow\downarrow\rangle & m = -1 \end{cases} \qquad \text{(triplet states)}$$

$$|00\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right) \qquad \text{(singlet state)}.$$

Derivation: The Triplet States

- First, we introduce the earlier arrow notation, which we generalize to a two-particle system with states $|m_1 m_2\rangle$ as follows:

$$\left|\tfrac{1}{2}\tfrac{1}{2}\right\rangle = |\uparrow\rangle \otimes |\uparrow\rangle \equiv |\uparrow\uparrow\rangle$$
$$\left|-\tfrac{1}{2}\tfrac{1}{2}\right\rangle = |\downarrow\rangle \otimes |\uparrow\rangle \equiv |\downarrow\uparrow\rangle$$
$$\left|\tfrac{1}{2}-\tfrac{1}{2}\right\rangle = |\uparrow\rangle \otimes |\downarrow\rangle \equiv |\uparrow\downarrow\rangle$$
$$\left|-\tfrac{1}{2}-\tfrac{1}{2}\right\rangle = |\downarrow\rangle \otimes |\downarrow\rangle \equiv |\downarrow\downarrow\rangle,$$

  where we have dropped the tensor product for conciseness.

- The largest possible value of m for the state $|\uparrow\uparrow\rangle$ is $m = \frac{1}{2} + \frac{1}{2} = 1$. The corresponding $|sm\rangle$ state for $m = 1$, i.e. $|sm\rangle = |11\rangle$ is also an eigenstate of S^2, as shown below:

$$S^2 |sm\rangle = S^2 |11\rangle = 1(1+1)\hbar^2 |11\rangle = 2\hbar^2 |\uparrow\uparrow\rangle.$$

  For the next step, we recall the action of the operator S_- on the state $|sm\rangle$, which reads

$$S_- |sm\rangle = \hbar\sqrt{s(s+1) - m(m-1)}\,|s, m+1\rangle$$

- Next, we apply the ladder operator S_- to the state $|11\rangle$, which gives

$$S_- |11\rangle = \hbar\sqrt{1(1+1) - 1(1-1)}\,|10\rangle = \sqrt{2}\hbar\,|10\rangle,$$

  which we can also write in the equivalent form

$$S_- |11\rangle = \left(S_{-_1} I_2 + I_1 S_{-_2}\right)|\uparrow\uparrow\rangle = \hbar\,|\downarrow\uparrow\rangle + \hbar\,|\uparrow\downarrow\rangle$$
$$= \hbar\left(|\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle\right).$$

  By equating the two equivalent forms of the same state, we have derived the identity

$$\sqrt{2}\hbar\,|10\rangle = \hbar\left(|\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle\right) \implies |10\rangle = \frac{\hbar}{\sqrt{2}}\left(|\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle\right).$$

- Following a similar procedure with $s = 1$ and $m = -1$, we could derive the identity

$$|1, -1\rangle = |\downarrow\downarrow\rangle.$$

  The above three states $|sm\rangle = |1m\rangle$ with $m = -1, 0, 1$ are called the *triplet* states for a two-particle spin system with individual spins $s_{1,2} = 1/2$.

Derivation: The Single State

- The fourth possible state for the two-particle spin $1/2$ system is the *singlet* state $|sm\rangle = |00\rangle$ with $s = 0$ and $m = 0$. To generate the singlet state, we expand the state $|sm\rangle = |00\rangle$ in the $|m_1 m_2\rangle$ basis, which reads

$$|00\rangle = c_{\frac{1}{2},-\frac{1}{2}} |\uparrow\downarrow\rangle + c_{-\frac{1}{2},\frac{1}{2}} |\downarrow\uparrow\rangle .$$

Because the states $|10\rangle$ and $|00\rangle$ have different values of s, they are orthogonal and obey $\langle 10|00\rangle = 0$, which implies

$$c_{-\frac{1}{2},\frac{1}{2}} = -c_{\frac{1}{2},-\frac{1}{2}} .$$

- For convenience, we define $c \equiv c_{\frac{1}{2},-\frac{1}{2}} = -c_{-\frac{1}{2},\frac{1}{2}}$, in terms of which the singlet state $|00\rangle$ reads

$$|00\rangle = c|\uparrow\downarrow\rangle - c|\downarrow\uparrow\rangle = c\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right) .$$

We then apply the normalization condition $\langle 00|00\rangle \equiv 1$ together with the orthonormality of the states $|\uparrow\rangle$ and $|\downarrow\rangle$ to fully derive the singlet state:

$$1 = 2c^2 \implies c = \tfrac{1}{\sqrt{2}} \implies |00\rangle = \tfrac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right).$$

- The complete set of triplet and singlet states is thus

$$|1m\rangle = \begin{cases} |\uparrow\uparrow\rangle & m = 1 \\ \tfrac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\right) & m = 0 \\ |\downarrow\downarrow\rangle & m = -1 \end{cases}$$

$$|00\rangle = \tfrac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right).$$

9.3 Heisenberg Coupling

Explain the Heisenberg coupling interaction between two particles with spin $s = 1/2$. Be sure to discuss the relevant eigenvalue equation and energy eigenvalues.

- The Heisenberg coupling involves two particles with spin $s_i = 1/2$ and reads

$$H = J_0 \boldsymbol{S}_1 \cdot \boldsymbol{S}_2$$

where J_0 is called the exchange coupling constant. When $J_0 < 0$, system's spins $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ are aligned in the ground state (ferromagnetic coupling). When $J_0 > 0$, the spins $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ are oppositely oriented in the ground state (anti-ferromagnetic coupling).

- Next, we use the identity

$$S^2 = \boldsymbol{S} \cdot \boldsymbol{S} = (\boldsymbol{S}_1 + \boldsymbol{S}_2) \cdot (\boldsymbol{S}_1 + \boldsymbol{S}_2) = S_1^2 + S_2^2 + 2\boldsymbol{S}_1 \cdot \boldsymbol{S}_2$$

to write the Heisenberg coupling term in the form

$$H = \frac{J_0}{2}\left(S^2 - \frac{3}{2}\hbar^2\right).$$

- Note that the coupling term is rotationally invariant, which implies the commutation relation $[\boldsymbol{S}, H] = 0$. Because of the relation $[\boldsymbol{S}, H] = 0$, the quantities S^2 and S_z are conserved, and the singlet and triplet states are eigenstates of the equation

$$H \left|sm\right\rangle = \frac{J_0 \hbar^2}{2} \left(s(s+1) - \frac{3}{2} \right) \left|sm\right\rangle$$

with corresponding energy eigenvalues

$$E_s = J_0 \hbar^2 \begin{cases} \frac{1}{4} & s = 1 \quad \text{(triplet)} \\ -\frac{3}{4} & s = 0 \quad \text{(singlet)} \end{cases}$$

The ground state (with lowest energy) is thus the singlet state, while the triplet states are triply degenerate with energy $\Delta E = J_0 \hbar^2$ above the ground state.

9.4 The Clebsch-Gordan Coefficients

What are the Clebsch-Gordan coefficients? From where do they arise? Explain what they are used for.

- The Clebsch-Gordan coefficients are change-of-basis coefficients used to transform between the product basis and total angular momentum basis when analyzing two particles with angular momenta $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$ and any value of spin—not just $s = 1/2$.

- We consider a two-particle system with two angular momentum-like operators $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$ that obey the usual angular momentum commutation relation

$$\left[J_{\alpha_i}, J_{\beta_j} \right] = i\hbar \delta_{ij} \epsilon_{\alpha\beta\gamma} J_{\gamma_j}.$$

Leaving the tensor product and identity operators implicit, the total angular momentum operator $\boldsymbol{J}$ and corresponding basis states are

$$\left|j_1 m_1 j_2 m_2\right\rangle = \left|j_1 m_1\right\rangle \left|j_2 m_2\right\rangle \qquad \text{and} \qquad \boldsymbol{J} = \boldsymbol{J}_1 + \boldsymbol{J}_2.$$

- The eigenstatates $\left|j_1 j_2 j m\right\rangle$ of the operators $\boldsymbol{J}^2$ and J_z are written in the product basis $\left\{ \left|j_1 m_1 j_2 m_2\right\rangle \right\}$ in the form

$$\left|j_1 j_2 j m\right\rangle = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} \left\langle j_1 m_1 j_2 m_2 | j m \right\rangle \left|j_1 m_1 j_2 m_2\right\rangle$$

The coefficients $\left\langle j_1 m_1 j_2 m_2 | j m \right\rangle$ in the expansion are called the Clebsch-Gordon coefficients and encode the transformation between the total angular momentum basis $\left|j m\right\rangle$ and the product basis $\left|m_1 m_2\right\rangle$.

9.4.1 Example: Addition of Angular Momenta in the Case $1 \times 1/2$

Background

- As an exercise in the addition of angular momentum, we will find the total angular momentum basis representation of an electron in the ground state and first excited state of a hydrogen atom. The excited electron has angular momentum $\boldsymbol{J}_1 = \boldsymbol{L}$

and basis states $\{|lm_l\rangle\}$, while the proton in the hydrogen nucleus has angular momentum $\boldsymbol{J}_2 = \boldsymbol{S}$ and basis states $\{|\uparrow\rangle, |\downarrow\rangle\}$.

The total angular momentum is $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$, and the corresponding basis is

$$\{|j_1 j_2 jm\rangle\} \equiv \{|lsjm\rangle\} = \left\{\left|l\tfrac{1}{2}jm\right\rangle\right\},$$

where the spin $1/2$ proton's angular momentum j_2 is fixed at $j_2 \equiv s = 1/2$. For both $l = 0$ and $l = 1$, the electron's basis states are the spherical harmonics

$$Y_{lm_l}(\theta, \phi) = \langle \boldsymbol{r}|lm_l\rangle.$$

Finally, recall that the angular momentum quantum numbers must satisfy

$$|s + l| \leq j \leq s + l.$$

Simple Example: The Ground State

- We begin with the ground state $l = 0$ and $s = 1/2$, in which case the total angular momentum quantum number can be only $j = 1/2$.

- The corresponding states in both the total angular momentum basis $|lsmj\rangle$ and the product basis $|lm_l sm_s\rangle$ are written

$$
\begin{array}{llll}
j = \tfrac{1}{2} & |lsmj\rangle = \left|0\tfrac{1}{2}\tfrac{1}{2}\tfrac{1}{2}\right\rangle & \Longleftrightarrow & |lm_l sm_s\rangle = \left|00\tfrac{1}{2}\tfrac{1}{2}\right\rangle = |00\rangle |\uparrow\rangle \\
j = -\tfrac{1}{2} & |lsmj\rangle = \left|0\tfrac{1}{2}\tfrac{1}{2} - \tfrac{1}{2}\right\rangle & \Longleftrightarrow & |lm_l sm_s\rangle = \left|00\tfrac{1}{2} - \tfrac{1}{2}\right\rangle = |00\rangle |\downarrow\rangle
\end{array}
$$

The First Excited State

- In the first excited electron state with $l = 1$, we can have both $j = 3/2$ and $j = 1/2$. Here is the plan: We will begin with the state with the largest value of m, which occurs for $j = 3/2$, and generate the $j = 3/2$ states with lower values of m using the ladder operator J_-. Finally, we find the $j = 1/2$ state using the orthonormality of the basis states.

- The state with the largest possible value of m, i.e. $m = 3/2$, occurs when $j = 3/2$ and thus $l = 1$ and $s = 1/2$ and is written $\left|1\tfrac{1}{2}\tfrac{3}{2}\tfrac{3}{2}\right\rangle$ in the good spin basis and $|11\rangle |\uparrow\rangle$ in the product basis. The plan is to then use the ladder operator $J_- = S_- + L_-$ to determine the state with $m = 1/2$, i.e. the state $|jm\rangle = \left|\tfrac{3}{2}\tfrac{1}{2}\right\rangle$.

  To do this, we first apply J_- to the good-spin basis expression $\left|1\tfrac{1}{2}\tfrac{3}{2}\tfrac{3}{2}\right\rangle$ and $S_- + L_-$ to the equivalent product basis expression $|11\rangle |\uparrow\rangle$ to get

$$J_- \left|1\tfrac{1}{2}\tfrac{3}{2}\tfrac{3}{2}\right\rangle = \hbar\sqrt{3}\left|1\tfrac{1}{2}\tfrac{3}{2}\tfrac{1}{2}\right\rangle$$

$$(L_- + S_-)|11\rangle |\uparrow\rangle = \hbar\sqrt{2}|10\rangle |\uparrow\rangle + \hbar|11\rangle |\downarrow\rangle$$

  We then equate the two expressions, which represent the same state, just written in different bases, and divide common terms to get

$$\hbar\sqrt{3}\left|1\tfrac{1}{2}\tfrac{3}{2}\tfrac{1}{2}\right\rangle \equiv \hbar\sqrt{2}|10\rangle |\uparrow\rangle + \hbar|11\rangle |\downarrow\rangle \implies \left|1\tfrac{1}{2}\tfrac{3}{2}\tfrac{1}{2}\right\rangle = \sqrt{\tfrac{2}{3}}|10\rangle |\uparrow\rangle + \sqrt{\tfrac{1}{3}}|11\rangle |\downarrow\rangle.$$

  The coefficients $\sqrt{\tfrac{2}{3}}$ and $\sqrt{\tfrac{1}{3}}$ are Clebsch-Gordan coefficients, are encode the transformation between the good spin basis expression $\left|1\tfrac{1}{2}\tfrac{3}{2}\tfrac{1}{2}\right\rangle$ to the equivalent state, written in the product basis.

- Next, using the just-derived expression for $\left|1\frac{1}{2}\frac{3}{2}\frac{1}{2}\right\rangle$ in the product basis, we repeat the procedure to generate the state with $|jm\rangle = \left|\frac{3}{2} - \frac{1}{2}\right\rangle$. This reads

$$J_- \left|1\frac{1}{2}\frac{3}{2}\frac{1}{2}\right\rangle = 2\hbar \left|1\frac{1}{2}\frac{3}{2} - \frac{1}{2}\right\rangle$$

$$(L_- + S_-)\left[\sqrt{\frac{2}{3}}\left|10\right\rangle\left|\uparrow\right\rangle + \sqrt{\frac{1}{3}}\left|11\right\rangle\left|\downarrow\right\rangle\right] = \frac{2}{\sqrt{3}}\hbar\left|1-1\right\rangle\left|\uparrow\right\rangle + \sqrt{\frac{2}{3}}\hbar\left|10\right\rangle\left|\downarrow\right\rangle + \sqrt{\frac{2}{3}}\hbar\left|10\right\rangle\left|\downarrow\right\rangle + 0$$

$$= \frac{2}{\sqrt{3}}\hbar\left|1-1\right\rangle\left|\uparrow\right\rangle + 2\sqrt{\frac{2}{3}}\hbar\left|10\right\rangle\left|\downarrow\right\rangle$$

As before, we equate the two different basis expressions of the same state to get

$$2\hbar\left|1\frac{1}{2}\frac{3}{2} - \frac{1}{2}\right\rangle = \frac{2}{\sqrt{3}}\hbar\left|1-1\right\rangle\left|\uparrow\right\rangle + 2\sqrt{\frac{2}{3}}\hbar\left|10\right\rangle\left|\downarrow\right\rangle,$$

and then divide like terms to get

$$\left|1\frac{1}{2}\frac{3}{2} - \frac{1}{2}\right\rangle = \sqrt{\frac{1}{3}}\left|1-1\right\rangle\left|\uparrow\right\rangle + \sqrt{\frac{2}{3}}\left|10\right\rangle\left|\downarrow\right\rangle.$$

The resulting Clebsch-Gordan coefficients for the state $\left|1\frac{1}{2}\frac{3}{2} - \frac{1}{2}\right\rangle$ are $\sqrt{\frac{1}{3}}$ and $\sqrt{\frac{2}{3}}$.

- On more application of the ladder operators J_- and $L_- + S_-$ would lead to the state with $|jm\rangle = \left|\frac{3}{2} - \frac{3}{2}\right\rangle$, which is

$$\left|1\frac{1}{2}\frac{3}{2} - \frac{3}{2}\right\rangle = \left|1-1\right\rangle\left|\downarrow\right\rangle.$$

- Finally, we determine the states with $j = 1/2$ and $|jm\rangle = \left|\frac{1}{2} \pm \frac{1}{2}\right\rangle$ with the orthogonality condition

$$\langle lsjm_j|lsj'm_j\rangle = \delta_{jj'},$$

which in our case using $j = 3/2$ and $j = 1/2$ with $m_j = \pm 1/2$ reads

$$\left\langle 1\frac{1}{2}\frac{3}{2} \pm \frac{1}{2}\middle|1\frac{1}{2}\frac{1}{2} \pm \frac{1}{2}\right\rangle = 0.$$

since these two states have different values of j. The desired states are then

$$\left|1\frac{1}{2}\frac{1}{2}\frac{1}{2}\right\rangle = \sqrt{\frac{2}{3}}\left|10\right\rangle\left|\uparrow\right\rangle - \sqrt{\frac{1}{3}}\left|11\right\rangle\left|\downarrow\right\rangle$$

$$\left|1\frac{1}{2}\frac{1}{2} - \frac{1}{2}\right\rangle = -\sqrt{\frac{2}{3}}\left|1-1\right\rangle\left|\uparrow\right\rangle - \sqrt{\frac{1}{3}}\left|10\right\rangle\left|\downarrow\right\rangle.$$

10 Perturbation Theory

10.1 Rayleigh-Schrödinger Method

Discuss the Rayleigh-Schrödinger method for first-order perturbative analysis of a quantum system with a non-degenerate spectrum. Be sure to state and derive the relevant eigenvalue and eigenfunction formulas.

- The Rayleigh-Schrödinger method is used to approximate the energy eigenvalues E_n and eigenfunctions $|n\rangle$ of a quantum system with a Hamiltonian of the form

$$H = H_0 + H',$$

where the term H_0 is a good first approximation for the system, while the term H' is a *perturbation* term with a secondary effect on the system. We assume we are able to analytically solve the stationary Schrödinger equation for H_0, which reads

$$H_0 = |n_0\rangle = E_n^{(0)} |n_0\rangle,$$

More so, we assume H_0's eigenstates are orthonormal and obey $\langle m_0|n_0\rangle = \delta_{mn}$, and that the energy eigenvalues $E_n^{(0)}$ are non-degenerate.

- The first and second-order energy approximations for the energy E_n are

$$E_n^{(1)} = \langle n_0|V|n_0\rangle \qquad \text{and} \qquad E_n^{(2)} = \langle n_0|V|n_1\rangle = \sum_{m \neq n} \frac{|V_{mn}|^2}{E_n^{(0)} - E_m^{(0)}},$$

while the first-order approximation for the eigenvalues $|n\rangle$ is

$$|n_1\rangle = \sum_{m \neq n} = \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} |m_0\rangle.$$

- The energy and eigenfunction corrections, respectively, for arbitrary order j are

$$E_n^{(j)} = \langle n_0|V|n_{j-1}\rangle \qquad \text{and} \qquad |n_j\rangle = \sum_{m \neq n} |m_0\rangle \langle m_0|n_j\rangle.$$

10.1.1 Derivation: Rayleigh-Schrödinger Method

First-Order Energy Correction

- We begin by writing the perturbation term in the form $H' = \lambda V$ where V has units of energy and λ is a dimensional parameter encoding the strength of the perturbation. Our goal is to solve the stationary Schrödinger equation for the *total* Hamiltonian $H = H_0 + H'$, which reads

$$H |n\rangle = E_n |n\rangle.$$

- The first step is to expand the desired eigenstates $|n\rangle$ and energy eigenvalues E_n in powers of the perturbation parameter λ in the form

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots$$
$$|n\rangle = |n_0\rangle + \lambda |n_1\rangle + \lambda^2 |n^2\rangle + \cdots,$$

where $E_n^{(j)}$ and $|n_j\rangle$ are progressively higher-order corrections to the total Hamiltonian H's n-th eigenvalue E_n and eigenstate $|n\rangle$, respectively.

- We proceed by multipliying the equation for $|n\rangle$ by $\langle n_0|$, which gives

$$\langle n_0|n\rangle = 1 + \lambda \langle n_0|n_1\rangle + \lambda^2 \langle n_0|n_2\rangle + \cdots .$$

We then make an important assumption: we assume the complete solution for $|n\rangle$ is well approximated by the lowest-order approximation $|n_0\rangle$, allowing us to temporarily assume $\langle n_0|n\rangle = 1$. Under the assumption $\langle n_0|n\rangle = 1$, the above equation simplifies to

$$\lambda \langle n_0|n_1\rangle + \lambda^2 \langle n_0|n_2\rangle + \cdots = 0,$$

which is satisfied by all $|n_j\rangle$ only if $\langle n_0|n_j\rangle = \delta_{0j}$ for all $j \in \mathbb{N}$.

- We then subsitute the expressions for H, $|n\rangle$ and E_n into the stationary Schrödinger equation $H|n\rangle = E_n|n\rangle$ to get

$$(H_0 + \lambda V)\big(|n_0\rangle + \lambda|n_1\rangle + \cdots\big) = \big(E_n^{(0)} + \lambda E_n^{(1)} + \cdots\big)\big(|n_0\rangle + \lambda|n_1\rangle + \cdots\big).$$

We then equate the coefficients of each power of λ. The result, up to $\lambda^j = \lambda^3$, is

$$
\begin{array}{c|rcl}
\lambda^0 & H_0|n_0\rangle & = & E_n^{(0)}|n_0\rangle \\
\lambda^1 & H_0|n_1\rangle + V|n_0\rangle & = & E_n^{(0)}|n_1\rangle + E_n^{(1)}|n_0\rangle \\
\lambda^2 & H_0|n_2\rangle + V|n_1\rangle & = & E_n^{(0)}|n_2\rangle + E_n^{(1)}|n_1\rangle + E_n^{(2)}|n_0\rangle \\
\lambda^3 & H_0|n_3\rangle + V|n_2\rangle & = & E_n^{(0)}|n_3\rangle + E_n^{(1)}|n_2\rangle + E_n^{(2)}|n_1\rangle + E_n^{(3)}|n_1\rangle .
\end{array}
$$

The first equation is trivial and represents the known eigenvalue relation for the Hamiltonian H_0. The second equation with λ^1 is more useful—we multiply this equation by $\langle n_0|$, use the identity $\langle n_0|H_0 = E_n^{(0)}\langle n_0|$ and apply the orthonormality condition $\langle n_0|n_j\rangle = \delta_{0j}$ to get

$$\langle n_0|H_0|n_1\rangle + \langle n_0|V|n_0\rangle = E_n^{(0)}\langle n_0|n_1\rangle + E_n^{(1)}\langle n_0|n_0\rangle$$
$$E_n^{(0)}\langle n_0|n_1\rangle + \langle n_0|V|n_0\rangle = 0 + E_n^{(1)}$$
$$0 + \langle n_0|V|n_0\rangle = 0 + E_n^{(1)}.$$

In terms of the matrix element $\langle n_0|V|n_0\rangle \equiv V_{nn}$, the end result is

$$E_n^{(1)} = \langle n_0|V|n_0\rangle = V_{nn},$$

which is the first-order energy correction quoted at the begining of the subsection.

Second-Order Energy and First-Order Eigenstate Correction

- We can perform an analogous procedure with the equations for higher powers λ^j to get expressions for the higher-order corrections $E_n^{(1)}$. For example, using the equation for λ^2 and multiplying by $\langle n_0|$ as before gives

$$E_n^{(0)}\langle n_0|n_2\rangle + \langle n_0|V|n_1\rangle = E_n^{(0)}\langle n_0|n_2\rangle + E_n^{(1)}\langle n_0|n_1\rangle + E_n^{(2)}\langle n_0|n_0\rangle ,$$

which produces $E_n^{(2)} = \langle n_0|V|n_1\rangle$ after applying $\langle n_0|n_j\rangle = \delta_{0j}$.

In general, the expression for the j-th correction to the energy E_n is

$$E_n^{(j)} = \langle n_0|V|n_{j-1}\rangle .$$

- Technically, the result $E_n^{(2)} = \langle n_0|V|n_1 \rangle$ is the desired second-order energy correction. However, this result is not particularly useful unless we know $|n_1\rangle$. To find $|n_1\rangle$ we multiply the earlier equation for λ^1 by $|m_0\rangle$, use the earlier identity $\langle m_0|H_0 = E_m^{(0)}\langle m_0|$, apply $\langle n_0|n_j\rangle = \delta_{0j}$ and rearrange to get

$$E_m^{(0)}\langle m_0|n_1 \rangle + \langle m_0|V|n_0\rangle = E_n^{(0)}\langle m_0|n_1\rangle + 0$$

$$\left(E_n^{(0)} - E_m^{(0)}\right)\langle m_0|n_1\rangle = \langle m_0|V|n_0\rangle = V_{mn}.$$

The coefficients in the expansion of $|n_1\rangle$ are thus

$$\langle m_0|n_1\rangle = \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}},$$

which implies the first-order correction $|n_1\rangle$ to the state $|n\rangle$ is

$$|n_1\rangle = \sum_{m\neq n} = \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}}|m_0\rangle,$$

as quoted in the begining of the subsection.

- Using the just-derived result for $|n_1\rangle$, the second-order energy correction $E_n^{(2)}$ is then

$$E_n^{(2)} = \langle n_0|V|n_1\rangle = \sum_{m\neq n} \frac{|V_{mn}|^2}{E_n^{(0)} - E_m^{(0)}}.$$

10.1.2 Extra: Third-Order Energy and Second-Order Eigenstate Correction

- As an exercise in the Rayleigh-Schrödinger method, we will determine the second-order wavefunction correction $|n_2\rangle$, which will also reveal the third-order energy correction $E_n^{(3)}$.

We begin by multiplying the earlier equation for λ^2 by $\langle m_0|$, which gives

$$\langle m_0|H_0|n_2\rangle + \langle m_0|V|n_1\rangle = E_n^{(0)}\langle m_0|n_2\rangle + E_n^{(1)}\langle m_0|n_1\rangle + E_n^{(2)}\langle m_0|n_0\rangle.$$

This equation, after substituting in the results derived earlier in this section, comes out to, in order,

$$E_m^{(0)}\langle m_0|n_2\rangle + \sum_{l\neq n}\frac{V_{ml}V_{ln}}{E_n^{(0)} - E_l^{(0)}} = E_n^{(0)}\langle m_0|n_2\rangle + E_n^{(1)}\frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} + E_n^{(2)}\cdot 0.$$

We then substitute in $E_n^{(1)} = V_{nn}$ and rearrange to get

$$\sum_{l\neq n}\frac{V_{ml}V_{ln}}{E_n^{(0)} - E_l^{(0)}} = \left(E_n^{(0)} - E_m^{(0)}\right)\langle m_0|n_2\rangle + \frac{V_{nn}V_{mn}}{E_n^{(0)} - E_m^{(0)}}.$$

The coefficients $\langle m_0|n_2\rangle$ in the expansion of $E_n^{(3)}$ are thus

$$\langle m_0|n_2\rangle = \sum_{n\neq l}\frac{V_{ml}V_{ln}}{\left(E_n^{(0)} - E_m^{(0)}\right)\left(E_n^{(0)} - E_l^{(0)}\right)} - \frac{V_{mn}v_{nn}}{\left(E_n^{(0)} - E_m^{(0)}\right)^2},$$

and the second-order wavefunction correction is

$$|n_2\rangle = \sum_{m\neq n}\left[\sum_{n\neq l}\frac{V_{ml}V_{ln}}{\left(E_n^{(0)} - E_m^{(0)}\right)\left(E_n^{(0)} - E_l^{(0)}\right)} - \frac{V_{mn}V_{nn}}{\left(E_n^{(0)} - E_m^{(0)}\right)^2}\right]|m_0\rangle.$$

- In terms of the just-derived $|n_2\rangle$, the third-order energy correction $E_n^{(3)}$ is

$$E_n^{(3)} = \sum_{m \neq n} \langle m_0|n_2\rangle \, V_{nm},$$

  where the expression for $|n_2\rangle$ is left out for conciseness.

Normalization

- Finally, we consider the normalization of the $|n\rangle$ states. Up to the first-order correction in λ, the $|n\rangle$ are already normalized, since

$$\langle n|n\rangle = \big(|n_0\rangle + \lambda \, |n_1\rangle \big)\big(|n_0\rangle + \lambda \, |n_1\rangle \big)$$
$$= \langle n_0|n_0\rangle + \lambda \, \langle n_0|n_1\rangle + \lambda \, \langle n_1|n_0\rangle + \mathcal{O}(\lambda^2)$$
$$= 1 + \mathcal{O}(\lambda^2)$$

  However, the $|n\rangle$ are not normalized in the second-order correction, where we have

$$\langle n|n\rangle = \langle n_0|n_0\rangle + \lambda^2 \, \langle n_1|n_1\rangle + \mathcal{O}(\lambda^3)$$
$$= 1 + \lambda^2 \sum_{m \neq n} \frac{|V_{nm}|^2}{\big(E_n^{(0)} - E_m^{(0)}\big)^2} + \mathcal{O}(\lambda^3)$$

  If we want the $|n\rangle$ states to be normalized we have two options:

  1. Renormalize the entire $|n\rangle$ state into

$$|n\rangle \rightarrow \left[\sum_j \langle n_j|n_j\rangle \right]^{-1/2} |n\rangle .$$

  2. Renormalize only the second-order correction via

$$|n_2\rangle \rightarrow |n_2\rangle - \frac{\lambda}{2} \sum_{m \neq n} \frac{|V_{nm}|^2}{\big(E_n^{(0)} - E_m^{(0)}\big)^2} |n_0\rangle .$$

  In terms of this renormalization, the product $\langle n|n\rangle$ comes out to

$$\langle n|n\rangle = 1 + \mathcal{O}(\lambda^3).$$

10.2 First-Order Perturbation for a Degenerate Spectrum

Explain the process for perturbatively finding a degenerate system's energy eigenvalues. Demonstrate the process for a doubly degenerate energy level, and use the result to generalize the procedure to arbitrary degeneracy.

10.2.1 Overview

- As in the Rayleigh-Schrödinger method, we consider a Hamiltonian of the form

$$H = H_0 + \lambda V,$$

where we now assume the energy eigenvalues of the unperturbed Hamiltonian H_0 are N-times degenerate. As a result, the eigenvalue equation for H_0 reads

$$H_0 \left(c_1 \ket{n_0^{(1)}} + \cdots + c_N \ket{n_0^{(N)}} \right) = E_n^{(0)} \left(c_1 \ket{n_0^{(1)}} + \cdots + c_N \ket{n_0^{(N)}} \right),$$

where $E_n^{(0)}$ is the N-time degenerate, n-th eigenvalue of the Hamiltonian H_0.

Our goal is to (approximately) solve the stationary Schrödinger equation

$$H \ket{n} = E \ket{n}$$

for the complete Hamiltonian H's energy eigenvalues E_n.

- Finding the first order energy and eigenstate corrections involves solving the $N \times N$ eigenvalue problem

$$\begin{pmatrix} V_{11} & \cdots & V_{1N} \\ \vdots & \ddots & \vdots \\ V_{N1} & \cdots & V_{NN} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_N \end{pmatrix} = E_n^{(1)} \begin{pmatrix} c_1 \\ \vdots \\ c_N \end{pmatrix},$$

where the matrix elements V_{ij} are given by

$$V_{ij} = \bra{n_0^{(i)}} V \ket{n_0^{(j)}}.$$

The matrix V_{ij}'s eigenvalues are the first-order energy eigenvalue corrections $E_{n_l}^{(1)}$, where $l = 1, 2, \ldots N$, to the total Hamiltonian H's energies E_n.

The energy corrections $E_{n_l}^{(1)}$ and the diagonal matrix elements V_{ii} obey

$$\sum_{l=1}^{N} E_{n_l}^{(1)} = \sum_{i=1}^{N} V_{ii}.$$

- *Note*: the matrix elements $\bra{n_0^{(i)}} V \ket{n_0^{(j)}}$ are small (e.g. much less than one), it is also important to consider transitions between excited states with different values of n and m. In this case, we redefine the matrix elements to be

$$V_{ij} \to V_{ij} + \sum_{m \neq n} \frac{V_{im} V_{mj}}{E_n^{(0)} - E_m^{(0)}},$$

where we leave out the case $m = n$ to avoid a zero in the denominator.

10.2.2 Derivation: The Doubly Degenerate Case

- We begin by assuming the unperturbed Hamiltonian H_0's energy eigenvalues $E_n^{(0)}$ are doubly degenerate. We first expand H's eigenstates and eigenvalues in the form

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots$$
$$\ket{n} = c_1 \ket{n_0^{(1)}} + c_1 \ket{n_0^{(2)}} + \lambda \ket{n_1} + \cdots.$$

Because the unperturbed Hamiltonian H_0's energy levels are doubly degenerate, the zeroth-order energy $E_n^{(0)}$ corresponds to two linearly independent wavefunctions, which we have written as the linear combination $c_1 \ket{n_0^{(1)}} + c_1 \ket{n_0^{(2)}}$.

- Next, working only up to first order in λ, we substitute the expansions for $|n\rangle$ and E_n into the stationary Schrödinger equation and equate the coefficients for λ^0 and λ^1 to get the three equations

$$H_0 |n_0^{(1)}\rangle = E_n^{(0)} |n_0^{(1)}\rangle$$

$$H_0 |n_0^{(2)}\rangle = E_n^{(0)} |n_0^{(2)}\rangle$$

$$H_0 |n_1\rangle + c_1 V |n_0^{(1)}\rangle + c_2 V |n_0^{(2)}\rangle = E_n^{(0)} |n_1\rangle + E_n^{(1)} \big(c_1 |n_0^{(1)}\rangle + c_2 |n_0^{(2)}\rangle \big).$$

- Next, we multiply the last equation through by $|n_0^{(1)}\rangle$ to get an equation for the coefficients c_1 and c_2:

$$\langle n_0^{(1)}|H_0|n_1\rangle + c_1 \langle n_0^{(1)}|V|n_0^{(1)}\rangle + c_2 \langle n_0^{(1)}|V|n_0^{(2)}\rangle$$
$$= E_n^{(0)} \langle n_0^{(1)}|n_1\rangle + E_n^{(1)} \big(c_1 \langle n_0^{(1)}|n_0^{(1)}\rangle + c_2 \langle n_0^{(1)}|n_0^{(2)}\rangle \big).$$

We then apply the orthonormality of the $|n_j\rangle$ states, simplifying the equation to

$$0 + c_1 V_{11} + c_2 V_{12} = 0 + E_n^{(1)}(c_1 + 0),$$

where we have defined the matrix elements

$$V_{ij} = \langle n_0^{(i)}|V|n_0^{(j)}\rangle .$$

- We then perform a similar procedure in which we multiply the last equation through by $|n_0^{(2)}\rangle$ to get a second equation for the coefficients c_1 and c_2. In one place, these two equations are

$$c_1 V_{11} + c_2 V_{12} = E_n^{(1)} c_1$$

$$c_1 V_{21} + c_2 V_{22} = E_n^{(1)} c_2.$$

This system of equations is a 2×2 eigenvalues problem of the form

$$\begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = E_n^{(1)} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

Solving this eigenvalue problem determines the desired first-order perturbation theory correction—the eigenvalues are the two first-order energy corrections $E_{n_1}^{(1)}$ and $E_{n_2}^{(1)}$ and the corresponding eigenvectors determine the coefficients c_1 and c_2 of the eigenstate corrections $|n_0^{(1)}\rangle$ and $|n_0^{(2)}\rangle$.

10.2.3 The General N-Times Degenerate Case

- The procedure for an N-times degenerate energy level is a straightforward generalization of the $N = 2$ case and leads to a corresponding $N \times N$ eigenvalue problem with matrix elements V_{ij} given by

$$V_{ij} = \langle n_0^{(i)}|V|n_0^{(j)}\rangle$$

- Solving the $N \times N$ eigenvalue problem's corresponding characteristic polynomial

$$\det \big[V_{ij} - E_n^{(1)}\delta_{ij}\big] = 0$$

leads to the first-order energy eigenvalue corrections $E_{n_l}^{(1)}$ where $l = 1, 2, \ldots N$.

Solving the associated eigenvector problem leads to the coefficients c_{i_l} for the eigenstate corrections $|n_0^{(i)}\rangle$.

10.3 Perturbation Theory for a Time Dependent Hamiltonian

Explain the process for perturbatively finding the eigenfunctions of a system with a time-dependent Hamiltonian.

Explain the interaction picture and state and derive the relevant equations.

What are Rabi oscillations? Give a physical situation where the phenomenon might apply, and state the relevant equations.

- We consider a quantum system with a time-dependent Hamiltonian of the form

$$H(t) = H_0 + \lambda V(t),$$

  where H_0, is a time-independent, diagonalizable Hamiltonian with known eigenvalue relation $H_0 |n\rangle = E_n |n\rangle$ and orthonormal eigenstates that reasonably approximates the system, while $V(t)$ is a time-dependent perturbation term of secondary influence.

- In time dependent perturbation theory, the goal is to solve for the system's time-dependent wavefunction $|\psi(t)\rangle$ given an initial state $|\psi(0)\rangle$.

- In the interaction picture we write the states $|\psi(0)\rangle$ and $|\psi(t)\rangle$ in the H_0 basis $\{|n\rangle\}$ in the form

$$|\psi(0)\rangle = \sum_n c_n(0) |n\rangle \qquad \text{and} \qquad |\psi(t)\rangle = \sum_n c_n(t) e^{-i\frac{E_n}{\hbar}t} |n\rangle,$$

  where the coefficients $c_n(t)$ are given by the coupled system of differential equations

$$i\hbar \frac{\partial c_m(t)}{\partial t} e^{-i\frac{E_m}{\hbar}t} = \lambda \sum_n \langle m|V(t)|n\rangle \, e^{-i\frac{E_n}{\hbar}t} c_n(t),$$

  or the equivalent matrix system of differential equations

$$i\hbar \frac{\partial}{\partial t} \boldsymbol{c}(t) = \lambda \mathbf{V}(t)\boldsymbol{c}(t), \qquad \text{where} \qquad V_{mn}(t) = \langle m|V(t)|n\rangle \, e^{-i\frac{E_n - E_m}{\hbar}t}.$$

- Assuming the initial state is $|\psi(0)\rangle = |m\rangle$, the first order approximation for the coefficients $c_n(t)$ in the expression for the time-dependent wavefunction $|\psi(t)\rangle$ is

$$c_n^{(1)}(t) = \frac{1}{i\hbar} \int_0^t V_{nm}(t') \, \mathrm{d}t'.$$

  In terms of $c_n^{(1)}$, the state $|\psi(t)\rangle$ is given to first-order in λ as

$$|\psi(t)\rangle = \sum_n \left[c_n^{(0)}(t) + \lambda c_n^{(1)}(t) + \mathcal{O}(\lambda^2) \right] e^{-i\frac{E_n}{\hbar}t} |n\rangle.$$

  The first-order approximation is valid in the regime $\left| c_n^{(1)}(t) \right| \ll 1$. In practice, the first-order approximation gives an acceptable result for many physical problems.

10.3.1   Derivation: The Interaction Picture

- We begin our analysis by writing the initial and time-dependent states $|\psi(0)\rangle$ and $|\psi(t)\rangle$, respectively, in the H_0 basis $\{|n\rangle\}$ in the form

$$|\psi(0)\rangle = \sum_n c_n(0)\,|n\rangle \qquad \text{and} \qquad |\psi(t)\rangle = \sum_n c_n(t)e^{-i\frac{E_n}{\hbar}t}\,|n\rangle\,,$$

  where the coefficients $c_n(t)$ are time-dependent, since $H(t)$ is time-dependent.

- We then substitute the general expression for $|\psi(t)\rangle$ into the Schrödinger equation,

$$i\hbar\frac{\partial}{\partial t}\,|\psi(t)\rangle = H\,|\psi(t)\rangle\,,$$

  which leads to

$$i\hbar\sum_n\left(\frac{\partial c_n(t)}{\partial t}e^{-i\frac{E_n}{\hbar}t} - i\frac{E_n}{\hbar}c_n(t)e^{-i\frac{E_n}{\hbar}t}\right)|n\rangle = \sum_n\left(E_n + \lambda V(t)\right)c_n(t)e^{-i\frac{E_n}{\hbar}t}\,|n\rangle\,.$$

  Next, we multiply the equation by the basis vector $\langle m|$, apply the orthonormality relation $\langle m|n\rangle = \delta_{mn}$, and cancel like terms, and divide through by $e^{i\frac{E_m}{\hbar}t}$ to get

$$i\hbar\frac{\partial c_m(t)}{\partial t}e^{-i\frac{E_m}{\hbar}t} = \lambda\sum_n\langle m|V(t)|n\rangle\,e^{-i\frac{E_n}{\hbar}t}c_n(t)$$

$$i\hbar\frac{\partial c_m(t)}{\partial t} = \lambda\sum_n\langle m|V(t)|n\rangle\,e^{-i\frac{(E_n - E_m)}{\hbar}t}c_n(t),$$

  which is a system of coupled differential equations for the coefficients $c_n(t)$.

- Next, for shorthand, we define the modified matrix elements

$$V_{mn}(t) = \langle m|V(t)|n\rangle\,e^{-i\frac{E_n - E_m}{\hbar}t},$$

  in terms of which the above system of equations for the coefficients c_n reads

$$i\hbar\frac{\partial c_m(t)}{\partial t} = \lambda\sum_n V_{mn}(t)c_n(t).$$

- Alternatively, we can write the above system of equations in the matrix form

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix} c_1(t) \\ c_2(t) \\ \vdots \\ c_m(t) \\ \vdots \end{pmatrix} = \lambda\begin{pmatrix} V_{11}(t) & V_{12}(t) & \cdots & V_{1n}(t) & \cdots \\ V_{21}(t) & V_{22}(t) & \cdots & V_{2n}(t) & \cdots \\ \vdots & \vdots & \ddots & \vdots & \cdots \\ V_{m1}(t) & V_{m2}(t) & \cdots & V_{mn}(t) & \cdots \\ \vdots & \vdots & & \vdots & \ddots \end{pmatrix}\begin{pmatrix} c_1(t) \\ c_2(t) \\ \vdots \\ c_m(t) \\ \vdots \end{pmatrix},$$

  or the more compact equation

$$i\hbar\frac{\partial}{\partial t}\boldsymbol{c}(t) = \lambda\mathbf{V}(t)\boldsymbol{c}(t).$$

  The time-dependent matrix elements $V_{mn}(t)$ are defined, as above, by

$$V_{mn}(t) = \langle m|V(t)|n\rangle\,e^{-i\frac{E_n - E_m}{\hbar}t},$$

  where E_n and E_m are H_0's energy eigenvalues. This formalism is called the interaction picture, and the system's quantum state is determined by the vector $\boldsymbol{c}(t)$.

10.3.2 Derivation: First-Order Time-Dependent Perturbation Approach

- We consider a perturbation that begins at the time $t = 0$, of the form

$$V(t) = \begin{cases} 0 & t < 0 \\ \lambda V(t) & t \geq 0 \end{cases}$$

  and begin by expanding the coefficients c_n in the interaction picture in the form

$$c_n = c_n^{(0)} + \lambda c_n^{(1)} + \lambda^2 c_n^{(2)} + \cdots.$$

- We then assume the system occurs in one of the H_0 eigenstates, i.e. $|\psi(0)\rangle = |m\rangle$, expand the initial and time-dependent states in the H_0 basis in the form

$$|m\rangle \equiv |\psi(0)\rangle = \sum_n c_n(0) |n\rangle \qquad \text{and} \qquad |\psi(t)\rangle = \sum_n c_n(t) e^{-\frac{E_n}{\hbar} t} |n\rangle.$$

  Next, we multiply the initial state equation through by $\langle m|$ and apply $\langle m|n\rangle = \delta_{mn}$, which in the zeroth-order approximation implies

$$c_n^{(0)}(0) = \delta_{nm}.$$

- We then approximate the interaction-picture system of equations for the first-order correction of the coefficients c_k as

$$i\hbar \frac{\partial c_k^{(1)}(t)}{\partial t} = \lambda \sum_n V_{kn}(t) c_n^{(0)}(t),$$

  and make an important assumption: We assume that the time-dependent coefficients $c_n^{(0)}(t)$ are well approximated by their initial values $c_n^{(0)}(0)$. Interpreted physically, this just means that the perturbation term $\lambda V(t)$ is weak enough that it does not considerably disturb the system from its initial configuration.

- Using the approximation $c_n^{(0)}(t) \approx c_n^{(0)}(0)$, we can substitute in $c_n^{(0)}(0) = \delta_{nm}$ into the system of equations for the first-order correction $c_k^{(1)}(t)$ to get

$$i\hbar \frac{\partial c_k^{(1)}(t)}{\partial t} \approx \lambda \sum_n V_{kn}(t) c_n^{(0)}(0) = \lambda \sum_n V_{kn}(t) \delta_{nm} = V_{km}(t).$$

  Comparing the first and last term in the equation and integrating the resulting differential equation gives us the expression for the coefficients $c_k^{(1)}(t)$ in the first-order perturbation approximation:

$$i\hbar \frac{\partial c_k^{(1)}(t)}{\partial t} = V_{km}(t) \implies c_k^{(1)}(t) = \frac{1}{i\hbar} \int_0^t V_{km}(t') \, dt'.$$

- Using the just-derived $c_k^{(1)}$, the state $|\psi(t)\rangle$ is given to first-order in λ as

$$|\psi(t)\rangle = \sum_k \left[c_k^{(0)}(t) + \lambda c_k^{(1)}(t) + \mathcal{O}(\lambda^2) \right] e^{-i\frac{E_k}{\hbar} t} |k\rangle.$$

10.3.3   Example: Two-State System and Rabi Oscillations

- As a simple example of the time-dependent formalism, we consider a two-state system corresponding to the two quantum states $|1\rangle$ and $|2\rangle$ and 2×2 Hamiltonian H_0. We take our time-dependent perturbation term to be a harmonic oscillation of the form $V(t) \sim e^{-i\omega t}$. A physical example of such a state would be an atom with two energy levels upon which we shine monochromatic light.

- We begin by writing the two state system's Hamiltonian in the form

$$H = H_0 + V(t),$$

  where H_0 and $V(t)$ are given by

$$H_0 = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} \qquad \text{and} \qquad V(t) = \hbar \begin{pmatrix} 0 & \Delta e^{i\omega t} \\ \Delta e^{-i\omega t} & 0 \end{pmatrix}.$$

- The system's time-dependent Schrödinger equation reads

$$i\hbar \begin{pmatrix} \dot{c}_1(t) \\ \dot{c}_2(t) \end{pmatrix} = \hbar \begin{pmatrix} 0 & \Delta e^{i(\omega - \tilde{\omega})t} \\ \Delta e^{i(\omega - \tilde{\omega})t} & 0 \end{pmatrix} \begin{pmatrix} c_1(t) \\ c_2(t) \end{pmatrix}$$

  where we have defined $\hbar\tilde{\omega} = E_2 - E_1$.

- We assume the system's initial state is the ground state $|\psi(0)\rangle = |1\rangle$, which implies

$$c_1(0) = 1 \qquad \text{and} \qquad c_2(t) = 0.$$

  In this case, the probabilities P_2 and P_1 of the system occupying the states $|2\rangle$ and $|1\rangle$, respectively, oscillate according to

$$P_2(t) = \frac{\Delta^2}{\Omega} \sin^2 \Omega t \qquad \text{and} \qquad P_1 = 1 - P_2,$$

  where we have defined the Rabi frequency Ω according to

$$\Omega = \sqrt{\Delta^2 + \tfrac{1}{4}(\omega - \tilde{\omega})^2}.$$

  The sinusoidal oscillations of the occupation probabilities P_1 and P_2 are called Rabi oscillations, and have maximum amplitude, exactly equal to one, when $\omega = \tilde{\omega}$.

10.4   Transition Between States and Fermi's Golden Rule

Discuss the probability for transition between states under the influence of a step-like potential. State, derive and interpret both the general first-order perturbative formula and Fermi's golden rule, and explain in which situations the latter applies.

- Consider a Hamiltonian of the form $H = H_0 + \lambda V(t)$ where H_0 is time-independent and diagonalizable with known eigenstates $|n\rangle$, the potential is the step-function

$$V(t) = \begin{cases} 0 & t < 0 \\ V_0 & t \geq 0, \end{cases}$$

  and the system occurs in the initial H_0 eigenstate $|\psi(0)\rangle = |m\rangle$.

114

- The probability of a transition from the initial state $|m\rangle$ to final H_0 eigenstate $|n\rangle$ is

$$P_{nm} = \frac{2\pi}{\hbar}|V_{nm}|^2 \delta_t(E_n - E_m)t \qquad \text{where} \qquad \delta_t(x) \equiv \frac{1}{\pi}\frac{\sin^2(xt)}{x^2 t},$$

where the matrix elements are defined, as in the interaction picture, via

$$V_{nm} = \langle n|V_0|m\rangle \, e^{-i\frac{E_n - E_m}{\hbar}t}.$$

- *Interpretation*: The function δ_t approaches the Dirac delta function in the limit $t \gg \frac{\hbar}{|E_n - E_m|}$. In this regime, there is appreciable likelihood only for transitions to states with energies in a neighborhood of the initial energy E_m.

  Meanwhile, for $t \ll \frac{\hbar}{|E_n - E_m|}$, the function δ_t is very wide, which corresponds to appreciable probability for transitions to states with energies in the range $E_m \pm \frac{\hbar}{t}$. In general, the width of the function $\delta_t(E_n - E_m)$ decreases with time according to

$$\Delta E \sim \frac{\hbar}{t}.$$

- Fermi's golden rule applies to the special case in which the system transitions to a final state $|n\rangle$ for which there are many states with energies close to the final energy E_n and for large times $t \gg \frac{\hbar}{|E_n - E_m|}$. In this case we can define the density of states

$$\rho(E_n) = \frac{\mathrm{d}N}{\mathrm{d}E_n},$$

  where $\mathrm{d}N$ represents the number of states available in a small energy band $\mathrm{d}E_n$ centered around E_n. In this case, Fermi's golden rule states that the probability for transition from the initial state $|m\rangle$ to the final state $|n\rangle$ chages with time as

$$P_{mn}(t) = \frac{2\pi}{\hbar}|V_{nm}|^2 \rho(E_m)t \implies \frac{\mathrm{d}P_{mn}(t)}{\mathrm{d}t} = \frac{2\pi}{\hbar}|V_{nm}|^2 \rho(E_m).$$

Derivation: General Transition Between States

- We begin with the first-order perturbative approximation to the interaction picture,

$$c_n^{(1)}(t) = \frac{1}{i\hbar}\int_0^t V_{nm}(t')\,\mathrm{d}t' \qquad \text{where} \qquad V_{mn}(t) = \langle m|V(t)|n\rangle \, e^{-i\frac{E_n - E_m}{\hbar}t}.$$

  Since $V(t > 0) = V_0$ is constant, the matrix elements $V_{km} = \langle k|V_0|m\rangle$ are also constant, and the coefficients in the first-order perturbation approximation are

$$c_n^{(1)}(t) = \frac{1}{i\hbar}\int_0^t V_{nm}e^{-i\frac{E_m - E_n}{\hbar}t'}\,\mathrm{d}t' = \frac{V_{nm}}{i\hbar}\frac{e^{i\omega_{mn}t} - 1}{-i\omega_{mn}},$$

  where we have defined $\hbar\omega_{mn} = E_n - E_m$.

- In terms of $c_n^{(1)}(t)$ and ω_{mn}, the probability of a transition from $|m\rangle$ to $|n\rangle$ is

$$P_{nm} = \left|c_n^{(1)}\right|^2 = \frac{|V_{nm}|^2}{\hbar^2}\frac{\left|e^{-i\omega_{mn}t} - 1\right|^2}{\omega_{mn}^2} = \frac{|V_{nm}|^2}{\hbar^2}\frac{\sin^2\left(\frac{1}{2}\omega_{mn}t\right)}{\left(\frac{1}{2}\omega_{mn}\right)^2}$$

$$\equiv \frac{2\pi}{\hbar}|V_{nm}|^2 \delta_t(E_n - E_m)t,$$

where we have defined the function

$$\delta_t(x) \equiv \frac{1}{\pi} \frac{\sin^2(xt)}{x^2 t}.$$

Keep in mind that this analysis of the transition probability P_{km}, which uses first-order time-dependent perturbation theory, rests on the assumption $P_{nm}(t) \ll 1$, which is necessary for the first-order approximation to hold.

Derivation: Fermi's Golden Rule

- We consider the special case in which our system transitions from the intial H_0 eigenstate $|m\rangle$ to a final state $|n\rangle$ for which there are many states with energies very close to the final energy E_n. We describe the densely-spaced states using the concept of a density of states (as in e.g. thermodynamics), which we define as

$$\rho(E_k) = \frac{\mathrm{d}N}{\mathrm{d}E_k},$$

  where $\mathrm{d}N$ represents the number of states available in a small energy band $\mathrm{d}E_n$ centered around E_n.

- The total probability for a transition to any state in the neighborhood of E_n is an integral over E_n of the form

$$P(t) = \int P_{nm}(t)\rho(E_n)\,\mathrm{d}E_n.$$

For large $t \gg \frac{\hbar}{|E_n - E_m|}$, the transition probability increases with time as

$$P(t) = \frac{2\pi}{\hbar}|V_{nm}|^2 \rho(E_m)t.$$

The rate of change of probability with respect to time is thus constant and obeys

$$\frac{\mathrm{d}P_{mn}(t)}{\mathrm{d}t} = \frac{2\pi}{\hbar}|V_{nm}|^2 \rho(E_m),$$

which is Fermi's golden rule.

10.5   The WKB Approximation

Provide an overview of the WKB *method in the context of quantum mechanics. Define the* WKB *approximation, and explain how it connects quantum and classical mechanics.*

- In the context of quantum mechanics WKB method is an expansion in terms of $\hbar$, which represents an expansion with respect to the classical limit $\hbar \to 0$.

- We begin by writing the system's wavefunction Ψ with the ansatz

$$\Psi(\boldsymbol{r},t) = e^{\frac{i}{\hbar}S(\boldsymbol{r},t)}, \qquad S(\boldsymbol{r},t) \in \mathbb{C}.$$

- In the limit $\hbar \to 0$, the Schrödinger equation for the above wavefunction reduces to

$$\frac{\partial S}{\partial t} = \frac{1}{2m}(\nabla S)^2 + V,$$

which is a classical Hamilton-Jacobi equation for the principle Hamiltonian function S for a classical particle with velocity $\boldsymbol{v} = \frac{1}{m}\nabla S$. From classical mechanics, we know that Hamilton-Jacobi theory is equivalent to the formalism of Newton's second law, so the limit $\hbar \to 0$ essentially recovers classical mechanics.

- To first order in $\hbar$, the Schrödinger equation for Ψ becomes

$$-\frac{\partial S_1}{\partial t} = \frac{1}{2m}\big(2\nabla S_0 \cdot \nabla S_1 - i\nabla^2 S_0\big),$$

Analysis in the regime is called the WKB or semi-classical approximation, and can be thought of as a first-order quantum correction to classical mechanics.

Derivation: The WKB Approximation

- We begin by writing the system's wavefunction Ψ with the ansatz

$$\Psi(\boldsymbol{r}, t) = e^{\frac{i}{\hbar}S(\boldsymbol{r},t)}, \qquad S(\boldsymbol{r}, t) \in \mathbb{C}.$$

We then substitute this ansatz into the Schrödinger equation,

$$i\hbar\frac{\partial \Psi}{\partial t} = i\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi,$$

which leads to the partial differential equation

$$-\frac{\partial S}{\partial t}\Psi = \left(\frac{1}{2m}(\nabla S)^2 - i\frac{\hbar^2}{2m}\nabla^2 S + V\right)\Psi.$$

Assuming $\Psi \neq 0$, we can then divide through by Ψ to get

$$-\frac{\partial S}{\partial t} = \frac{1}{2m}(\nabla S)^2 + V - i\frac{\hbar^2}{2m}\nabla^2 S.$$

- Next, we note that in the limit $\hbar \to 0$, the above equation reduces to

$$\frac{\partial S}{\partial t} = \frac{1}{2m}(\nabla S)^2 + V,$$

which is the classical Hamilton-Jacobi equation quoted at the begining of the subsection.

- We proceed with the WKB method by expanding S in powers of $\hbar$, in the form

$$S = S_0 + \hbar S_1 + \hbar^2 S_2 + \cdots,$$

where the zeroth order expansion in $\hbar$ produces the aforementioned classical equation

$$\frac{\partial S}{\partial t} = \frac{1}{2m}(\nabla S)^2 + V,$$

while the first order in $\hbar$ corresponds to the first-order WKB approximation.

$$-\frac{\partial S_1}{\partial t} = \frac{1}{2m}\big(2\nabla S_0 \cdot \nabla S_1 - i\nabla^2 S_0\big).$$

Example: The WKB Approximation

- As an example, we consider the one-dimensional stationary states

$$\Psi(x, t) = e^{-i\frac{E}{\hbar}t}\psi(x).$$

We will write S and ψ as the time-independent functions

$$S(x) = S_0(x) + \hbar S_1(x) + \mathcal{O}(\hbar^2) \qquad \text{and} \qquad \psi(x) = e^{\frac{i}{\hbar}S(x)}.$$

- We begin by solving for $S_0(x)$ using the zeroth-order classical equation, which gives

$$E = \frac{1}{2m}\left(\frac{\mathrm{d}S_0(x)}{\mathrm{d}x}\right)^2 + V(x)$$

and

$$S_0(x) = \pm \int_{x_0}^{x} \sqrt{2m\left[E - V(x')\right]}\, \mathrm{d}x' \equiv \pm \int_{x_0}^{x} p(x')\, \mathrm{d}x',$$

where we have defined $p(x) = \sqrt{2m[E - V(x)]}$ for shorthand. We then use this expression for $S_0(x)$ to solve for S_1 via

$$2\frac{\mathrm{d}S_0(x)}{\mathrm{d}x}\frac{\mathrm{d}S_1(x)}{\mathrm{d}x} = i\frac{\mathrm{d}^2 S_0(x)}{\mathrm{d}x^2},$$

which leads to

$$S_1(x) = \frac{i}{2}\int_{x_0}^{x} \frac{S_0''(\tilde{x})}{S_0'(\tilde{x})}\, \mathrm{d}\tilde{x} = \frac{i}{2}\ln\frac{p(x)}{p(x_0)}.$$

The approximate solution for the wavefunction in the WKB approximation is thus

$$\psi_{\mathrm{wkb}}(x, t) = \frac{C}{\sqrt{p(x)}}\exp\left[-i\frac{E}{\hbar}t \pm \frac{i}{\hbar}\int_{x}^{x_0} p(x')\, \mathrm{d}x'\right].$$

This solution represents a modulated wave, with a position-dependent wave vector

$$k(x) = \frac{p(x)}{\hbar}.$$

The plus/minus sign in the exponent encodes the wave's direction, just like in a plane wave, and depends on the problem's concrete boundary conditions. The constants C and x_0 are determed by normalization and boundary conditions, respectively.

10.6 The Variational Method

Explain the variational method and what it is used for. State and derive the relevant equations.

- The variational method is used to find the ground state energy E_0 of a quantum system with Hamiltonian H and wavefunction ψ, and rests on the inequality

$$E_0 \leq \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}.$$

In other words, the H ground state energy E_0 is smaller than the energy expectation value for any state $|\psi\rangle$ expanded in the H basis $\{\,|n\rangle\,\}$.

- The variational method proceeds as follows:

  1. First, we choose $|\psi\rangle$ to be a good approximation to the Hamiltonian's ground state, and require that $|\psi\rangle$ has the same symmetry properties as H.

  2. We then parametrize the wavefunction ψ with n variational parameters α_i in the form $\psi = \psi(\alpha_1, \alpha_2, \ldots, \alpha_n; x)$.
     Using this trial function, we then calculate the expectation value

     $$\langle \psi | H | \psi \rangle = \mathcal{E}(\alpha_1, \ldots, \alpha_n).$$

  3. Next, by varying the parameteris α_i, we determine the minimum energy $\mathcal{E}_{\min}$, which is the best approximation to the ground state within the space of trial functions generated by the parameters $\alpha_1, \ldots, \alpha_n$.

  4. Because of the earlier inequality, the energy $\mathcal{E}_{\min}$ must obey

     $$E_0 \leq \mathcal{E}_{\min},$$

     which theoretically allows us to arbitrarily increase the number of parameters and trial functions, and repeat the process of finding the minimum energy expectation value $\mathcal{E}_{\min}$ until $\mathcal{E}_{\min}$ is an arbitrarily close approximation of E_0, since $E_0 \leq \mathcal{E}_{\min}$ gaurantees we will never drop below E_0.

- The variational method can be used to show that in the Rayleigh-Schrödinger method for solving a Hamiltonian of the form

  $$H = H_0 + \lambda V,$$

  the Hamiltonian H's true ground state energy E_0 is less than the sum of the first two perturbative energy corrections $E_0^{(0)}$ and $E_0^{(1)}$, i.e.

  $$E_0 \leq E_0^{(0)} + E_0^{(1)}.$$

  Derivation: Let H_0 have the known eigenvalue relation

  $$H_0 \, |\varphi_n^{(0)}\rangle = E_n^{(0)} \, |\varphi_n^{(0)}\rangle \, .$$

  We then calculate the expectation value $\langle H \rangle$ to first order in λ, which reads

  $$\langle H \rangle = \langle \psi | H | \psi \rangle = E_0^{(0)} + \lambda \, \langle \varphi_0^{(0)} | V | \varphi_0^{(0)} \rangle + \mathcal{O}(\lambda^2) = E_0^{(0)} + E_0^{(1)} + \mathcal{O}(\lambda^2).$$

  The inequality underlying the variational principle gaurantees $E_0 \leq \langle H \rangle$, which implies

  $$E_0 \leq E_0^{(0)} + E_0^{(1)}.$$

Derivation: The Variational Method Inequality

- We consider a Hamiltonian with orthornormalized but otherwise unknown eigenstates and eigenvalues satisfying

  $$H \, |n\rangle = E_n \, |n\rangle \, , \qquad \langle m | n \rangle = \delta_{mn}.$$

  We then expand an arbitrary wavefunction $|\psi\rangle$ in the $\{\, |n\rangle \,\}$ basis in the form

  $$|\psi\rangle = \sum_n c_n \, |n\rangle = \sum_n \langle n | \psi \rangle \, |n\rangle \, .$$

- The energy expectation value $\langle H \rangle$ for the state $|\psi\rangle$ is

$$\langle\psi|H|\psi\rangle = \left\langle \sum_m \langle m|\psi\rangle \langle m| \middle| H \middle| \sum_n \langle n|\psi\rangle |n\rangle \right\rangle$$

$$= \left\langle \sum_m \langle m|\psi\rangle \langle m| \middle| \sum_n \langle n|\psi\rangle E_n |n\rangle \right\rangle$$

$$= \sum_n E_n |c_n|^2.$$

- We can bound this result from above according to

$$E_0 \sum_n |c_n|^2 \le \sum_n E_n |c_n|^2,$$

  where E_0 is the ground state energy, which obeys $E_0 \le E_1 \le \cdots \le E_n \le \cdots$.

- The above upper bound implies

$$\langle\psi|H|\psi\rangle \ge E_0 \sum_n |c_n|^2 = E_0 \langle\psi|\psi\rangle,$$

  which we finally rearrange to get the desired variational method inequality

$$E_0 \le \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}.$$

11 Scattering

11.1 The Scattering Matrix in 1D

Discuss the analysis of the one-dimensional scattering problem. Be sure to discuss the general ansatz for stationary scattering states, the scattering matrix, and the transfer matrix.

11.1.1 Introductory Formalism

- We consider one-dimensional scattering problems involving a Hamiltonian of the form

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x).$$

  In the absence of a potential, i.e. for $V(x) = 0$, the Hamiltonian H's stationary states are plane waves with wavefunctions and energies of the form

$$\langle x|\psi\rangle \sim e^{\pm ikx} \qquad \text{with} \qquad E_p = \frac{p^2}{2m} > 0.$$

  Each energy eigenvalue E_p is doubly degerate, since the linearly independent states with momenta $\pm p = \pm\hbar k$ have the same energy.

- In scattering problems we have $V(x) \neq 0$; however, we assume the potential is nonzero only in a finite region $x \in [x_a, x_b]$ and also assume the potential does not allow for bound states. In this case, we write stationary scattering states with the ansatz

$$\psi(x) = \begin{cases} A_1 e^{ikx} + B_1 e^{-ikx} & x < x_a \\ \psi_{ab}(x) & x \in [x_a, x_b] \\ A_2 e^{-ikx} + B_2 e^{ikx} & x > x_a. \end{cases}$$

  In other words, we assume the wavefunction is a linear superposition of plane waves in the region of zero potential and some wavefunction ψ_{ab} in the region with nonzero potential $V(x), x \in [x_a, x_b]$. The A coefficients denote waves moving toward the potential barrier and the B coefficients denote waves moving away from the barrier.

Incoming and Outgoing Vectors

- Next, using the coefficients A_1, B_1, A_2 and B_2, we introduce two two-dimensional column vectors, Ψ_{in} and Ψ_{out}, which encode the waves moving toward and away from the potential, respectively. The vectors are:

$$\Psi_{\text{in}} = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \qquad\qquad \Psi_{\text{out}} = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$$

$$\Psi_{\text{in}}^\dagger = \begin{pmatrix} A_1^*, A_2^* \end{pmatrix} \qquad\qquad \Psi_{\text{out}}^\dagger = \begin{pmatrix} B_1^*, B_2^* \end{pmatrix}.$$

  Interpretation: The A coefficients in the ansatz for ψ correspond to plane waves moving toward the potential, while the B coefficients correspond to plane waves away from the potential, which motivates the definitions of Ψ_{in} and Ψ_{out}.

- The coefficients A and B obey the fundamental relationship

$$|A_1|^2 + |A_2|^2 = |B_1|^2 + |B_2|^2,$$

Interpretation: The sum of the currents entering the potential equals the sum of the currents leaving the potential, which may also be written

$$\Psi_{\text{out}}^\dagger \Psi_{\text{out}} = \Psi_{\text{in}}^\dagger \Psi_{\text{in}}.$$

As a result, the system's probability current density is constant and equal to

$$j_a = \frac{\hbar k}{m}\left(|A_1|^2 - |B_1|^2\right) = j_b = \frac{\hbar k}{m}\left(|B_2|^2 - |A_2|^2\right).$$

11.1.2   The Scattering Matrix

- The coefficient vectors Ψ_{in} and Ψ_{out} are related by a scattering matrix $\mathbf{S}$ via

$$\Psi_{\text{out}} = \mathbf{S}\Psi_{\text{in}}.$$

The scattering matrix is unitary, which we can prove using the relationship

$$\Psi_{\text{out}}^\dagger \Psi_{\text{out}} = \Psi_{\text{out}}^\dagger \mathbf{S}\Psi_{\text{in}} = \Psi_{\text{in}}^\dagger \mathbf{S}^\dagger \mathbf{S}\Psi_{\text{in}}.$$

Since $\Psi_{\text{out}}^\dagger \Psi_{\text{out}}^\dagger = \Psi_{\text{in}}^\dagger \Psi_{\text{in}}^\dagger$, it follows that $\mathbf{S}^\dagger \mathbf{S} = \mathbf{I}$ and thus $\mathbf{S}^\dagger = \mathbf{S}^{-1}$, which means $\mathbf{S}$ is indeed unitary.

- By convention, we parameterize the scattering matrix in the form

$$\mathbf{S} = \begin{pmatrix} r & t' \\ t & r' \end{pmatrix} \in \mathbb{C}^{2\times 2},$$

from which follow the relationships

$$\begin{pmatrix} r \\ t \end{pmatrix} = \mathbf{S}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} t' \\ r' \end{pmatrix} = \mathbf{S}\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The parameters t and t' encode the probability for a wave packet to pass through the potential (to transmit, hence the letter t), while the parameters r and r' encode the probability for a wave packet to reflect off the potential barrier.

- The scattering matrix parameters are related according to the relationships

$$\mathbf{S}^\dagger \mathbf{S} = \mathbf{I} \iff \begin{cases} 1 = |t|^2 + |r|^2 = |t'|^2 + |r'|^2 \\ 0 = r^*t' + t^*r' = (t')^*r + (r')^*t. \end{cases}$$

$$\mathbf{S}\mathbf{S}^\dagger = \mathbf{I} \iff \begin{cases} 1 = |t'|^2 + |r|^2 = |t|^2 + |r'|^2 \\ 0 = r^*t + (t')^*r' = t*r + (r')^*t', \end{cases}$$

which follow from the fact that $\mathbf{S}$ is unitary.

- The scattering matrix is parameterized by four complex numbers, t, t', r and r', which corresponds to eight independent real-valued parameters. However, the fact that $\boldsymbol{S}$ is unitary creates certain restrictions on the parameters, namely:

$$|t| = |t'| \qquad |r| = |r'| \qquad r' = -r^* \frac{t'}{t^*}.$$

These three relationships between the parameters imply that $\mathbf{S}$ is fact described by five real-valued parameters instead of eight. With these relationships in mind, the scattering matrix may be written

$$\mathbf{S} = \begin{pmatrix} r & t' \\ t & r' \end{pmatrix} = \begin{pmatrix} r & t' \\ t & -r^* \frac{t'}{t^*} \end{pmatrix} = \begin{pmatrix} r & e^{i\phi}t^* \\ t & -e^{i\phi}r^* \end{pmatrix}, \qquad \det \boldsymbol{S} = -e^{i\phi}.$$

11.1.3   The Transfer Matrix

- In addition to the scattering matrix approach, we often analyze one-dimensional systems in terms of transfer matrices. Scattering matrices relate the waves on one side of a potential barrier to the waves on the other side in the form

$$\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \mathbf{M} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \qquad \text{where} \qquad \mathbf{M} = \begin{pmatrix} \mathrm{M}_{11} & \mathrm{M}_{12} \\ \mathrm{M}_{21} & \mathrm{M}_{22} \end{pmatrix}.$$

- To relate the scattering matrix $\mathbf{S}$ to the transfer matrix $\mathbf{M}$, we begin with

$$\begin{pmatrix} r \\ t \end{pmatrix} = \mathbf{S} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} t' \\ r' \end{pmatrix} = \mathbf{S} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The scattering matrix obeys the equations

$$\begin{pmatrix} 1 \\ r \end{pmatrix} = \mathbf{M} \begin{pmatrix} t \\ 0 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 0 \\ t' \end{pmatrix} = \mathbf{M} \begin{pmatrix} r' \\ 1 \end{pmatrix},$$

from which follows the relationship

$$\mathbf{M} = \begin{pmatrix} \frac{1}{t} & -\frac{r'}{t} \\ \frac{r}{t} & t' - \frac{rr'}{t} \end{pmatrix} \qquad \text{where} \qquad \det \mathbf{M} = \frac{t'}{t}.$$

Under the condition $t = t'$, the relationship between $\mathbf{S}$ and $\mathbf{M}$ simplifies to

$$\mathbf{M} = \begin{pmatrix} \frac{1}{t} & \frac{r^*}{t^*} \\ \frac{r}{t} & \frac{1}{t^*} \end{pmatrix} \qquad \text{and} \qquad \mathbf{S} = \frac{1}{\mathrm{M}_{11}} \begin{pmatrix} \mathrm{M}_{21} & 1 \\ 1 & -\mathrm{M}_{21}^* \end{pmatrix}.$$

- The transfer matrix approach is well-suited to systems involving a series of multiple (e.g. N) potential barriers, with endpoints at the positions $x_1, x_2, \ldots, x_{2N}$. In this case, we define the wave amplitudes A and B with respect to the shifted plane waves $e^{\pm ik(x-x_i)}$, split the x axis into $2N + 1$ regions, and calculate the $2N - 1$ transfer matrices $\mathbf{M}_{ij} = \mathbf{M}_{12}, \mathbf{M}_{23}, \ldots, \mathbf{M}_{(2N-1)2N}$ connecting the regions. In this case, the transfer matrix for the entire x axis is simply the product

$$\mathbf{M}_{1(2N)} = \mathbf{M}_{12}\mathbf{M}_{23} \cdots \mathbf{M}_{(2N-1)2N}.$$

We can then use the total transfer matrix together with the equations in the previous bullet point to calculate the corresponding total transfer matrix $\mathbf{S}_{1(2N)}$.

11.2   Scattering and Time Reversal and Parity Invariance

Discuss the one-dimensional scattering problem for a systems invariant under time reversal and parity transformation. State and derive how the scattering matrix simplifies.

- For a system invariant under time reversal, the scattering matrix is symmetric with $t = t'$ and can be written in terms of only four real parameters in the form

$$\mathbf{S} = \begin{pmatrix} r & t' \\ t & r' \end{pmatrix} \rightarrow \begin{pmatrix} r & t \\ t & -r^* \frac{t'}{t^*} \end{pmatrix} = \begin{pmatrix} r & |t|e^{i\frac{\phi}{2}} \\ |t|e^{i\frac{\phi}{2}} & -e^{i\phi}r^* \end{pmatrix}, \qquad \det \mathbf{S} = -\frac{t}{t^*}.$$

- For a system invariant under parity transformation, the scattering matrix is symmetric with respect to both diagonals, and thus obeys both $t = t'$ and $r = r'$, and can be written in terms of only three real parameters in the form

$$\mathbf{S} = \begin{pmatrix} r & t' \\ t & r' \end{pmatrix} \rightarrow \begin{pmatrix} r & t \\ t & r \end{pmatrix} = -e^{i\tau} \begin{pmatrix} i|r| & |t| \\ |t| & i|r| \end{pmatrix}, \qquad t = e^{i\tau}|t|.$$

11.2.1   Derivation: Scattering Problems with Time-Reversal Invariance

- For a time-invariant system, for any ψ solving the stationary Schrödinger equation in the regions $x < x_a$ and $x > x_b$, the time-transformed wavefunction

$$\mathcal{T}\psi(x) = \psi^*(x) = \begin{cases} A_1^* e^{-ikx} + B_1^* e^{ikx} & x < x_a \\ A_2^* e^{ikx} + B_2^* e^{-ikx} & x > x_a \end{cases}$$

also solves the stationary Schrödinger equation. Time invariance thus switches and conjugates the ingoing and outgoing vectors via

$$\mathcal{T}\Psi_{\text{in}} = \begin{pmatrix} B_1^* \\ B_2^* \end{pmatrix} = \Psi_{\text{out}}^* \qquad \text{and} \qquad \mathcal{T}\Psi_{\text{out}} = \begin{pmatrix} A_1^* \\ A_2^* \end{pmatrix} = \Psi_{\text{in}}^*.$$

- Since the time-reversed states also solve the Schrödinger equation, the scattering matrix must be symmetric. To prove this, we begin with the relationship $\Psi_{\text{out}} = \mathbf{S}\Psi_{\text{in}}$. We then replace Ψ with the equivalent solution $\mathcal{T}\Psi$ and apply the above transformation rules of Ψ_{in} and Ψ_{out} under $\mathcal{T}$ to get

$$\mathcal{T}\Psi_{\text{out}} = \mathbf{S}\mathcal{T}\Psi_{\text{in}} \implies \Psi_{\text{in}}^* = \mathbf{S}\Psi_{\text{out}}^*.$$

We then multiply the equation from the left by $\mathbf{S}^\dagger$, apply the unitary identity $\mathbf{S}^\dagger \mathbf{S} = \mathbf{I}$, and take the equation's complex conjugate to get

$$\mathbf{S}^\dagger \Psi_{\text{in}}^* = \mathbf{S}^\dagger \mathbf{S}\Psi_{\text{out}}^* = \Psi_{\text{out}} \implies (\mathbf{S}^\dagger)^* \Psi_{\text{in}} \equiv \mathbf{S}^T \Psi_{\text{in}} = \Psi_{\text{out}}.$$

Comparing this to the original relationship $\mathbf{S}\Psi_{\text{in}} = \Psi_{\text{out}}$ implies $\mathbf{S} = \mathbf{S}^T$, i.e. the scattering matrix is symmetric.

- The symmetry condition $\mathbf{S} = \mathbf{S}^T$ implies $t = t'$, which allows us to write the scattering matrix in the form

$$\mathbf{S} = \begin{pmatrix} r & t \\ t & r' \end{pmatrix} = \begin{pmatrix} r & t \\ t & -r^* \frac{t}{t^*} \end{pmatrix} = \begin{pmatrix} r & |t|e^{i\frac{\phi}{2}} \\ |t|e^{i\frac{\phi}{2}} & -e^{i\phi}r^* \end{pmatrix}, \qquad \det \mathbf{S} = -\frac{t}{t^*}.$$

11.2.2 Derivation: Scattering Problems with Parity Invariance

- A system with invariance under parity transformation obeys $V(-x) = V(x)$ and thus $x_a = -x_b$. In this case, for any ψ solving the stationary Schrödinger equation in the regions $x < x_a$ and $x > x_b$, the parity-transformed wavefunction

$$\mathcal{P}\psi(x) = \psi(-x) = \begin{cases} B_2 e^{-ikx} + A_2 e^{ikx} & x < x_a \\ B_1 e^{ikx} + A_1 e^{-ikx} & x > |x_a| \end{cases}$$

also solves the stationary Schrödinger equation. The parity operator thus transforms the roles of the coefficients in the form $A_{1,2} \longleftrightarrow B_{1,2}$. As a result, the parity operator transforms the ingoing and outgoing vectors in the form

$$\mathcal{P}\Psi_{\text{in}} = \sigma_x \Psi_{\text{in}} \qquad \text{and} \qquad \mathcal{P}\Psi_{\text{out}} = \sigma_x \Psi_{\text{out}},$$

where $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is the first Pauli spin matrix.

As an example, we consider the transformation of Ψ_{in}, which reads

$$\mathcal{P}\Psi_{\text{in}} = \mathcal{P}\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \begin{pmatrix} A_2 \\ A_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \sigma_x \Psi_{\text{in}}.$$

- The parity-transformed states are related by the scattering matrix according to

$$\sigma_x \Psi_{\text{out}} = \mathbf{S}\sigma_x \Psi_{\text{in}} \implies \sigma_x^2 \Psi_{\text{out}} \equiv \mathbf{I}\Psi_{\text{out}} = \sigma_x \mathbf{S}\sigma_x \Psi_{\text{in}}.$$

Comparing the last equality to the general relationship $\Psi_{\text{out}} = \mathbf{S}\Psi_{\text{in}}$ implies $\sigma_x \mathbf{S}\sigma_x$, which in turn implies

$$\sigma_x \mathbf{S}\sigma_x = \sigma_x \begin{pmatrix} r & t' \\ t & r' \end{pmatrix}\sigma_x = \begin{pmatrix} r' & t \\ t' & r \end{pmatrix} \mathbf{S} = \begin{pmatrix} r & t' \\ t & r' \end{pmatrix},$$

or, in other words, $r = r'$ and $t = t'$. This condition on r and t implies the scattering matrix for a parity-invariant system is symmetric with respect to both diagonals and can be written in the form

$$\mathbf{S} = \begin{pmatrix} r & t \\ t & r \end{pmatrix} = -e^{i\tau}\begin{pmatrix} i|r| & |t| \\ |t| & i|r| \end{pmatrix}, \qquad t = e^{i\tau}|t|.$$

11.3 Scattering States and Normalization of Plane Waves

Discuss scattering states and the various methods for normalizing plane waves.

- We now consider plane waves—eigenstates of the momentum operator $\hat{\boldsymbol{p}}$ and also of the stationary Schrödinger equation for a free particle. Such states obey the eigenvalue equations

$$\hat{p}|\psi_p^0\rangle = p|\psi_p^0\rangle \qquad \text{and} \qquad H_0|\psi_p^0\rangle = E_p|\psi_p^0\rangle.$$

Keep in mind the momentum eigenstates $|\psi_p^0\rangle$ don't directly represent physical states, but instead form a basis for expanding wave packets encoding the system's actual quantum state.

The superscript ^{0} is used to denote plane waves—we will use $^+$ and $^-$ superscripts later in this section to denote scattering states.

- The momentum eigenstates $\left|\psi_p^0\right\rangle$ are plane waves of the form

$$\langle x|p\rangle = \psi_p^0(x) = Ce^{i\frac{p}{\hbar}x}.$$

As discussed in [Subsection 2.8.2](#), we can normalize momentum eigenstates using the Dirac normalization convention

$$\langle x|p\rangle = \psi_p^0(x) = \frac{1}{\sqrt{2\pi\hbar}}e^{i\frac{p}{\hbar}x},$$

where the eigenstates obey $\langle p_1|p_2\rangle = \delta(p_1 - p_2)$ and the identity operator is written

$$\mathrm{I} = \int_{-\infty}^{\infty} |p\rangle \langle p| \, \mathrm{d}p.$$

This notation allows for orthogonal eigenstates in the case of a continuous spectrum, with momentum eigenvalues $p \in \mathbb{R}$.

- Alternatively, for a system with the periodic boundary conditions $\psi_p(x + L) = \psi_p(x)$, implying a discretization of the possible momentum eigenvalues p_n according to

$$p_n = \frac{2\pi\hbar}{L}, \quad n = 0, 1, 2, \ldots,$$

we can normalize plane waves using the formula

$$\langle m|n\rangle = \frac{e^{2\pi i(n-m)} - 1}{2\pi i(n - m)} L|C|^2 \quad \text{for } n \neq m.$$

For a choice of normalization constant $C = L^{-1/2}$, the eigenstates are orthonormal and obey $\langle m|n\rangle = \delta_{mn}$. Although the energy spectrum is discrete, in the limit of large L the spacing between energy levels grows arbitrarily small.

- Regardless of the choice of normalization, in the presence of the potential barrier V we define three dimensional scattering states $\left|\psi_{\boldsymbol{p}}^+\right\rangle$ according to

$$(H_0 + V)\left|\psi_{\boldsymbol{p}}^+\right\rangle = E_{\boldsymbol{p}}\left|\psi_{\boldsymbol{p}}^+\right\rangle,$$

where $H_0 = \frac{p^2}{2m}$. This definition ensures that in the limit of a vanishing potential $V(\boldsymbol{r}) \to 0$, the scattering states approach plane waves $\left|\psi_{\boldsymbol{p}}^0\right\rangle$, as expected for a free particle with $V(\boldsymbol{r}) = 0$.

- For a one-dimensional system, a scattering state is written in the form

$$\psi_p^+(x) = C \begin{cases} e^{i\frac{p}{\hbar}x} + re^{-i\frac{p}{\hbar}x} & x < x_a \\ \psi_{ab}(x) & x \in [x_a, x_b] \\ te^{i\frac{p}{\hbar}x} & x > x_b, \end{cases}$$

where r and t are parameters from the scattering matrix $\mathbf{S}$ and encode reflection and transmission from the potential barrier, respectively.

- A one dimensional system has two degenerate states, corresponding to the momentum eigenvalues $\pm|p|$. This double denegeracy motives the definition of a quantum number $\eta = \pm 1$ encoding the direction of wave propagation. The corresponding states, i.e.

$$\big||p|, \eta\big\rangle = |p\rangle, \quad \text{with eigenvalues} \quad p = \eta|p|,$$

are eigenstates of the system's Hamiltonian and can thus be described in terms of the eigenvalue E instead of p. As a result, we can write the completeness relation in two equivalent forms:

$$I = \sum_\eta \int_0^\infty |p, \eta\rangle \langle p, \eta|\, \mathrm{d}p = \sum_\eta \int_0^\infty |E, \eta\rangle \langle E, \eta|\, \mathrm{d}E,$$

where the bases $\{|p, \eta\rangle\}$ and $\{E, \eta\}$ differ only in the normalization of the basis vectors. The state $|E, \eta\rangle$ is normalized according to

$$\langle E_1, \eta_1 | E_2, \eta_2 \rangle = \delta_{\eta_1 \eta_2} \delta(E_1 - E_2).$$

- For momentum eigenvalues $p_i \sqrt{2mE_i} > 0$ we rewrite the Dirac delta function as

$$\delta(E_1 - E_2) = \delta\left(\frac{p_1^2}{2m} - \frac{p_2^2}{2m}\right) = 2m\delta\big[(p_1 - p_2)(p_1 + p_2)\big] = \frac{2m}{p_1 + p_2}\delta(p_1 - p_2)$$
$$= \frac{m}{p_1}\delta(p_1 - p_2),$$

where we have used the delta function properties $\delta(ax) = \frac{1}{|a|}\delta(x)$ and $f(x)\delta(x - a) = f(a)\delta(x - a)^3$. In terms of the above expression for the delta function, we have

$$\langle x|E, \eta\rangle = \sqrt{\frac{m}{p}} \langle x|p, \eta\rangle = \frac{1}{\sqrt{2\pi\hbar}}\left(\frac{m}{2E}\right)^{1/4} \exp\left(i\eta\sqrt{\frac{2mE}{\hbar^2}}\right),$$

which is a third way to normalize a plane wave. An analogous normalization relationship holds between the corresponding scattering scates $\psi_{E\eta}^+ = \sqrt{\frac{m}{p}}\psi_{p\eta}^+$.

11.4 Overview: The Scattering Problem in Three Dimensions

Give an overview of the generalization of the scattering problem to three dimensionsions. Discuss the important differences between the one and three-dimensional cases, and discuss the expansion of scattering states in terms of spherical waves.

11.4.1 Overview

- First, we give an overview of the scattering process:

  - An incident quantum particle, described by a wave packet, approaches a scatterer (a target or potential barrier) and scatters from the target in a given spatial direction, where we then detect the particle.

  - The incident wave packet must initially be wide enough that its width does not increase dramatically (with respect to the initial width) over the course of the scattering process. The wavepacket must be large with respect to the target and small with respect to the dimensinos of the laboratory—the second

condition ensures the wave packet does not simultaneously cover the scatterer and the detector.

The wavepacket's width is determed by the cross-sectional width (cross-sectional area) of the beam of incident particles.

– After the scattering interaction between incident particle and the target, we observe two wave packets: one continues past the target in the original direction of incidence, and one continues at an angle with respect to the direction of incidents and corresponds to the scattered particles.

– The scattering cross section is the fundamental measurable quantity used to analyze scattering processes, and represents the number of particles scattered in a given element of solid angle per unit time per unit incident current.

• We will work with particles expanded in a plane wave basis

$$\psi_k(\boldsymbol{r}) = Ce^{i\boldsymbol{k}\cdot\boldsymbol{r}},$$

where the constant C is independent of the choice of orthonormalization of the basis vectors. The probability current density for these plane waves is given by

$$\boldsymbol{j}_0 = j\frac{\boldsymbol{p}}{p}, \qquad j_0 = \frac{p}{m}|C|^2,$$

where j_0 encodes the probability of a particle being incident on the surface S_0 in the time interval Δt. In this case we can choose the normalization constant

$$|C|^2 = \frac{m}{pS_0\Delta t},$$

which represents a fourth possible way to normalize plane waves.

• A typical scattering experiment involves a beam of incident particles with cross section S_0 moving along the z axis and described by the probability current density $j_N = \frac{N}{S_0\Delta t} = Nj_0$.

11.4.2   Expansion In Terms of Spherical Waves

• We consider the scattering state $\psi_k^+(\boldsymbol{r})$, which, in the limit of a vanishing scattering potential, approaches the plane wave $\psi_k(\boldsymbol{r}) = Ce^{ikz}$. We begin our analysis by writing the plane wave in a spherical wave basis, i.e.

$$e^{ikz} = e^{ikr\cos\theta} = \sum_{l=}^{\infty}(2l+1)i^l j_l(kr)P_l(\cos\theta),$$

where $j_l(x)$ are the spherical Bessel functions and $P_l(x)$ are the Legendre polynomials.

Next, we expand the spherical wave representation in the limit of large distance of the particle from the target, in which case the the Bessel function approaches the asymptotic limit $j_l(x) \sim \frac{1}{x}\sin\left(x - l\frac{\pi}{2}\right)$. In this limit, the plane wave becomes

$$\psi_k^0 \equiv Ce^{ikz} \rightarrow C\sum_{l=0}^{\infty}(2l+1)\frac{i^l}{2ikr}\left[e^{i(kr-l\frac{\pi}{2})} - e^{-i(kr-l\frac{\pi}{2})}\right]P_l(\cos\theta).$$

- We assume the potential is spherically symmetric, in which case the problem's angular momentum is conserved, meaning we can expand the scattering state λ_k^+ in terms of the angular momentum eigenstates analogously to the expansion of the plane wave ψ_k^0. However, when expanding the scattering states in the angular momentum basis, we must also account for the scattering amplitude $S_l(k)$. With the consideration of $S_l(k)$ in mind, the expansion of the scattering states reads

$$\psi_k^+(\boldsymbol{r}) \to C \sum_{l=0}^{\infty} (2l+1) \frac{i^l}{2ikr} \left[S_l(k) e^{i(kr-l\frac{\pi}{2})} - e^{-i(kr-l\frac{\pi}{2})} \right] P_l(\cos\theta).$$

- Expansion in terms of spherical waves is needed in three dimensions because, in three dimensions, a particle can scatter in any spatial direction. Every partial wave has an associated reflection amplitude in the radial direction.

  For a given l, a scattering state has the same structure as a wave reflected from an infinite potential barrier in the previous section, i.e.

$$\psi_k^+(\boldsymbol{r}) = C \left(e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + B_1 e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \right).$$

  where the origin $r = 0$ corresponds to the infinite potential barrier.

- In elastic scattering, probability is conserved in each scattering channel and the problem's scattering matrix is thus unitary, which allows us to introduce the phase term $\delta_l(k)$ and write the scattering matrix components as

$$\mathrm{S}_l(k) = e^{2i\delta_l(k)} = 1 + 2ie^{i\delta_l(k)} \sin\delta_l(k), \qquad |S_l(k)| = 1.$$

  We then decompose the scattering states into a non-scattered plane wave, which continues in the initial direction of incidence, and a scattered spherical wave. This decomposition reads

$$\psi_{\boldsymbol{k}}^+(\boldsymbol{r}) \to Ce^{ikz} + C \left[\sum_{l=0}^{\infty} (2l+1) \frac{S_l(k)-1}{2ik} P_l(\cos\theta) \right] \frac{e^{ikr}}{r}$$

$$= C \left(e^{ikz} + \frac{f(\theta,\phi)}{r} e^{ikr} \right),$$

  where we have defined the scattering amplitude

$$f(\theta,\phi) = \frac{1}{k} \sum_{l=0}^{\infty} (2l+1) e^{i\delta_l(k)} \sin\delta_l(k) P_l(\cos\theta).$$

  For a spherically symmetric potential, the scattering amplitude depends only on the polar angle because of rotational invariance about the z axis.

- Finally, we note: the symbol $+$ in the scattering state $\psi_{\boldsymbol{k}}^+$ denotes the outgoing waves $\frac{f}{r} e^{ikr}$, while the scattering state $\psi_{\boldsymbol{k}}^-$ corresponds to the incident wave $\frac{f}{r} e^{-ikr}$.

11.5 Scattering Cross Section

What are the scattering cross section and differential cross section? Discuss their definition and physical interpretation. State and derive the optical theorem.

- Consider a beam of incident particles in a scattering experiment, where N particles in the beam pass through the beam's cross section S_0 in the time Δt, of which

$$N_\mathrm{s} = PN$$

  are scattered, where P is the probability of a given particle in the incident beam scattering off the target in any spatial direction.

- The scattering cross section, typically denoted by σ, is defined as the ratio between the number of scattered particles per unit time and the number current density of incoming particles, defined as $j_N = \frac{N}{S_0 \Delta t}$. In equation form, the definition reads

$$\sigma = \frac{1}{j_N} \frac{N_\mathrm{s}}{\Delta t} = \frac{1}{j_0} \frac{N_\mathrm{s}}{N} \frac{1}{\Delta t} = \frac{1}{j_0} \frac{P}{\Delta t},$$

  where $j_N = N j_0$. Alternatively, we can write the scattering cross section in the form

$$\sigma = \frac{N_\mathrm{s}}{N} S_0 = P S_0.$$

  Interpretation: The scattering cross section represents a surface, with surface area σ, through which a particle that has passed through the beam cross section S_0 passes through with probability $P = \frac{N_\mathrm{s}}{N} = \frac{\sigma}{S_0}$.

- The differential cross section, denoted by $\mathrm{d}\sigma$, is defined as

$$\mathrm{d}\sigma = \frac{1}{j_N} \frac{\mathrm{d}N_\mathrm{s}}{\Delta t} = \frac{1}{j_0} \frac{\mathrm{d}P}{\Delta t} = |f(\theta, \phi)|^2 \, \mathrm{d}\Omega,$$

  and is related to scattering amplitude $f(\theta, \phi)$ via

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = |f(\theta, \phi)|^2.$$

  Note that the expression is not a derivative of σ with respect to Ω—all of $\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}$ together is the differential cross section.

- The total and differential cross section are related by

$$\sigma = \iint \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \, \mathrm{d}\Omega,$$

  which should be read as an integral of the differential cross section $\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}$ over the entire solid angle $\mathrm{d}\Omega$.

- The optical theorem states that the scattering amplitude $f(\theta, \phi)$ for scattering in the forward direction is related to the total scattering cross section σ by

$$\mathrm{Im}\left[f(\theta, \phi)\big|_{\theta=0} \right] = \frac{k}{4\pi} \sigma,$$

  where k is the wave vector encoding the scattering states.

Discussion: Scattering Cross Section and the Optical Theorem

- The probability $dP(\boldsymbol{r})$ of a particle scattering into the element of solid angle $d\Omega$ in the time interval Δt is defined as

$$dP(\boldsymbol{r}) = j_{\boldsymbol{r}}\, dS \Delta t = j_0 |f(\theta, \phi)|^2\, d\Omega \Delta t,$$

where $j_{\boldsymbol{r}}$ is the probability current density in the direction $\hat{\mathbf{r}} = \boldsymbol{r}/r$ and $dS = r^2\, d\Omega$. We have related the current $j_{\boldsymbol{r}}$ to the scattering amplitude $f(\theta, \phi)$ via

$$j_{\boldsymbol{r}}(\theta, \phi) = \boldsymbol{j}(\boldsymbol{r}) \cdot \hat{\mathbf{r}} \overset{r \to \infty}{\longrightarrow} j_0 \frac{|f(\theta, \phi)|^2}{r^2}.$$

- The probability P of a particle scattering in any spatial direction in the time interval Δt is determined by the total cross section

$$\frac{1}{j_0}\frac{P}{\Delta t} = \sigma = \iint \frac{d\sigma}{d\Omega}\, d\Omega = \int_0^{2\pi}\int_0^{\pi} |f(\theta, \phi)|^2 \sin\theta\, d\theta\, d\phi$$

$$= \int_0^{2\pi}\int_0^{\pi} \left| \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)e^{i\delta_l(k)}\sin\delta_l(k)P_l(\cos\theta) \right|^2 \sin\theta\, d\theta\, d\phi$$

$$= \frac{4\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\sin^2\delta_l(k),$$

where we have evaluated the integral using the orthogonality of the Legendre polynomials, i.e.

$$\iint P_l(\cos\theta)P_{l'}(\cos\theta)\, d\Omega = \frac{4\pi}{2l+1}\delta_{ll'},$$

and the Legendre polynomial identity $P_l(1) = 1$.

- Because the total number of particles is conserved in scattering processes, the number of particles continuing past the scatterer in the original direction of incident is naturally less than the number in the incident beam, which is encoded in the unitary nature of the scattering matrix in the relationship $|S_l(k)| = 1$.

- Finally, the scattering amplitude for scattering in the forward direction is described by

$$\mathrm{Im}\left[f(\theta, \phi)\big|_{\theta=0} \right] = \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)\,\mathrm{Im}\left[e^{i\delta_l(k)} \right]\sin\delta_l(k)$$

$$= \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)\sin^2\delta_l(k)$$

$$= \frac{k}{4\pi}\sigma.$$

This result is the optical theorem, and is derived using the Legendre polynomial identity $P_l(1) = 1$.