Quantum Mechanics I

New York Lectures

Jesse Liu

Department of Physics, New York University,
726 Broadway, New York, NY 10003, USA

These lecture notes accompany the NYU Quantum Mechanics I (PHYS-UA 123) undergraduate class for Fall 2026 semester. This course presents a standard introduction to quantum mechanics aimed at physics majors. Topics include foundational experiments, probability amplitudes, Schrödinger's equation, wave mechanics, operators and observables, spin and angular momentum, and the hydrogen atom. Pre-requisites are the calculus-based undergraduate physics sequence, mathematical physics, and classical dynamics.

0 Syllabus: Quantum Mechanics I

Course: NYU PHYS-UA 123 Quantum Mechanics I
Semester: Fall 2026
Instructor: Jesse Liu (jesse.liu2 [at] nyu [dot] edu)
Lectures: Tuesday and Thursdays, 11:00am–12:15pm, Meyer 122
Office hours: Thursday, 2:00–3:00pm, 726 Broadway Room 852

Recitations:
Section 002: Broadway 1045 Mondays 2:00–3:15pm. Recitation leader: Jaewoo Lee
(Office hours: Monday & Tuesday 4–5pm Broadway 944, jl6264 [at] nyu [dot] edu)
Section 003: Broadway 1045 Tuesdays 2:00–3:15pm. Recitation leader: Jaewoo Lee
Section 004: Meyer 264 Tuesdays 4:55–6:10pm. Recitation leader: Paul John Balderston
(Office hours: Wednesdays 11am–1pm Broadway 1035, pjb9896 [at] nyu [dot] edu)

Undergraduate Program Administrator: Bill LePage (wlp1 [at] nyu [dot] edu)

0.1 Course overview

This is an upper-division undergraduate introduction to quantum mechanics. The course is primarily aimed at physics majors who have taken a calculus-based sequence of physics classes. The course covers the topics listed in the official NYU Bulletin:

> *foundational experiments; wave-particle duality; wave functions; the uncertainty principle; the time-independent Schrödinger equation and its applications to one-dimensional problems and the hydrogen atom; angular momentum; spin; Hilbert Space, operators, and observables; time independent perturbation theory; atomic spectra.*

Lecture and homework schedule

Lecture and recitation attendance is expected. In case of absences, it is the student's responsibility to catch up by reading the relevant lecture notes. Tentative lecture and exam schedule (subject to change) along with homework (HW) due dates are:

Introduction

1. Sep 3: Foundational experiments. Wave-particle duality. **Set HW1: due Sep 10**.
2. Sep 8: Atomic spectra. Radioactivity and nucleus. Bohr model.
3. Sep 10: Wavefunction, probability amplitudes. **Set HW2: due Sep 17**.

4. Sep 15: Schrödinger equation. Probability conservation. Stationary states.

   Wave mechanics

5. Sep 17: Infinite well and Dirac delta potential. **Set HW3: due Sep 24**.
6. Sep 22: Bound states in finite well.
7. Sep 24: Fourier transform. Gaussian wavepacket. **Set HW4: due Oct 1**.
8. Sep 29: Scattering and tunnelling.

   Operator formalism

9. Oct 1: State vectors. Dirac notation. Inner product. **Set HW5: due Oct 8**.
10. Oct 6: Hermitian operators and observables.
11. Oct 8: Measurement. Commutation relations. **Set HW6: due Oct 15**.
12. Oct 13: Harmonic oscillator. Ladder operators.
13. Oct 15: Uncertainty principle. Ehrenfest theorem. **HW: study for mid-term 1**.
14. Oct 20: Mid-course review.
15. Oct 22: ***Mid-term exam 1 (lectures 1–11).*** **Set HW7: due Oct 29**

    Spherical systems

16. Oct 27: Orbital angular momentum.
17. Oct 29: Quantizing angular momentum. **Set HW8: due Nov 5**.
18. Nov 3: Spherical harmonics.
19. Nov 5: Hydrogen radial equation. **Set HW9: due Nov 12**.
20. Nov 10: Stern-Gerlach experiment. Spin half.
21. Nov 12: Spin algebra. Magnetic precession. **HW: study for mid-term 2**.
22. Nov 17: Identical particles. Two spin-half particles.
23. Nov 19: ***Mid-term exam 2 (lectures 12–21)***.

    Atomic spectra

24. Nov 24: Hydrogen gross structure. **Set HW10: due Dec 3**.

    Nov 26: No class — Thanksgiving recess.

25. Dec 1: Time-independent perturbation theory.
26. Dec 3: Fine structure. Spin-orbit coupling. **HW: study for final**.
27. Dec 8: Helium. If time permits: Bell's inequality (optional).
28. Dec 10: Comprehensive review of course.

    Dec 17 (tentative): ***Final exam*****. 12pm–1:50pm following Albert schedule TBC.**

Recitation schedule

Recitation sections on Monday/Tuesdays discuss solutions to homework problems and corresponding lectures, with a tentative schedule below as guidance, but recitation leaders have discretion on content and how they are run:

Sep 7/8: No recitations (Labor Day).

1. Sep 14/15: Discuss homework 1 and lecture 1.
2. Sep 21/22: Discuss homework 2 and lectures 2–3.
3. Sep 28/29: Discuss homework 3 and lectures 4–5.
4. Oct 5/6: Discuss homework 4 and lectures 6–7.

   Oct 13/14: No recitations (Fall break; Oct 14 is Legislative Monday).
5. Oct 19/20: Discuss homework 5 and lectures 8–9.
6. Oct 26/27: Discuss homework 6 and lectures 10–11.
7. Nov 2/3: Discuss homework 7 and lectures 12–13.
8. Nov 9/10: Discuss homework 8 and lectures 16–17.
9. Nov 16/17: Discuss homework 9 and lectures 18–19.

   Nov 23/24: No recitations (Thanksgiving week).

   Nov 30/Dec 1: No recitations (no homework due week before).
10. Dec 7/8: Discuss homework 10 and lectures 20–24.

Prerequisites

Mathematical Physics (PHYS-UA 106)
Dynamics (PHYS-UA 120)
Calculus III (MATH-UA 123)

Assessment

Assessed coursework components are:

- Homework assignments (40%). These are due at 11:59pm on the Thursday noted at the assignment header. Solutions are presented in the recitations.
- Two in-class midterm exams (25% each).
- In-person final exam (30%). Date and time follows the university schedule on Albert.

Assessed component weights sum to 120%, but you will only be graded out of 100%. For example, if you complete all coursework but miss one mid-term exam, this incomplete mid-term scores zero but drops 5% rather than 25% of your final grade. This generous policy is designed to accommodate brief illness or unexpected absences during the semester. It means that out of fairness to every student, there are strictly **no make-up exams and no extra credit** for this class. All exams are closed book. Final assigned letter grades are not negotiable.

All homework assignments are made available at the start of the semester to accommodate flexibility in their completion. Assignment "set" dates reflect one week before they are due, but students requiring more time can start them well in advance. Therefore, **no extensions are granted** beyond the due date and **late homework submissions receive zero credit**. Submitted work is assessed by these guidelines at the grader's discretion:

- 4: Excellent; complete, accurate solutions; easy to follow working that is clearly and logically organized; sound explanations and physical justifications.
- 3: Good; mostly complete solutions with minor inaccuracies; organized working; some explanations and physical justifications.
- 2: Fair; partly complete solutions with some inaccuracies; working requires effort to follow; few explanations and physical justifications given.
- 1: Poor; partial attempts with inaccurate solutions; no working or explanations given.
- 0: Late or incomplete submissions.

Academic integrity

You should attempt assignments yourself first before collaborating with others in the class. You must show all your working and **hand write** your own assignments that represent ***your* original work** before turning in scans on Brightspace. Your work must be legible at the grader's discretion to receive credit. You may not share with others or upload course materials, problem sets or solutions to online forums or artificial intelligence software.

In completing homework assignments, you may use textbooks, lecture notes, scientific calculators and software with equivalent functionality, such as `Mathematica` or `Python`, to aid and check numerical calculations. However, copying directly from other students or tutors, solutions manuals, online forums, artificial intelligence software or any other source is prohibited and conflicts with academic integrity policy. Generative artificial intelligence may be appropriate in some courses, but its use is beyond the scope of this class. Just as we still learn vocabulary or arithmetic by heart and hand despite the invention of spellcheck and calculators, learning quantum physics requires thinking and problem solving by hand.

0.2 Literature

Textbooks While I provide lecture notes and there is no required textbook, I recommend picking one or two to read alongside from Bobst Library: QC174 on 9th floor or many are available online marked [NYU library] below via NYU login. Undergraduate textbooks I found useful in preparing this course include:

- James Binney and David Skinner, *The Physics of Quantum Mechanics*, Oxford University Press (2013). An exacting book from the course I took, which emphasizes operator methods. Oxford recorded Prof Binney's lectures online. [NYU library]
- Ana María Cetto and Luis de la Peña, *Quantum Mechanics: A Physical Approach*. Cambridge University Press (2025). New insightful textbook published open access.
- David Griffiths and Darrell Schroeter, *Introduction to Quantum Mechanics*, 3rd Edition, Cambridge University Press (2018). Popular textbook. [NYU library]
- David McIntyre, Corinne Manogue and Janet Tate, *Quantum Mechanics: A Paradigms Approach*. Educators encouraging "spins first" use this standard. [NYU library]
- Alastair Rae and Jim Napolitano, Quantum Mechanics, 6th Edition, CRC Press (2015). Well-regarded textbook I used as an undergraduate. [NYU library]
- Ramamurti Shankar, *Principles of Quantum Mechanics*, 2nd Edition, Springer (1996). I found the discussion on vector spaces clear as an undergraduate. [NYU library]
- David Tong, *Quantum Mechanics*, Cambridge University Press (2025). New textbook in a series based on the Cambridge theoretical physics course.
- John Townsend, *A Modern Approach to Quantum Mechanics*, 2nd edition, University Science Books (2012). Classic textbook I found illuminating as a student.

Mathematical methods This course is mathematical so have a standard review on hand:

- George Arfken, Hans Weber and Frank Harris, Mathematical Methods for Physicists, 7th Edition, Academic Press (2012). [NYU library]
- Mary Boas, Mathematical Methods in the Physical Sciences, 3rd Edition, Wiley (2006).

Online materials There is also high-quality literature online, for example:

- James Binney and David Skinner. Robust lecture notes I used as an undergraduate.
- Henning Schomerus. High quality notes in web friendly format.

- Daniel Schroeder. Insightful `Mathematica` approach with beautiful diagrams.
- Ben Simon. For advanced quantum mechanics course with experimental perspectives.
- David Tong and its advanced sequel. Theoretical yet lucid and engaging.
- LibreTexts Quantum Mechanics. Collection of open access online textbooks.

0.3 Structure and approach

Quantum mechanics is profound yet practical science. It defies intuition but the payoff is enormous: chemistry, astronomy, subatomic physics, materials, lasers, solar cells, phones, cameras, computers, transistors, light emitting diodes, medical imaging, atomic clocks, superconducting magnets. There are no surprises in the standard topics for this course, but textbooks still debate ordering, whether to teach waves, operators, spins, postulates, experiments, or history first. Mathematical methods at the level of PHYS-UA 106 is assumed.

The course has five parts. *Introduction* unabashedly underscores empiricism in the discovery of quantum mechanics, starting with the Davisson-Germer experiment both because it took place in New York and the electron is the protagonist of this class. *Wave mechanics* builds intuition by solving the time-independent Schrödinger equation in one dimension with step potentials to introduce quantization, uncertainty principle, and tunnelling. *Operator formalism* generalizes wavefunction intuition with Dirac notation into theoretical postulates then solves the harmonic oscillator via ladder operators. *Spherical systems* extends to the three-dimensional two-body problem of hydrogen before introducing spin via the Stern-Gerlach experiment and identical particles. *Atomic spectra* presents hydrogen gross structure before introducing time-independent perturbation theory up to fine structure as stepping stones for Quantum Mechanics II (PHYS-UA 124). Interpretations and Bell inequalities are not on syllabus, but optional reading serves as invitations to wider literature.

The semester has 26 lectures and two in-class mid-term exams scheduled. Each 75-minute lecture typically introduces four to six pages of new material, where this pace means reading ahead is encouraged. There is also no substitute from completing the ten homework assignments and recitation participation, which are designed to consolidate learning. Two planned lectures review course material preceding the first mid-term and final exams.

Lastly, while the physics topics are standard, the notes endeavour to extend conventional textbooks in the history of quantum physics by also highlighting contributions from often-overlooked scientists beyond Nobel laureates and local New York developments.

About these notes

These notes organize the lecture plan around the idiosyncratic NYU Fall 2026 semester schedule. Each section corresponds to one in-class session. I make the work-in-progress typeset version available to students in advance as a courtesy. I follow my notes closely (hence they're called lecture notes), which also helps if it is hard to hear or see in lectures. Being hand typed without generative artificial intelligence, these notes inevitably have artisanal rough edges so feel free to send corrections to typographical errors. The typesetting uses LaTeX, the JHEP template, and British English spelling. I draw diagrams with `figma`, `tikz.net`, `matplotlib`, and cite image sources otherwise. The notes may not be used to train artificial intelligence models. The title-page image shows an electron diffraction photo by G. P. Thomson.

Contents

I Introduction

1 Foundational experiments

Quantum mechanics did not emerge from philosophical nor mathematical deduction alone. It was empirical evidence, looking at nature, that established a new paradigm of physics.

1.1 Davisson-Germer experiment

Our story begins in the West Village, New York City, 1925. Amid the roaring twenties, quantum mechanics was starting with an explosion no less. Clinton Davisson and Lester Germer discovered the wave behaviour of electrons nearby at Bell Laboratories on 463 West Street (figure 1). It should be essential New York tourism on the centenary of the discovery that electrons are waves! Published in 1927, their groundbreaking paper begins humbly [1, 2]

> "The investigation reported in this paper was begun as the result of an accident which occurred in this laboratory in April 1925... a liquid-air bottle exploded at a time when the target was at a high temperature; the experimental tube was broken, and the target heavily oxidized by the inrushing air."

That may not sound like an auspicious start to a scientific breakthrough. But this accident was key to establishing the concept of matter waves.

Davisson and Germer were originally characterizing materials with industrial aims for heated filaments of cathode-ray tubes. After the explosion, the inrush of air oxidized their polycrystalline nickel crystal. To remove the oxide layer, they heated the nickel in hydrogen that serendipitously made the crystal monocrystalline. They scattered electrons onto nickel again and observed telltale peaks at specific scattering angles and accelerating voltages (figure 2). To understand this momentous discovery, we time jump to two backstories:

- **X-ray diffraction**. In 1912, Max von Laue discovered X-rays diffract as waves, which William and Lawrence Bragg extended to establish Bragg's law: X-rays with wavelength λ incident on crystals produce constructive interference at specific angles ϕ

$$n\lambda = d\sin\phi. \tag{1.1}$$

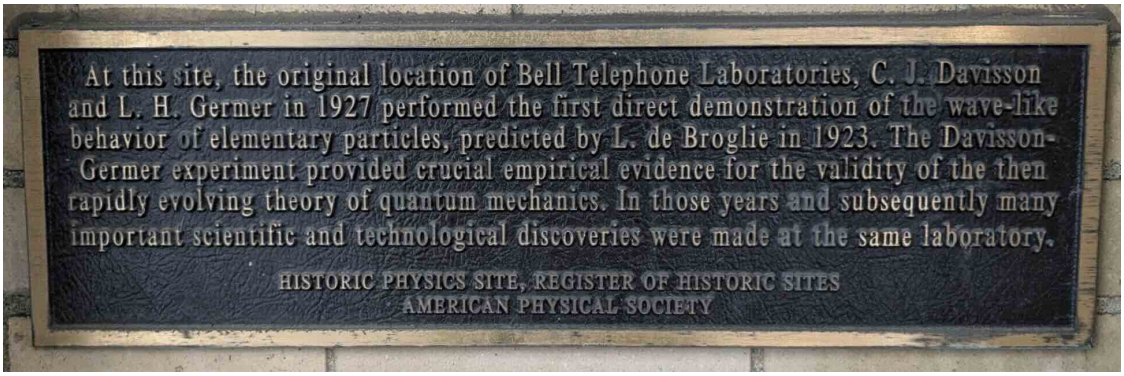


Figure 1: Plaque at the former Manhattan site of Bell Laboratories commemorating the discovery of the electron waves. Hint: find Bethune Street between 11th and 12th Street.

This relates the diffraction order n and distance between rows of atoms d like the grating spacing in wave optics. Scattering X-rays of known wavelength $\lambda = 0.16$ nm and observing the $n = 1$ peak at angle $\phi = 50^\circ$ determines

$$d = \frac{\lambda}{\sin\phi} = \frac{0.16\text{ nm}}{\sin(50^\circ)} \simeq 0.215\text{ nm}. \tag{1.2}$$

The next decades saw X-ray crystallography revolutionize biochemistry, from the structure of penicillin by Dorothy Hodgkin and Barbara Low at Oxford to deoxyribonucleic acid (DNA) by Rosalind Franklin, Raymond Gosling and Maurice Wilkins in London.

- **Matter wave hypothesis**. In his 1924 thesis, Louis de Broglie hypothesized that electrons and matter behaved as waves [3]. The *de Broglie wavelength* λ is inversely proportional to its linear momentum $p = mv$

$$\lambda = \frac{h}{p}. \tag{1.3}$$

The proportionality factor h is *Planck's constant*, which has a value of

$$h \approx 6.626 \times 10^{-34}\text{ J s} = 4.136 \times 10^{-16}\text{ eV s}. \tag{1.4}$$

Planck's constant h is a fundamental constant of nature that is determined empirically. Electron-volt units are defined by the energy gained $E = eV \approx 1.6 \times 10^{-19}$ J by an electron with charge e accelerated through 1 V of potential difference V.

We then relate the kinetic energy E_{kin} acquired during acceleration through a potential V to the electron momentum $p = mv$

$$eV = E_{\text{kin}} = \frac{1}{2}mv^2 = \frac{p^2}{2m} \quad \Rightarrow \quad p = \sqrt{2meV}. \tag{1.5}$$

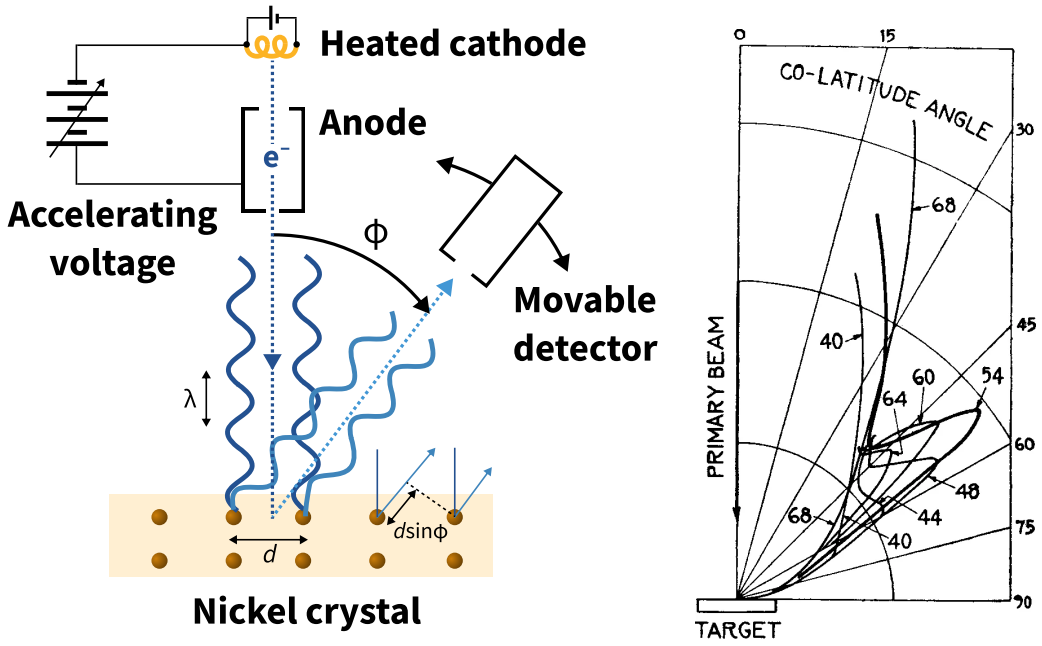


Figure 2: Left: sketch of Davisson-Germer experiment, where electron beams incident on a nickel crystal target are scattered into a movable detector. Right: measured scattered intensities labelled by acceleration voltages for various angles [1].

Then apply the de Broglie relation $p = h/\lambda$ and rearrange for λ. Accelerating electrons through a potential of 54 V, Davisson and Germer observed their diffraction results are consistent with electrons having a de Broglie wavelength of

$$\lambda = \frac{h}{\sqrt{2meV}} = \frac{1.23\text{ nm}}{\sqrt{V}} = 0.16\text{ nm}. \tag{1.6}$$

Nanometre units set the characteristic length scale of quantum mechanics.

1.2 A tale of two Thomsons

Independent of Davisson and Germer, George Paget Thomson and his student Alexander Reid discovered electron wave properties at the University of Aberdeen in Scotland. Published in 1927, they scattered electrons through celluloid film and observed a [6]

> "central spot formed by the undeflected rays is surrounded by rings, recalling in appearance the haloes formed by mist round the sun".

Figure 3 shows a remarkable photograph of these electron diffraction rings. In polycrystalline film, electrons are scattered in all directions equally. The rings have radii that obey a

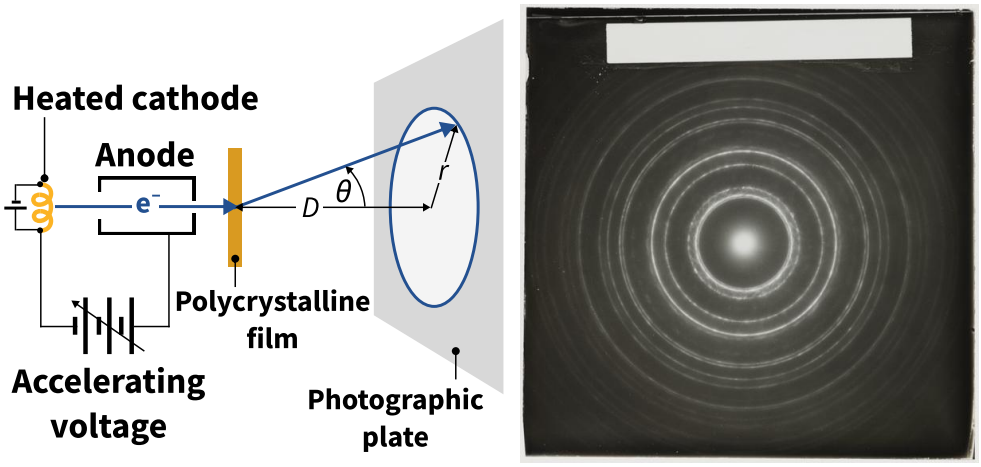


Figure 3: Sketch of Thomson-Reid electron diffraction experiment conducted at the University of Aberdeen, Scotland. An electron beam is incident on a thin piece of polycrystalline film. Small crystals are oriented in random directions, so the diffraction pattern comprises concentric rings shown in the photograph by George Paget Thomson. CC-BY 4.0.

geometric equation from wave optics $r = D\tan\theta$. These electron diffraction results require a new concept: **wave-particle duality**. This juxtaposition is a tale of two Thomsons:

> J. J. Thomson discovered electrons are particles in 1897.
> G. P. Thomson, his son, discovered electrons are waves in 1927.

Let us review the evidence for the particle nature of matter. In 1869, William Crookes found that electrical current flows from a heated cathode to anode in a vacuum chamber. These Crookes tubes emanated mysterious rays that made glass and phosphorescent materials fluoresce, which revealed their straight-line paths and are blocked by metal obstacles to cast shadows (figure 4). At Cambridge, J. J. Thomson studied cathode rays with charge q and mass m by balancing the forces $qE = qvB$ from a magnetic B and electric field E. The magnetic force qvB induces a centripetal force mv^2/r that causes a circular path with radius r, $mv^2/r = qvB$. Eliminating v relates the charge-to-mass ratio q/m to the E and B fields

$$\frac{q}{m} = \frac{1}{r}\frac{E}{B^2} \approx -1.7 \times 10^{11}\ \mathrm{C\,kg^{-1}}. \tag{1.7}$$

This is three orders of magnitude larger than that of a hydrogen ion determined via electrolysis. This q/m ratio implies either the charge is 1000 times larger or the mass is 1000 smaller than hydrogen ions. Disambiguation requires directly measuring the electron charge. Thomson called the electron "corpuscles" and caused water vapour to condense into droplets (figure 4 upper right). Balancing the droplet gravitational force mg against an electric field

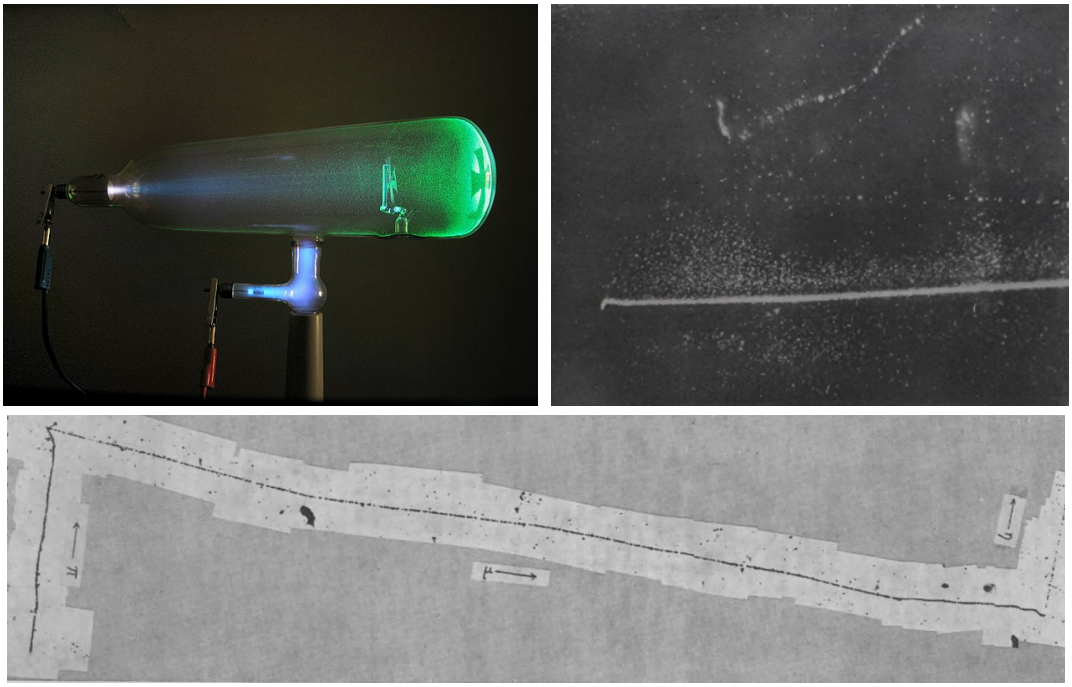


Figure 4: Experimental evidence for particle nature of matter constituents. Upper-left: cathode rays travelling like particles in straight lines in a modern Crookes tube. Upper-right: Wilson cloud chamber showing tracks of an helium nucleus (thick line) and electrons (upper lines) [4]. Bottom: photo of pion (π) decaying into a muon (μ) then an electron (e labelled η in the image), observed by Rosemary Brown and Minnie van der Merwe [5].

qE fixes the charge to be around $q = mg/E \approx 10^{-19}$ C, the same as hydrogen ions [7]. This was the discovery of the electron as a subatomic particle.

Throughout the twentieth century, scientists discovered new forms of subatomic matter. For example, pions π are particles responsible for the strong force binding protons and neutrons. Marietta Kurz and Irene Roberts first identified pions in photographic plates that underpinned its discovery in 1947 [8]. Figure 4 (lower) shows a photo of a pion particle turning into a muon and electron, uncovered by Rosemary Brown and Minnie van der Merwe in 1949 with Cecil Powell at Bristol University. These photos show the particle behaviour of matter, moving like points and leaving lines that show their trajectories.

1.3 Photoelectric effect

In 1801, Thomas Young challenged the "corpuscular" theory of light favoured by Isaac Newton with a wave theory that explains interference [9]. Interference is exemplified by Newton's

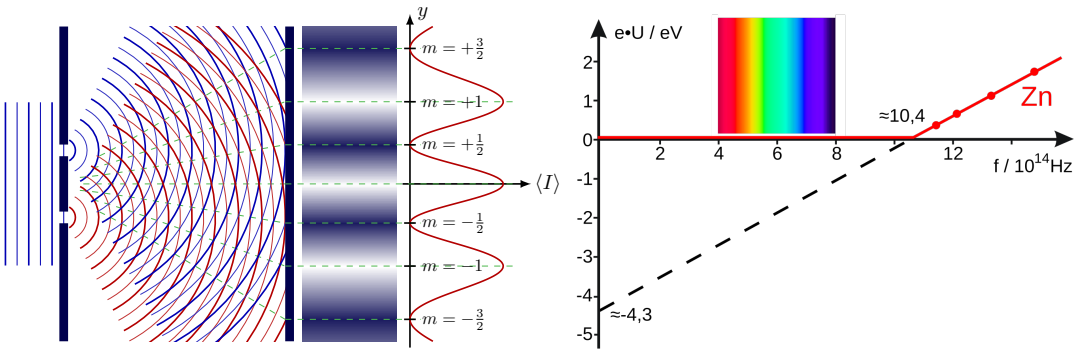


Figure 5: Wave-particle duality of light. Left: schematic of wave interference in double slit experiment. Right: photoelectric effect showing energy of emitted electrons from zinc versus frequency of incident light. Images: Izaak Neutelings, Klaus-Dieter Keller, CC BY 3.0.

rings observed by Robert Hooke in 1665 and the double-slit experiment (figure 5 left). But if light is a wave, what exactly is oscillating? The answer came in 1865 and is among the most stunning results of physics. Spoiler alert, Maxwell's equations in vacuum[1] imply the electric $\mathbf{E}$ and magnetic $\mathbf{B}$ fields obey the wave equation (see Electromagnetism PHYS-UA 131)

$$\frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} = \nabla^2 \mathbf{E}, \quad \frac{1}{c^2}\frac{\partial^2 \mathbf{B}}{\partial t^2} = \nabla^2 \mathbf{B}, \tag{1.8}$$

where the Laplacian is the sum of double derivatives $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$. Writing the wave as travelling in the z direction, the $\mathbf{E}, \mathbf{B}$ oscillations are mutually orthogonal

$$E_x = E_0 \cos(kz - \omega t), \quad B_y = B_0 \cos(kz - \omega t). \tag{1.9}$$

We relate $c = \omega/k$ to wavenumber $k = 2\pi/\lambda$ and angular frequency $\omega = 2\pi f$. Maxwell found the waves propagate at the speed of light $c = 1/\sqrt{\varepsilon_0 \mu_0} = 299\,792\,458$ m s^{-1} set by the vacuum electric permittivity ε_0 and magnetic permeability μ_0. The unification of electricity, magnetism and optics seemed to settle, once and for all, that light is a wave. However…

In 1887, Heinrich Hertz observed that ultraviolet light incident on certain materials induces an electrical current that increased spark discharge rates. If light were a wave, the induced current would be: (i) proportional to the intensity (amplitude square) of incident light, and (ii) independent of its frequency. The photoelectric effect contradicted this expectation. Wilhelm Hallwachs and Philipp von Lenard subsequently measured electron stopping voltages with light incident on metals (figure 5, right) and established:

[1] These are Gauss' law $\nabla \cdot \mathbf{E} = 0$, no magnetic monopoles $\nabla \cdot \mathbf{B} = 0$, Faraday's law $\nabla \times \mathbf{E} = -\partial \mathbf{B}/\partial t$, and the Ampère–Maxwell law $\nabla \times \mathbf{B} = \mu_0 \varepsilon_0 \partial \mathbf{E}/\partial t$. Taking the curl $\nabla\times$ of Maxwell's equations and applying the vector identity $\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}$ yields the wave equation.

- no electrons are emitted if the incident light is below a certain frequency;
- the energy of each emitted electron increases with light frequency not its intensity.

In 1905, Albert Einstein proposed [10] the electron kinetic energy E_e equals the photon energy plus a material-specific work function W: $E_e = hf - W$. The Planck-Einstein relation relates the photon energy E to its frequency $f = c/\lambda$ (wavelength λ for speed c):

$$E = hf = \frac{hc}{\lambda} \simeq \frac{1240 \text{ eV nm}}{\lambda}. \tag{1.10}$$

The combination $hc \simeq 1240$ eV nm is an extremely useful conversion factor. It is convenient to absorb factors of 2π by defining the reduced Planck's constant

$$\hbar = \frac{h}{2\pi} \simeq 1.054 \times 10^{-34} \text{ J s}, \quad \hbar c \simeq 197 \text{ eV nm}. \tag{1.11}$$

We relate energy to angular frequency $\omega = 2\pi f$ and momentum to wavenumber $k = 2\pi/\lambda$:

$$E = \hbar\omega, \quad p = \hbar k. \tag{1.12}$$

1.4 Compton scattering

Compton scattering provides direct evidence for the particle nature of light. In 1923, Arthur Compton fired X-rays [11, 12] with wavelength λ onto effectively free electrons in a carbon target and measured the wavelength change versus scattering angle θ (figure 6). Classical electromagnetic fields cause free electrons to oscillate via the Lorentz force $\mathbf{f} = -e(\mathbf{E} + \mathbf{v} \times \mathbf{B})$. The oscillating electron then emits radiation with the same wavelength λ'

$$\lambda' - \lambda = 0, \quad \text{classical electromagnetism}. \tag{1.13}$$

The electron does not recoil in this elastic interaction called Thomson scattering[2]. At $\theta = 0$, there is indeed only one peak at the incident X-ray wavelength. But at larger $\theta > 0$, Compton observed a second peak that grows with angle, in tension with classical predictions.

If we instead model light as a particle, it can transfer energy and momentum onto an electron, causing it to recoil. To see this, consider a photon with energy E_γ and three-momentum $\mathbf{p}_\gamma$. This collides with a stationary electron of energy E_e and momentum $\mathbf{p}_e$. Conserving energy and momentum before (unprimed) and after (primed) the collision gives

$$E_\gamma + E_e = E'_\gamma + E'_e, \tag{1.14}$$

$$\mathbf{p}_\gamma + \mathbf{p}_e = \mathbf{p}'_\gamma + \mathbf{p}'_e. \tag{1.15}$$

[2] In classical electromagnetism, the Thomson cross-section is $\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{2} r_e^2 (1 + \cos^2\theta)$, where the classical electron radius $r_e = e^2/4\pi\varepsilon_0 m_e c^2$ is the distance where electrostatic energy equals the electron rest mass.

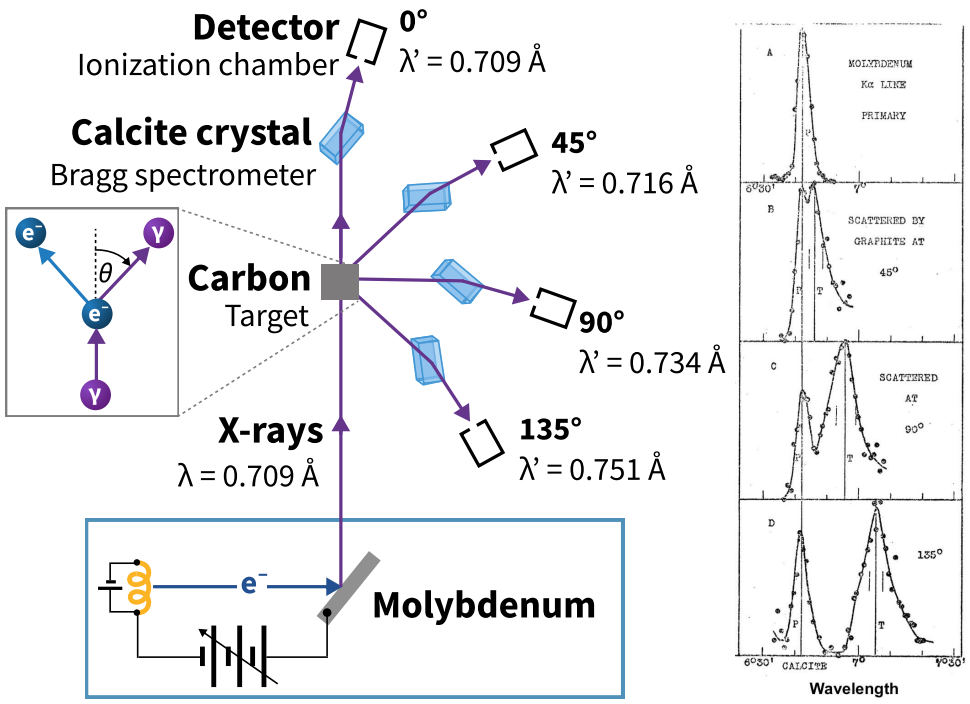


Figure 6: Compton scattering: a cathode-ray tube generates the "K-α" line of 0.709 Å X-rays from molybdenum, which strike a carbon graphite target and scatter into a calcite crystal spectrometer. Dots show the measured data at various scattering angles, where the classical Thomson (quantum Compton) prediction is marked P (T) [12].

Next is a little special relativity. Working in the initial electron's rest frame, we have $E_e = m_e c^2$ and $\mathbf{p}_e = \mathbf{0}$. Making the primed electron $E'_e, \mathbf{p}'_e$ the subject then squaring yields

$$(E'_e)^2 = (E_\gamma - E'_\gamma + m_e c^2)^2, \tag{1.16}$$

$$(\mathbf{p}'_e)^2 = (\mathbf{p}_\gamma - \mathbf{p}'_\gamma)^2 = \mathbf{p}_\gamma^2 + \mathbf{p}_\gamma'^2 - 2|\mathbf{p}_\gamma||\mathbf{p}'_\gamma|\cos\theta_{\gamma\gamma'}, \tag{1.17}$$

Subtract these equations and apply the Einstein energy-momentum relation $(E'_e)^2 - (\mathbf{p}'_e c)^2 = (m_e c^2)^2$ then relate the photon momentum to its energy $|\mathbf{p}_\gamma| = E_\gamma / c$. Expand the brackets so all the squared terms $(m_e c^2)^2$, E_γ^2, $(E'_\gamma)^2$ cancel. Dividing through by $E_\gamma E'_\gamma$ gives

$$m_e c^2 \left(\frac{1}{E'_\gamma} - \frac{1}{E_\gamma} \right) = 1 - \cos\theta_{\gamma\gamma'}. \tag{1.18}$$

The final step is to invoke the particle nature of light. Applying the Planck-Einstein relation

(1.10), $E = hf = hc/\lambda$, gives the Compton scattering formula supported by data (figure 6)

$$\lambda' - \lambda = \frac{h}{m_e c}\left(1 - \cos\theta_{\gamma\gamma'}\right). \tag{1.19}$$

We define the electron *Compton wavelength* λ_C corresponding to its rest-mass energy $m_e c^2$, which is very small hence high-energy X-rays are needed

$$\lambda_C = \frac{hc}{m_e c^2} \simeq \frac{1240 \text{ eV nm}}{m_e c^2} \simeq 2.43 \times 10^{-12} \text{ m}. \tag{1.20}$$

Ångström units 1 Å = 10^{-10} m are often used in spectroscopy. These experiments establish wave-particle duality, where matter and light exhibit both particle and wave behaviour.

HW1: Units and wave-particle duality (set Sep 3, due Sep 10)

1. Planck's constant $h = 6.626 \times 10^{-34}$ J s, speed of light $c = 2.9979 \times 10^{8}$ m s^{-1}, and 1 eV $= 1.602 \times 10^{-19}$ J. Find the numerical value of hc and $\hbar c$ in (a) Joule-metres J m and (b) show that $hc = 1240$ eV nm in atomic units of electronvolt-nanometers eV nm.

2. Sketch a diagram for and describe the setup of the Davisson-Germer experiment.

   (a) Optional field trip: visit the site where this experiment occurred and find figure 1.

   (b) How does this experiment show the wave behaviour of electrons? What is meant by wave-particle duality? What is the de Broglie hypothesis and its formula?

   (c) Derive the expression for the momentum p of an electron with mass m and charge Q accelerated through a potential V. Show that the de Broglie wavelength of an electron accelerated through a voltage V is $\lambda \simeq \frac{1.23 \text{ nm}}{\sqrt{V}}$.

   (d) What is the numerical value of the voltage required such that the electron de Broglie wavelength is 100 times smaller than (i) a virus with size 20 nm and (ii) the mean diameter of a silicon atom of 0.2 nm?

   (e) Describe the quantities in the Bragg formula $d \sin\phi = n\lambda$. Dialling the incident electron beam energy of their experiment, Davisson and Germer also found an $n = 1$ peak at $\phi = 44^\circ$ for monocrystalline nickel with lattice spacing $d = 0.215$ nm. What is the de Broglie wavelength and energy $E = p^2/2m$ of the incident electron beam? Express your answers in units of eV and nm.

3. Using $E = mc^2$, show that the electron mass $m_e = 9.11 \times 10^{-31}$ kg is 511 keV/c^2.

   (a) Convert the following rest masses into eV/c^2 units:
   (i) proton $m_p = 1.67 \times 10^{-27}$ kg, (ii) onion cell 10^{-15} kg, (iii) cat $m_{\text{cat}} = 5$ kg.

   (b) Show that de Broglie wavelength is $\lambda_p = (hc)/(0.01mc^2)$ for a particle with mass m and speed $v = 0.01c$. For the electron and objects (i)–(iii), estimate numerical values for the Compton wavelength λ_C and λ_p for $v = 0.01c$.

4. Describe the Compton experiment and write down the Compton scattering formula.

   (a) Taking the incident wavelength as $\lambda = 0.01$ nm, what are the outgoing scattered wavelengths λ'_θ for scattering angles $\theta = 0^\circ, 90^\circ, 180^\circ$ in the cases of: (i) classical electromagnetism and (ii) treating light as a particle?

   (b) How does this experiment and its results demonstrate the particle nature of light?

2 Atomic structure and radiation

The development of quantum mechanics is intertwined with experimentally probing atomic structure and its radiation. Matter emits thermal radiation, electromagnetic line spectra, and nuclear radiation. Understanding all of these phenomena requires quantum mechanics.

2.1 Line spectra

In the nineteenth century, advances in diffraction grating manufacturing enabled precision spectroscopy of light. Emission and absorption at discrete wavelengths were observed in terrestrial gases while Anders Ångström carefully measured hydrogen spectra from the Sun's light (figure 7). In 1885, Johann Balmer devised an empirical formula that fit the data

$$\lambda_n = \lambda_B \frac{n^2}{n^2 - 4}. \tag{2.1}$$

This relates the observed wavelengths λ_n to the Balmer constant $\lambda_B = 364.56$ nm by squares of integers n. It allowed Balmer to predict the $n = 7$ line that Ångström observed $\lambda_7 \approx 397$ nm. Three years later, Johannes Rydberg generalized this formula to

$$\frac{1}{\lambda} = R_H \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right). \tag{2.2}$$

This relates the light wavelength λ, the Rydberg constant for hydrogen $R_H \approx 13.6\ \text{eV}/hc \simeq 1.09 \times 10^7\ \text{m}^{-1}$ with a difference of two integers n_1, n_2. The R_H is related to the Rydberg constant in the limit of heavy nucleus R_∞ by the reduced mass

$$R_\infty = \frac{m_e + m_p}{m_p} R_H = \frac{m_e e^4}{8\varepsilon_0^2 h^3 c} = 10\,973\,731.568\,157\,(12)\ \text{m}^{-1}, \tag{2.3}$$

which is among the most precisely measured quantities in nature. The Balmer formula corresponds to the case of $n_1 = 2$. The origins of this empirical formula remained unexplained for over two decades. Classical theories of radiation predict continuous, not discrete, spectra. These mysteries required further knowledge about the structure of matter. Stellar spectroscopy is foundational to astronomy, where Cecilia Payne published in 1925 that hydrogen and helium are the dominant elements in stars and thus the cosmos [13].

2.2 Radioactivity and nuclei

Hints for atomic substructure came from the surprising discovery of radioactivity. In 1897, Henri Becquerel discovered radiation emitted by uranium salts after leaving them in his desk

Figure 7: Balmer series of visible light from line emission of hydrogen atoms. The four lines visible to the human eye have wavelength (colour, energy): 410.2 nm (indigo, 3.02 eV); 434.0 nm (blue, 2.86 eV); 486.1 nm (cyan, 2.55 eV); 656.2 nm (red, 1.89 eV). There are further lines in the ultraviolet. Image: Jan Homann CC BY-SA 3.0.

drawer on a cloudy Paris day and found photographic plates were unexpectedly exposed. Soon after in 1898, Marie Skłodowska-Curie and Pierre Curie discovered new radioactive elements radium and polonium. In 1901, Harriet Brooks and Ernest Rutherford discovered that the gas emitted by radium remained radioactive and is now known as radon.

The Curies carefully measured the radioactivity rate depends on the quantity of atoms

$$N(t) = N_0 \mathrm{e}^{-\lambda t}. \tag{2.4}$$

The emission of radiation is stochastic. It is impossible to predict exactly when or which atom will decay, only the aggregate average that follow the exponential decay of follows Poisson statistics. An electroscope detects clicks of ionization that are randomly distributed during a radioactive decay. The decay is fundamentally probabilistic and cannot be explained classically. The radiation was a million times more energetic than light emitted from atoms

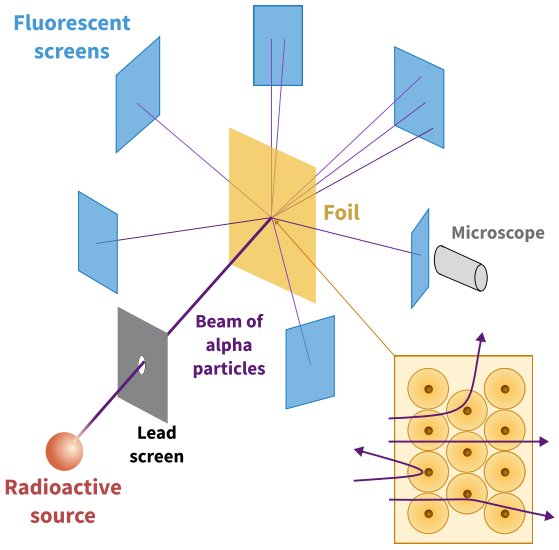
$$E_{\text{radioactivity}} \sim \mathscr{O}(10^6)\ \text{eV}, \quad E_{\text{line emission}} \sim \mathscr{O}(1)\ \text{eV}. \tag{2.5}$$

This indirectly evidenced extra nuclear forces beyond electromagnetism deep inside atoms.

Hans Geiger invented the ionisation chamber in 1908 to measure radioactivity, but the device behaved erratically from traversing alpha particles. At Manchester, Rutherford decided to get to bottom of this misbehaving instrument in 1911 as a new student Ernest Marsden arrived. They fired alpha particles from a radium source at gold foil and measured their scattering angles ϕ (figure 8a) by counting fluorescent hits on a zinc sulphide screen in a darkened laboratory. One day, Rutherford suggested they see how many reflected back on itself and were surprised to see thousands reflect back.

Rutherford proposed electrons surrounding a positively charged nucleus [15]. A key result from classical mechanics textbooks assuming Coulomb's law with energy and momentum conservation gives the elastic differential cross-section

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left(\frac{zZe^2}{4\pi\varepsilon_0}\frac{1}{4E}\right)^2 \frac{1}{\sin^4(\phi/2)}. \tag{2.6}$$



(a) Schematic of scattering experiment

610 Dr. H. Geiger *and* Mr. E. Marsden *on the Laws of*

TABLE II.

Variation of Scattering with Angle. (Collected results.)

I. Angle of deflexion, ϕ.	II. $\frac{1}{\sin^4\phi/2}$	III. Silver. Number of scintillations, N.	IV. Silver. $\frac{N}{\sin^4\phi/2}$	V. Gold. Number of scintillations, N.	VI. Gold. $\frac{N}{\sin^4\phi/2}$
150°	1·15	22·2	19·3	33·1	28·8
135	1·38	27·4	19·8	43·0	31·2
120	1·79	33·0	18·4	51·9	29·0
105	2·53	47·3	18·7	69·5	27·5
75	7·25	136	18·8	211	29·1
60	16·0	320	20·0	477	29·8
45	46·6	989	21·2	1435	30·8
37·5	93·7	1760	18·8	3300	35·3
30	223	5260	23·6	7800	35·0
22·5	690	20300	29·4	27300	39·6
15	3445	105400	30·6	132000	38·4
30	223	5·3	0·024	3·1	0·014
22·5	690	16·6	0·024	8·4	0·012
15	3445	93·0	0·027	48·2	0·014
10	17330	508	0·029	200	0·0115
7·5	54650	1710	0·031	607	0·011
5	276300	...	...	3320	0·012

(b) Deflection results [14]

Figure 8: Setup of Geiger–Marsden–Rutherford experiment, showing alpha particles emitted from a radioactive source scattering off gold foil (left) and deflection results (right).

Here, ze (Ze) is the incident (target nucleus) electric charge, $E = mv^2/2$ is the incident alpha-particle energy, ϕ is the scattering angle and $\mathrm{d}\Omega$ is the solid angle. Figure 8b shows the results from their 1913 paper, where the meticulously collected data of scintillation hits $N(\phi)$ support the predicted $\sin^{-4}(\phi/2)$ dependence from Rutherford's calculations. This established the existence of the atomic nucleus and the discovery of the proton.

We complete the story of the nucleus with the neutron. Its discovery took another two decades. At Paris in 1932, Irène Curie (M. Curie's daughter) and Frédéric Joliot studied radioactivity by bombarding beryllium with alpha particles. They discovered powerful radiation that could eject protons from hydrogen-rich paraffin wax, the telltale signature of a heavy neutral particle. James Chadwick reproduced the experiment in England and correctly interpreted the radiation is the neutral counterpart to the proton, discovering the neutron.

2.3 Franck-Hertz experiment

In 1914, James Franck and Gustav Hertz reported their experiment [16] accelerating electrons in a vapour of mercury gas (figure 9 left). Gradually increasing the voltage, they observed sudden sharp drops in arrival electrons in discrete intervals of 4.9 V. The electron loses its kinetic energy of 4.9 eV of energy to transition the mercury atom into its excited state. Mercury

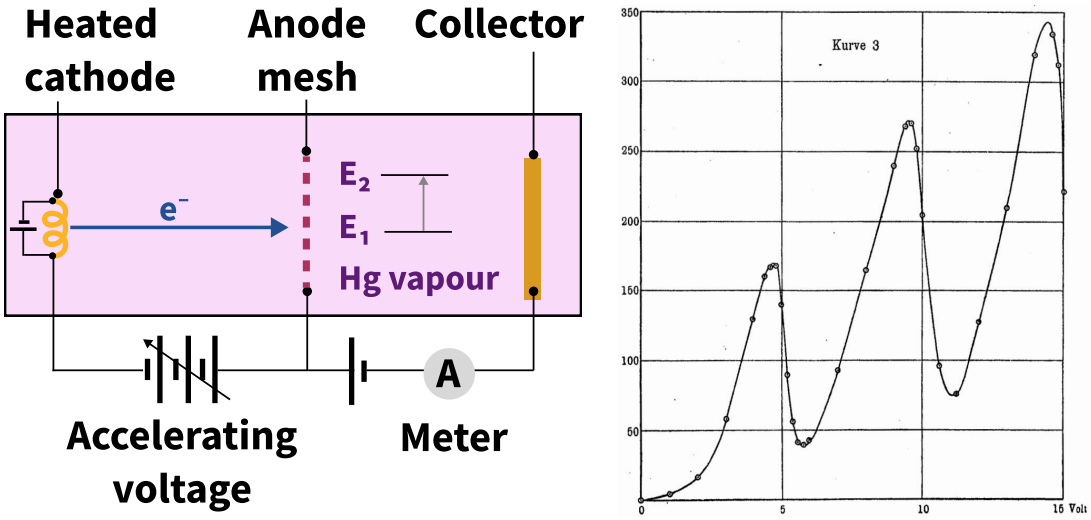


Figure 9: Left: Sketch of Frank-Hertz experimental setup. Right: Data from scattering electrons onto mercury ions with regular dips in current (vertical axis) at 4.9 V intervals in the wire-mesh potential (horizontal axis) [16].

atoms can only accept discrete integer n units of energy to enter excited states:

$$E^e_{\text{incident}} \neq n(\Delta E_{\text{Hg}} = 4.9\text{ eV}) \quad \Rightarrow \quad \text{electrons travel freely to detector,} \tag{2.7}$$

$$E^e_{\text{incident}} \approx n(\Delta E_{\text{Hg}} = 4.9\text{ eV}) \quad \Rightarrow \quad \text{electrons stopped so current drops.} \tag{2.8}$$

This discovery was direct evidence for quantized energies of electrons in atoms.

Motivated by nuclei and Frank-Hertz discoveries, Niels Bohr proposed in 1913 a planetary model for atoms (figure 10). The physics of the Bohr model is not quite correct, but it explains the Rydberg formula and is pedagogically useful as dimensional analysis [17]. Applying Newtonian mechanics to an electron orbiting a proton, the attractive electrostatic force provides the electron's centripetal force $F_{\text{Coulomb}} = F_{\text{centripetal}}$:

$$-\frac{e^2}{4\pi\varepsilon_0 r^2} = -\frac{m_e v^2}{r}. \tag{2.9}$$

The classical angular momentum is $L = m_e v r$, so the orbital radius follows

$$r = \frac{4\pi\varepsilon_0 L^2}{e^2 m_e}. \tag{2.10}$$

The total electron energy is the sum of the kinetic and Coulomb potential energies

$$E = \frac{1}{2}m_e v^2 - \frac{e^2}{4\pi\varepsilon_0 r} = -\frac{e^4 m_e}{32\pi^2\varepsilon_0^2 L^2}. \tag{2.11}$$

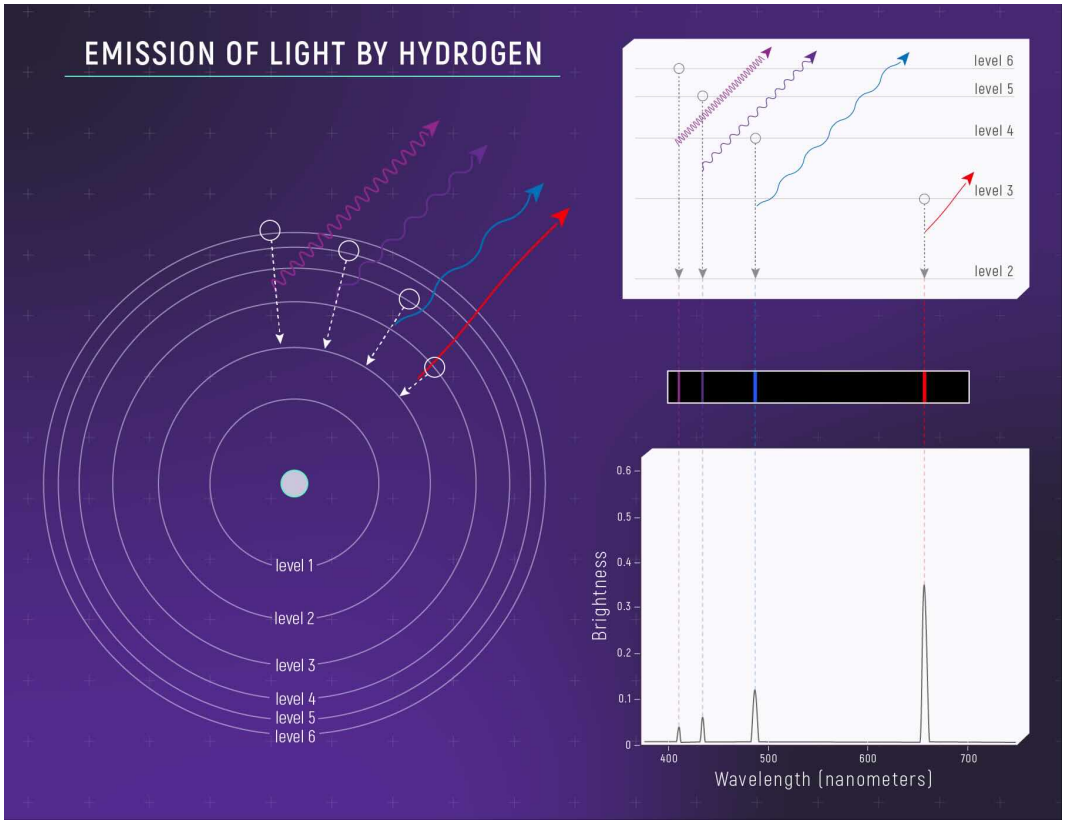


Figure 10: Bohr model of atom to explain Balmer series of visible light from line emission of hydrogen atom. Image: NASA, ESA, CSA, Leah Hustak (STScI).

Next is the quantum guess. Motivated by ideas of Planck-Einstein quantization from the photoelectric effect, Bohr hypothesized that the electron angular momentum is quantized

$$L \stackrel{?}{=} n\hbar, \quad n = 1, 2, 3, \ldots \tag{2.12}$$

This is *not* the correct quantization of angular momentum (sections 16 and 19 correctly derive the quantization of angular momentum and energy using quantum mechanics). But this was 1913 and quantum mechanics was only successfully formulated in 1926, so let's give them a break with this semi-classical guess. Using $L \stackrel{?}{=} n\hbar$ in equation (2.11), the energy is quantized

$$E_n = -\frac{m_e}{2}\left(\frac{e^2}{4\pi\varepsilon_0\hbar}\right)^2 \frac{1}{n^2} = -\frac{E_R}{n^2}. \tag{2.13}$$

Remarkably, this *is* the correct expression for hydrogen energies, where $E_R \simeq 13.6$ eV is the Rydberg energy. It provides a physical basis for the empirical Rydberg formula, where

photons have discrete energies due to electrons transitioning between quantized energy levels

$$E_{nm}^{\text{photon}} = hf = E_n - E_m = E_R\left(\frac{1}{n^2} - \frac{1}{m^2}\right). \tag{2.14}$$

The energy-level transitions are fundamentally probabilistic. The minimum orbit for $n = 1$ is called the Bohr radius

$$r_B = \frac{4\pi\varepsilon_0\hbar^2}{e^2 m_e}. \tag{2.15}$$

This story could not be correct because classical orbiting electrons lose kinetic energy by emitting electromagnetic radiation. Quoting the Larmor formula from classical electromagnetism, it states the emitted power P proportional to the acceleration $a = v^2/r$ squared

$$P = \frac{2}{3}\frac{e^2}{4\pi\varepsilon_0 c^3}a^2. \tag{2.16}$$

This implies Bohr model atoms are unstable as electrons lose kinetic energy and spiral into the nucleus. This deficiency motivated de Broglie to hypothesize that electrons are waves and the circular orbits are actually electron standing waves. We are familiar with standing waves from guitar strings. Vibrating strings fixed at both ends can only oscillate in discrete normal modes. The dimensionless combination of quantities

$$\alpha = \frac{e^2}{4\pi\varepsilon_0\hbar c} \simeq \frac{1}{137} \tag{2.17}$$

is called the fine-structure constant and is very useful to simplify calculations, which we revisit in section 19.4.

2.4 Thermal radiation

Textbooks often start with Planck's photon hypothesis to motivate $E = hf$, but precise statements require Thermal and Statistical Physics (PHYS-UA 140) beyond course pre-requisites. For completeness, let us briefly sketch the arguments. Matter at finite temperature emits electromagnetic thermal radiation. In an ideal box, the number of vibrating modes of waves between f and $f + \mathrm{d}f$ per unit volume[3] is proportional to f^2

$$n(f)\mathrm{d}f = Af^2\mathrm{d}f. \tag{2.18}$$

[3]Standing waves in a box of length L carry wavenumber $k_{n_x} = \pi n_x/L$ for integer n_x. Extending to three dimensions, the number of modes in a spherical shell is $4\pi r^2\mathrm{d}r$, where $r^2 = n_x^2 + n_y^2 + n_z^2$. Change variables to frequency $f = ck/2\pi$ gives $r = 2Lf/c$, so the number of modes is $n(f)\mathrm{d}f = 4\pi(2L/c)^3 f^2\mathrm{d}f$.

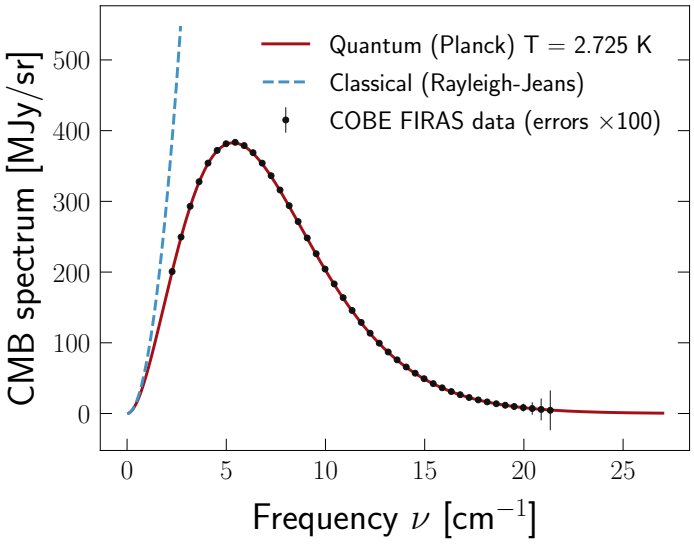


Figure 11: Cosmic microwave background (CMB) data versus frequency from Ref. [18] compared with quantum Planck (red solid line) vs classical Rayleigh-Jeans (blue dashed line) models of thermal radiation for $T = 2.725$ K plotted in `matplotlib`.

The numerical constant $A = 8\pi/c^3$ was derived by John William Strutt (Lord Rayleigh) and James Jeans in 1905. The energy density u is proportional to this mode density $n(f)$, which predicts the Rayleigh-Jeans law of radiation emitted from a blackbody

$$u(f,T) = Af^2\langle E\rangle = Af^2 k_B T, \tag{2.19}$$

for temperature T and Boltzmann constant k_B. To find the total energy for a given temperature, we must integrate over frequencies. A useful heuristic is sound waves, whose energy is related to the material heat energy. The smallest physical sound wavelength is bound by the material inter-atomic distance. We restrict the integration to a maximum frequency (minimum wavelength) $f_{\max} = c/\lambda_{\min}$ corresponding to the inter-atomic spacing $\lambda_{\min} \sim 0.1$ nm:

$$\int_0^{f_{\max}} f^2 \mathrm{d}f \sim f_{\max}^3 < \infty. \tag{2.20}$$

For classical electromagnetic waves in vacuum, the analogue of this minimum inter-atomic distance (maximum frequency) is unclear. Therefore, the integral diverges $u \to \infty$ for high frequencies $f \to \infty$, in contradiction with observations. This divergent behaviour is sometimes melodramatically called the *ultraviolet catastrophe*.

In 1900, Max Planck hypothesized that light is quantized with energy proportional to the frequency $E = hf$. Quoting results from statistical mechanics, the probability of a particle

i being in energy E_i is proportional to an exponentially weighted function of the system temperature $k_B T$

$$p(E_i) = \frac{1}{N} \mathrm{e}^{-E_i/k_B T}, \quad \text{where } N = \sum_i \mathrm{e}^{-E_i/k_B T}, \tag{2.21}$$

which is called the Boltzmann distribution. For those who have not taken statistical mechanics, just trust me on this for now. The mean energy $\langle E \rangle$ for the whole system is given by summing over all energies E_i weighted by the probability $p(E_i)$ of being in that energy

$$\langle E \rangle = \sum_i E_i \, p(E_i). \tag{2.22}$$

The Planck photon hypothesis states that each oscillating mode n of electromagnetic radiation is related to the energy by $E = nhf$ to give

$$\langle E(f,T) \rangle = \frac{1}{N} \sum_n (nhf) \mathrm{e}^{-nhf/k_B T} = \frac{hf}{\mathrm{e}^{hf/k_B T} - 1}. \tag{2.23}$$

The sums are evaluated using the geometric series formula $\sum_{n=0}^{\infty} ar^n = a/(1-r)$ for $|r| < 1$ and its first derivative with respect to r. The energy density gives the Planck radiation law that agrees with measurements (figure 11)

$$u(f,T) = Ahf^2 \langle E \rangle = Ahf^3 \frac{1}{\mathrm{e}^{hf/k_B T} - 1}. \tag{2.24}$$

The main point of this statistical mechanics detour is: at high frequencies $hf \gg k_B T$, the $\mathrm{e}^{hf/k_B T}$ term dominates and the energy density is exponential suppressed

$$u(f,T) \sim f^3 \mathrm{e}^{-hf/k_B T} \to 0. \tag{2.25}$$

The assumption that photons are quantized into discrete packets $E = hf$ successfully describes thermal radiation.

3 Wavefunction and probability

Wave-particle duality and probabilistic phenomena require reformulating mechanics. Quantum mechanics introduces the **wavefunction**

$$\Psi(x,t). \tag{3.1}$$

The wavefunction Ψ has several defining features:

- It fully specifies the particle's *quantum state*. Classical mechanics specifies a particle state by its position x and velocity $\dot{x}$, whereas quantum mechanics only requires x.
- It is a *complex-valued function*: $\Psi(x,t)$ takes in real numbers of the space and time coordinates (x,t) and outputs a complex number.
- It satisfies the *superposition principle*. Similar to classical water waves, this means if Ψ_1 and Ψ_2 describe a particle state, then so does Ψ_3 for complex numbers A,B

$$\Psi_3(x,t) = A\Psi_1(x,t) + B\Psi_2(x,t). \tag{3.2}$$

3.1 Probability amplitudes

While $\Psi(x,t)$ encodes all information about a quantum state, the wavefunction itself is not directly observable. We instead measure the *probability* of a particle being in a specific location. The probability density $\rho(x,t)$ (probability per unit distance) of a particle at x for time $t = t_0$ is the modulus square of the wavefunction, where * denotes complex conjugation

$$\rho(x,t_0) = |\Psi(x,t_0)|^2 = \Psi^*\Psi. \tag{3.3}$$

This is the **Born rule**, named after Max Born who deduced the concept in 1926 [19]. We interpret $\Psi(x,t)$ as a **probability amplitude**, whose modulus square gives probabilities. The wavefunction must therefore follow rules of probability:

1. The probability density must be non-negative, i.e. positive definite

$$\Psi^*(x,t)\Psi(x,t) \geq 0. \tag{3.4}$$

2. At a particular time t_0, the probability P of finding a particle between positions $a < x < b$ is given by the definite integral (figure 12)

$$P(x_a < x < x_b, t_0) = \int_{x_a}^{x_b} \rho(x,t_0)\,\mathrm{d}x = \int_{x_a}^{x_b} |\Psi(x,t_0)|^2\,\mathrm{d}x. \tag{3.5}$$

   By implication, $\Psi(x,t)$ must be a smooth, square-integrable function.

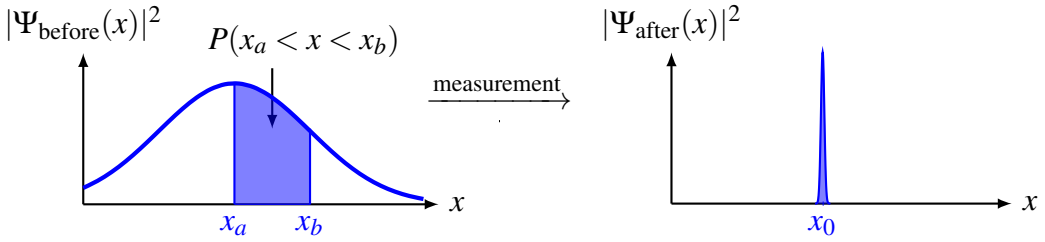


Figure 12: Measurement collapses the particle wavefunction $\Psi(x)$ from a probability amplitude distributed across x to a definite position x_0.

3. The sum over probabilities for all outcomes must be unity, which is called unitarity

$$\int_{-\infty}^{\infty} P(x,t_0)\,\mathrm{d}x = \int_{-\infty}^{\infty} |\Psi(x,t_0)|^2\,\mathrm{d}x = 1. \tag{3.6}$$

The Born rule means we can multiply the wavefunction by a constant complex phase with real α and it still describes the same physical state

$$\Psi' = e^{\mathrm{i}\alpha}\Psi. \tag{3.7}$$

The probabilities are the same because applying the Born rule gives

$$\Psi'^*\Psi' = e^{-\mathrm{i}\alpha}e^{\mathrm{i}\alpha}\Psi^*\Psi = \Psi^*\Psi. \tag{3.8}$$

For simplicity of calculations, we often represent particles using a non-normalized wavefunction $\Phi(x,t)$ with a non-unity integral over all space

$$\int_{-\infty}^{\infty} |\Phi(x,t)|^2\mathrm{d}x = N. \tag{3.9}$$

At the end of calculations, we then evaluate the normalization factor N needed to construct a normalized wavefunction $\Psi(x,t)$ by dividing

$$\Psi(x,t) = \frac{1}{\sqrt{N}}\Phi(x,t), \quad 0 < |N| < \infty. \tag{3.10}$$

To ensure the probabilistic interpretation, this normalization is formally only valid for a wavefunction with non-zero and finite N. This holds if the wavefunction vanishes $\Psi(x,t) \to 0$ sufficiently fast in the limit of infinity $|x| \to \infty$.

When we observe a particle, the wavefunction instantaneously turns into a single position

$$\Psi_{\text{before}} \to \Psi_{\text{after}} = A \lim_{\varepsilon \to 0} \mathrm{e}^{-(x-x_0)^2/2\varepsilon^2}. \tag{3.11}$$

Figure 12 (right) shows how we find a particle at a single point position after measurement. We cannot predict the exact observed position, only its probability. This discontinuous, non-deterministic transition is called **wavefunction collapse**, which we revisit in section 11.

3.2 Laws of probability

Quantum probability motivates a brief review of classical probability to highlight similarities and differences. For details, please see a full probability class e.g. MATH-UA 185.

1. **Independent events multiply**. Consider an event A occurring with probability $P(A)$ and another event B occurring with probability $P(B)$. If these two events are *independent*, the probability of both A and B occurring is multiplicative

$$P(A \text{ and } B) = P(A) \times P(B). \tag{3.12}$$

   Example: consider a biased coin with probability of landing heads being $P(H) = 0.4$. The first throw being heads $P(H_1) = 0.4$ and second throw being heads is $P(H_2) = 0.4$. Each successive throw is independent, so the probability of throw H_1 and H_2 being $P(H_1 \text{ and } H_2)$ heads is $P(H_1) \times P(H_2) = 0.4 \times 0.4 = 0.16$.

2. **Exclusive events add**. If event A occurs with probability $P(A)$ and excludes another event B from happening with probability $P(B)$, then these two events are *exclusive*. Then the probability of either A or B occurring is additive

$$P(A \text{ or } B) = P(A) + P(B). \tag{3.13}$$

   For example in a fair die, the probability of rolling a two is $P(2) = \frac{1}{6}$. If two occurs, then rolling a three cannot occur so these events are exclusive. The probability of rolling a two or three is additive $P(2 \text{ or } 3) = P(2) + P(3) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$. This additive law of real numbers is *insufficient to describe quantum probability amplitudes*.

3. **Total probabilities sum to unity**. For N exclusive events $i = 1, 2, \ldots N$, the sum of all probabilities is unity (or integral over real x for continuous distributions)

$$\sum_{i=1}^{N} P_i = 1, \quad \longleftrightarrow \quad \int P(x)\,\mathrm{d}x = 1. \tag{3.14}$$

   If each exclusive event has equal probability, the individual probabilities are $P_i = 1/N$.

For completeness, we recall key statistical quantities. For a random variable A with outcomes $\{A_i\}$, each occurring with a probability $P(A_i)$, the **expectation value** of A is

$$\langle A \rangle = \sum_{i=1}^{N} A_i P(A_i). \tag{3.15}$$

When N becomes very large, this becomes the mean. The expectation value for the sum of two random variables $\langle A+B\rangle$ is the sum of the individual expectation values

$$\langle A+B\rangle = \langle A\rangle + \langle B\rangle. \tag{3.16}$$

The variance σ^2 is the mean event-by-event differences of A from its expectation value

$$\sigma^2 = \langle (A-\langle A\rangle)^2\rangle = \langle A^2\rangle - 2\langle A\langle A\rangle\rangle + \langle \langle A\rangle^2\rangle = \langle A^2\rangle - \langle A\rangle^2. \tag{3.17}$$

Probability is no stranger to classical physics. When we flip a coin or model a hurricane, its motion and final state is *in principle* determined with 100% certainty by Newton's laws once we specify the initial position and velocity. Classical probabilities arise because we lack the full knowledge of the initial state. It is simpler *in practice* to describe coin flip outcomes as 50% chance of heads and tails. However, quantum theories are fundamentally different:

> Probabilistic observables are an *intrinsic feature* of quantum mechanics.

Quantum probabilities are not due to incomplete knowledge of a deterministic system. The Born rule is an intrinsic feature of quantum theories. Quantum mechanics can only predict probabilities of a particle being in position x, not its precise location with certainty.

3.3 Double-slit experiment

The double-slit experiment is a beautiful demonstration of wave-particle duality. Claus Jönsson in Tübingen [20] first performed laboratory demonstrations of electron double-slit interference in 1961 then in 1989 led by Akira Tonomura at Hitachi, Tokyo [21]. There is a lovely video of the double slit experiment in Ref. [22]. In 2025, the wave behaviour of matter was demonstrated for nanoparticles as large as 7,000 atoms in Vienna [23]. Experiment evidence for interference with single photons came in 1986 [24].

Consider a double-slit experiment where we randomly fire classical particles, such as billiard footballs or marbles, at a barrier with two slits (figure 13). The probability of finding an electron at x when only slit 1 is open is

$$\rho_1(x) = |\Psi_1(x)|^2. \tag{3.18}$$

Similarly, the probability of finding an electron at x when only slit 2 is open is

$$\rho_2(x) = |\Psi_2(x)|^2. \tag{3.19}$$

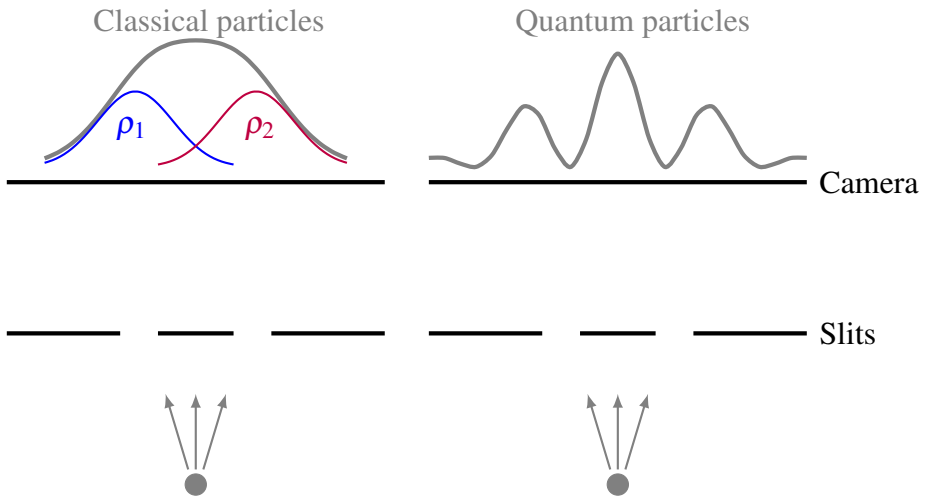


Figure 13: Sketch of double-slit experiment for particles thrown at a barrier with two slits. The classical particle distribution $\rho_{1+2} = \rho_1 + \rho_2$ is the sum of those for the individual slits ρ_1, ρ_2. Quantum particle distribution exhibits a wave-like interference pattern.

When both slits are open, the classical probability ρ_{1+2} of finding an electron at x is the addition of the probabilities of only one open slit

$$\rho_{1+2}^{\text{classical}}(x) = \rho_1(x) + \rho_2(x). \tag{3.20}$$

Classical particles have no concept of interference, so we expect two piles of particles behind the double-slit barrier. However, this simple addition does not hold in quantum mechanics

$$\rho_{1+2}^{\text{quantum}}(x) \neq \rho_1(x) + \rho_2(x). \tag{3.21}$$

Instead, wavefunction superposition implies that we add Ψ then take the modulus square

$$\rho_{1+2}(x) = |\Psi_1(x) + \Psi_2(x)|^2 = |\Psi_1(x)|^2 + |\Psi_2(x)|^2 + \Psi_1^*(x)\Psi_2(x) + \Psi_2^*(x)\Psi_1(x). \tag{3.22}$$

The non-zero cross-terms correspond to **quantum interference**. Letting the wavefunction have an amplitude A and complex phase $\varphi(x)$, we can write $\Psi_1 = A_1 e^{i\varphi_1}$ and $\Psi_2 = A_2 e^{i\varphi_2}$. The two-slit probability density of equation (3.22) then becomes

$$\rho_{1+2}(x) = |A_1(x)|^2 + |A_2(x)|^2 + 2A_1(x)A_2(x)\cos[\varphi_1(x) - \varphi_2(x)]. \tag{3.23}$$

Therefore, the probability of an electron being found at position x of the screen oscillates according to the *relative* phase difference $\varphi_1(x) - \varphi_2(x)$. This interference occurs *even for one electron*, just as water waves are delocalized. The wavefunction of each electron is extended so can take both paths through both slits simultaneously. Restating for emphasis, the additive law for two exclusive events is different for classical versus quantum probabilities:

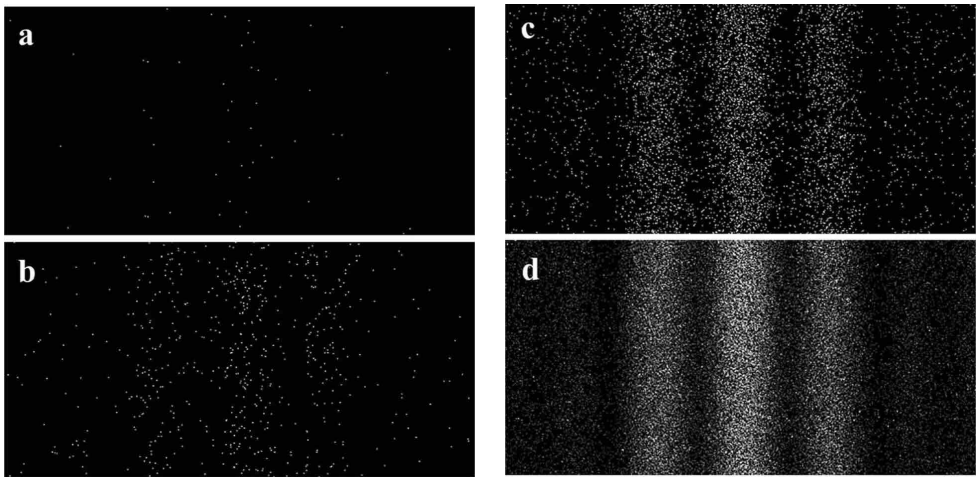


Figure 14: Double-slit experiment with single electrons from Ref. [25]. Fringes are observed after approximately the following number of electrons are recorded in a certain exposure time: (a) 50 in 0.1 s, (b) 500 in 1 s (c) 5,000 in 10 s, and (d) 50,000 in 100 s.

- Classical probabilities (3.13): $P(A \text{ or } B) = P(A) + P(B)$ for real numbers $P(A), P(B)$.
- Quantum probabilities: $P(A \text{ or } B) = |\Psi(A) + \Psi(B)|^2$ for complex numbers $\Psi(A), \Psi(B)$.

No other realm of science uses the quantum additive law for probability. Addition of complex probability amplitudes is unique to quantum mechanics.

HW2: Atoms and wavefunctions (set Sep 10, due Sep 17)

1. Describe the Frank-Hertz experiment and how the results demonstrate quantized energy levels of atoms. Consider the Bohr model of hydrogen:

   (a) What is the empirical evidence for the existence of an atomic nucleus?

   (b) By equating centripetal and electrostatic forces then using Bohr quantization $L = n\hbar$ of angular momentum $L = mvr$, show that the possible radii are $r_n = (\hbar n^2)/(mc\alpha)$ and find the expression for the dimensionless constant α.

   (c) Find an expression for the electron velocity v with circular Bohr orbits in terms of the speed of light c and quantum number n. Evaluate v/c for $n = 1$.

   (d) Evaluate the numerical pre-factors to show the energies are $E_n = -13.6\ \text{eV}/n^2$. What is the energy and wavelength of the photon for $n = 3$ transitioning to $n = 2$?

2. State what a wavefunction $\Psi(x)$ physically represents.

   (a) Using the normalization condition, what are the S. I. units of the wavefunction $\Psi(x)$ as a function of one-dimensional position x? Explain how you arrived at this answer. How does this change in three spatial dimensions?

   (b) Given a wavefunction $\Psi(x)$, what is the expression for the probability of finding a particle in $a < x < b$? What mathematical properties must $\Psi(x)$ satisfy to be considered a probability amplitude in quantum mechanics?

   (c) Given a wavefunction at $t = 0$ is $\Psi(x) = \frac{1}{\sqrt{2\pi}} e^{-(x-a)^2}$, where $a = 1$ m, write down the integral representing the probability $P(x)$ of finding a particle between position $x = 0.3$ m and $x = 1.2$ m and sketch the area that is represented by this integral. You do not need to calculate the integral.

3. The double-slit experiment has a source, barrier with slits of width a, and screen.

   (a) Qualitatively sketch and describe the distribution on the screen for: (i) classical marbles, (ii) classical water waves with wavelength $\lambda \approx a$.

   (b) Repeat these sketches for N quantum electrons with de Broglie wavelength $\lambda \approx a$ for (iii) $N = 1$ and (iv) $N \simeq 30$.

   (c) Discuss qualitatively the similarities and differences between the classical and quantum cases. Why are classical probabilities insufficient for quantum physics?

4 Schrödinger equation

One important aspect of quantum mechanics that nonetheless *is* deterministic is the time evolution of the wavefunction. In quantum mechanics, the **time-dependent Schrödinger equation** governs how a wavefunction $\Psi(x,t)$ changes with time

$$\mathrm{i}\hbar\frac{\partial\Psi}{\partial t} = \hat{H}\Psi. \tag{4.1}$$

Erwin Schrödinger published this landmark equation in 1926 via German [26] and English [27] papers. It is certainly momentous to start teaching this class on the centenary of such a pivotal equation in science! The right-hand side has an important object called the **Hamiltonian** $\hat{H}$, which is a differential operator that conventionally reads

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V. \tag{4.2}$$

Schrödinger's equation is the quantum mechanical analogue of Newton's second law $F = ma$. It governs how a particle moves in the presence of a force $F = \partial V/\partial x$ from a potential $V(x)$.

As with all differential equations in physics, they tell us how things change. We can read Schrödinger's equation out in words: the rate of temporal change of Ψ is related to the rate of change in the rate of spatial change of Ψ by a constant $\hbar$ and particle mass m. Showing the remarkable unity of physics, Schrödinger's equation looks similar (differing only by a factor of i) to the diffusion equation that we encounter in thermodynamics and fluid dynamics

$$\frac{\partial\Psi}{\partial t} = D\frac{\partial^2\Psi}{\partial x^2}, \tag{4.3}$$

and the wave equation we encounter in acoustics and electromagnetism

$$\frac{\partial^2\Psi}{\partial t^2} = c^2\frac{\partial^2\Psi}{\partial x^2}. \tag{4.4}$$

In these cases, the constant that relates the rates of change in time to position are the diffusion constant D and speed of sound or light c. These differential equations are linear in Ψ, which ensures the superposition principle holds. Heat, light, water, matter waves and their linear combinations are also solutions. Non-linear terms may look like Ψ^2 or $\Psi^3(\partial\Psi/\partial x)$, which are absent from Schrödinger's equation.

4.1 Energy and momentum operators

Given the evidence for the wave behaviour of quantum particles, we can model a particle as a plane wave using a complex exponential with angular frequency ω and wavenumber k:

$$\Psi(x,t) = A\mathrm{e}^{-\mathrm{i}(\omega t - kx)}, \tag{4.5}$$

where A is a fixed amplitude of the wavefunction. Recall the Euler identity relates this to trigonometric functions $e^{i\theta} = \cos\theta + i\sin\theta$. The Planck-Einstein $E = \hbar\omega$ and de Broglie $p = \hbar k$ relations map this to the particle's energy and momentum

$$\Psi(x,t) = A e^{-i(Et-px)/\hbar}. \tag{4.6}$$

We now use this wavefunction to illuminate the physics in the Schrödinger equation. If we act on this wavefunction with the left-hand side of the Schrödinger equation (4.1), we see the time derivative brings down the energy part of the wavefunction

$$i\hbar\frac{\partial\Psi}{\partial t} = EAe^{-i(Et-px)/\hbar} = E\Psi. \tag{4.7}$$

Similarly, the double partial derivative in space on the right-hand side of the Schrödinger equation (4.1) brings down the momentum part of Ψ

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V\right]\Psi = \left[\frac{p^2}{2m} + V\right]Ae^{-i(Et-px)/\hbar} = \left[\frac{p^2}{2m} + V\right]\Psi. \tag{4.8}$$

Comparing the two sides of our operations (4.7) and (4.8), we draw the connection

$$i\hbar\frac{\partial\Psi}{\partial t} = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V\right)\Psi \quad \longleftrightarrow \quad E = \frac{p^2}{2m} + V. \tag{4.9}$$

This looks an awful lot like the total energy of a system familiar from classical mechanics with momentum $p = mv$ and kinetic energy $T = \frac{1}{2}mv^2$. The classical Hamiltonian[4] is the sum of the kinetic and potential energy for a particle's motion

$$H = T + V = \frac{p^2}{2m} + V. \tag{4.10}$$

Schrödinger's equation identifies total energy and momentum with **differential operators**

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V, \quad \hat{p} = -i\hbar\frac{\partial}{\partial x}. \tag{4.11}$$

The hats above $\hat{H}$ and $\hat{p}$ remind us that these are not numbers, but rather operators that differentiate complex functions $\Psi(x,t)$. Relating observables to operators is a central concept in quantum theory that we formalize in section 10.

[4] In analytical dynamics, the Hamiltonian $H = T + V$ and Lagrangian $L = T - V$ are scalars defined as linear combinations of the kinetic T and potential V energies of a particle. The principle of least action $\delta\int L\mathrm{d}t = 0$ provides the vectorial equations of motion, namely Hamilton's equations or the Euler-Lagrange equation $\frac{\partial L}{\partial \mathbf{x}} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\mathbf{x}}}\right)$ that correspond to Newton's law $\mathbf{F} = m\ddot{\mathbf{x}} = \boldsymbol{\nabla} V$.

4.2 Conservation of probability

Schrödinger's equation modifies the wavefunction $\Psi(x,t)$ with time. It is important to check that changes to Ψ do not spoil the probabilistic interpretation of $\rho = \Psi^*\Psi$. We differentiate with respect to time using the product rule

$$\frac{\partial \rho}{\partial t} = \frac{\partial(\Psi^*\Psi)}{\partial t} = \Psi^*\frac{\partial \Psi}{\partial t} + \frac{\partial \Psi^*}{\partial t}\Psi. \tag{4.12}$$

The Schrödinger equation and its complex conjugate are

$$\frac{\partial \Psi}{\partial t} = -\frac{\mathrm{i}}{\hbar}\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V\right)\Psi, \quad \frac{\partial \Psi^*}{\partial t} = +\frac{\mathrm{i}}{\hbar}\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V\right)\Psi^*. \tag{4.13}$$

Inserting these expressions into the time-variation of probability density in equation (4.12) eliminates the potential, leaving

$$\frac{\partial \rho}{\partial t} = \frac{\mathrm{i}\hbar}{2m}\left(\Psi^*\frac{\partial^2 \Psi}{\partial x^2} - \Psi\frac{\partial^2 \Psi^*}{\partial x^2}\right) = \frac{\partial}{\partial x}\left[\frac{\mathrm{i}\hbar}{2m}\left(\Psi^*\frac{\partial \Psi}{\partial x} - \Psi\frac{\partial \Psi^*}{\partial x}\right)\right]. \tag{4.14}$$

Having moved one partial derivative in space out, we can define the quantity in square brackets as the probability current density $j(x,t)$

$$j(x,t) = -\frac{\mathrm{i}\hbar}{2m}\left(\Psi^*\frac{\partial \Psi}{\partial x} - \Psi\frac{\partial \Psi^*}{\partial x}\right). \tag{4.15}$$

Then we can write equation (4.14) compactly as

$$\frac{\partial \rho}{\partial t} + \frac{\partial j}{\partial x} = 0 \tag{4.16}$$

This is the **probability continuity equation**, which implies a conserved quantity. This equation expresses the conservation of probability density $\rho = \Psi^*\Psi$ in quantum mechanics. To see this manifestly, we perform a definite integral over an interval in space $a < x < b$ to find

$$\frac{\partial}{\partial t}\left(\int_a^b \rho \,\mathrm{d}x\right) = -\int_a^b \frac{\partial j}{\partial x}\mathrm{d}x = j(x=a,t) - j(x=b,t). \tag{4.17}$$

This tells us that the rate of change of probability to find a particle in $a < x < b$ is equal to the current density of particles entering minus that leaving this interval (figure 15). Importantly, this is a statement of *local* probability conservation. The probability density cannot just spontaneously disappear from one region of space and reappear in a distant corner of the universe. If the probability of finding a particle in this classroom is lower now than it was yesterday, that is because it had moved into a neighbouring part of space.

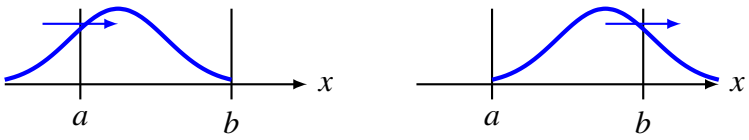


Figure 15: Flux of probability density entering and leaving an interval between two times.

The continuity equation generalizes to three dimensions with the divergence $\partial/\partial x \to \boldsymbol{\nabla}$

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{j} = 0, \tag{4.18}$$

which is reminiscent of of continuity equations describing mass density conservation $\rho = m/V$ in fluid dynamics or charge density conservation $\rho = Q/V$ in electromagnetism. The quantum probability current density $\mathbf{j}(\mathbf{r},t)$ is then given by

$$\mathbf{j} = \frac{\hbar}{2mi}(\Psi^* \boldsymbol{\nabla} \Psi - \Psi \boldsymbol{\nabla} \Psi^*). \tag{4.19}$$

4.3 Stationary states

This class only studies Hamiltonians that that are static. This means that the potential does not change with time and depends only on space $V = V(x)$:

$$\mathrm{i}\hbar \frac{\partial \Psi(x,t)}{\partial t} = \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x) \right] \Psi(x,t). \tag{4.20}$$

We can then solve the Schrödinger equation by separation of variables. Let $\Psi(x,t) = \psi(x)\phi(t)$ be a product of space- and time-dependent functions. Then we can separate the time-dependent and time-independent parts to read

$$\frac{\mathrm{i}\hbar}{\phi(t)}\frac{\partial \phi(t)}{\partial t} = \frac{1}{\psi(x)}\left[-\frac{\hbar^2}{2m}\frac{\partial^2 \psi(x)}{\partial x^2} + V(x)\psi(x) \right] = \text{constant}. \tag{4.21}$$

The left and right hand side must be the same and is equal to a constant that we identify as the energy E. The time-dependent side then reads

$$\frac{\mathrm{i}\hbar}{\phi(t)}\frac{\partial \phi(t)}{\partial t} = E. \tag{4.22}$$

We can solve this first-order differential equation by integrating both sides. The time evolution involves a change of phase with time at a rate set by the energy over $\hbar$

$$\phi(t) = \phi(0)\mathrm{e}^{-\mathrm{i}Et/\hbar}, \tag{4.23}$$

for some constant $\phi(0)$ that is the initial state at $t = 0$, which we could take as unity $\phi(0) = 1$ without loss of generality. The full time-dependent solution is given by

$$\Psi(x,t) = \mathrm{e}^{-\mathrm{i}Et/\hbar}\psi(x). \tag{4.24}$$

As time evolution is only a phase change, the Born rule implies that the probability remains constant in time $\Psi^*\Psi = |\psi(x)|^2$. We therefore call these solutions **stationary states**. The remaining right-hand side of Eq. (4.21) is the time-independent Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x)}{\partial x^2} + V(x)\psi(x) = E\psi(x). \tag{4.25}$$

With the Hamiltonian as the differential operator, we can write this compactly as

$$\hat{H}\psi(x) = E\psi(x). \tag{4.26}$$

This takes the form of an **eigenvalue equation**. In linear algebra formalism, the left-hand side is an operator and the right-hand side is the energy eigenvalue of the operator. Different forms of $V(x)$ govern distinct particle behaviour. The next part studies various solutions to illustrate wave properties governed by this equation.

II Wave mechanics

5 Nearly free particle

We now solve the time-independent Schrödinger equation $\hat{H}\psi = E\psi$ (4.26) for various $V(x)$. Consider the simplest case of a free particle where the potential is zero everywhere

$$V(x) = 0. \tag{5.1}$$

Then $\hat{H}\psi(x) = E\psi(x)$ becomes a second-order differential equation with the form of an undamped harmonic oscillator

$$\frac{\partial^2 \psi}{\partial x^2} = -k^2\psi, \quad k^2 = \frac{2mE}{\hbar^2}. \tag{5.2}$$

The solution is a plane wave with wavenumber k that can take any real number as its value

$$\psi_{\text{free}}(x) = \mathrm{e}^{\mathrm{i}kx}, \tag{5.3}$$

where the eigenvalue is the energy

$$E_k = \frac{\hbar^2 k^2}{2m}. \tag{5.4}$$

A delicate mathematical issue with this unconstrained plane-wave solution is that the wavefunction ψ_{free} is non-normalizable:

$$\int_{-\infty}^{\infty} |\psi_{\text{free}}|^2 \,\mathrm{d}x = \int_{-\infty}^{\infty} 1 \,\mathrm{d}x \to \infty \tag{5.5}$$

Of course, exactly "free" states do not exist in nature given there are always forces, but they are very useful models of physical systems. We resolve this by first considering two limiting cases of the potential: infinite well potential and the Dirac delta potential.

5.1 Infinite well potential

We confine the particle to be free in a well of finite length L with infinite potential outside. This is like being trapped in a well with infinite height: no matter how high you jump up,

there is no way out. The potential $V(x)$ is defined by

$$V(x) = \begin{cases} 0, & 0 < x < L, \\ \infty, & x \leq 0, x \leq L. \end{cases} \tag{5.6}$$

Outside this region where $V(x) \to \infty$, the wavefunction vanishes. For $0 < x < L$, the particle feels no potential so is still free but trapped inside this region. We could picture the wavefunction vanishing at the edge of the lab or universe.

We propose a solution that is a superposition of plane waves moving forward and backward with momentum $p = \pm\hbar k$

$$\psi(x) = A\mathrm{e}^{+\mathrm{i}kx} + B\mathrm{e}^{-\mathrm{i}kx}, \tag{5.7}$$

where A, B are normalization constants to be determined. Imposing the condition that $\psi = 0$ at $x = 0$ requires $B = -A$, which yields

$$\psi(x) = A\left(\mathrm{e}^{+\mathrm{i}kx} - \mathrm{e}^{-\mathrm{i}kx}\right) = 2\mathrm{i}A\sin(kx). \tag{5.8}$$

Imposing the condition that $\psi = 0$ at $x = L$ means $\sin(kL) = 0$, which implies k must take discrete values

$$k_n = \frac{n\pi}{L}, \quad \text{where } n = 1, 2, 3 \ldots \tag{5.9}$$

Here, n is a positive integer. Applying the energy of a free particle $E = \hbar^2 k^2/2m$ gives quantized energies

$$E_n = \frac{\hbar^2}{2m}\left(\frac{n\pi}{L}\right)^2. \tag{5.10}$$

The normalized wavefunctions are then

$$\psi_n(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi x}{L}\right). \tag{5.11}$$

Figure 16 shows the first few modes of these wavefunction solutions. They look like normal modes of a guitar string fixed at either end so that it cannot move outside its length, but can oscillate within $0 < x < L$. That's not a bad picture. We could call the wavefunctions of an infinite-wall potential guitar-string-like solutions. The quantized solutions arise from the boundary condition of fixing the wave so that it cannot move, just like the excited harmonics of a guitar string.

This infinite well may seem like a contrived toy model. Does it have any hope of describing nature? If we plugged in (one significant digit) values for the electron mass $m_e c^2 \approx$

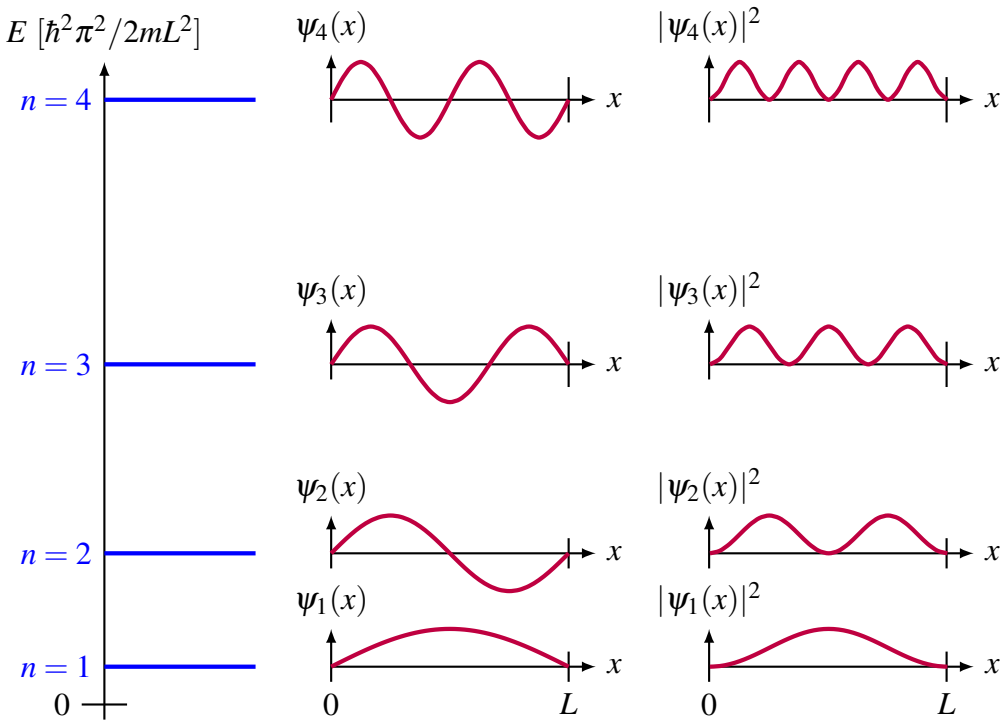


Figure 16: The energy eigenvalues E_n, wavefunctions $\psi_n(x)$, and its probability density $|\psi_n(x)|^2$ for the infinite-wall potential with width L. These solutions are analogous to normal modes of an oscillating guitar string.

500 keV, $\hbar c \approx 200$ eV nm, and take the width to be twice the Bohr radius $L = 2r_B \approx 0.1$ nm, we already get the right order of magnitude for the hydrogen energy levels

$$E_n = \frac{(\hbar c)^2}{2m_e c^2}\frac{\pi^2}{(2r_B)^2}\frac{1}{n^2} \simeq \frac{(200\text{ eV nm})^2}{2\times 5\times 10^5\text{ eV}}\frac{10}{(0.1\text{ nm})^2}\frac{1}{n^2} \simeq \frac{10\text{ eV}}{n^2}. \tag{5.12}$$

The correct value is 13.6 eV for $n = 1$. Despite appearing contrived, this dimensional analysis shows the relevant scales of hydrogen involve the electron mass and Bohr radius.

5.2 Dirac delta potential

Consider another limiting case where the potential is zero (so the particle is "free") everywhere except at the origin. We model this mathematically using a Dirac delta function potential centred at the origin $x = 0$ for simplicity

$$V(x) = \begin{cases} 0 & x \neq 0, \\ -V_0\delta(x) & x = 0. \end{cases} \tag{5.13}$$

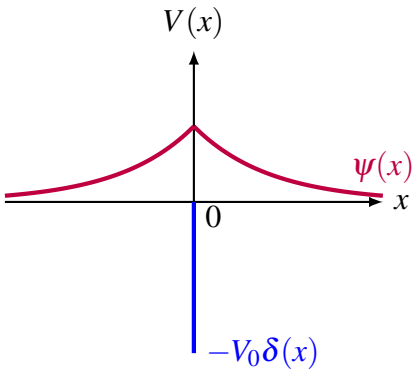


Figure 17: Delta function potential (blue) and the single wavefunction bound state (red).

The delta function is not a real function but rather a distribution. It is a spike with infinitesimal width, defined in terms of integration

$$\int_{-\infty}^{\infty} \delta(x)\mathrm{d}x = 1, \quad \delta(x) = 0 \text{ for } x \neq 0. \tag{5.14}$$

Heuristically, the delta function is "infinite" at $x = 0$ and zero elsewhere. We wish to determine the discontinuity of the wavefunction's first derivative. To do this, we integrate the time-independent Schrödinger equation either side of the δ function at the origin by a small amount $[-\varepsilon, +\varepsilon]$

$$-\frac{\hbar^2}{2m}\int_{-\varepsilon}^{+\varepsilon}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}\mathrm{d}x - V_0\int_{-\varepsilon}^{+\varepsilon}\delta(x)\psi(x)\mathrm{d}x = E\int_{-\varepsilon}^{+\varepsilon}\psi(x)\mathrm{d}x. \tag{5.15}$$

The integral on the right-hand side vanishes in the limit $\varepsilon \to 0$ because $\psi(x)$ is finite. The delta function has the sampling property

$$\int_{-\infty}^{\infty}\delta(x-a)f(x)\mathrm{d}x = f(a). \tag{5.16}$$

Therefore, we find the boundary condition across a δ function is

$$\lim_{\varepsilon\to 0}\left(\left.\frac{\mathrm{d}\psi}{\mathrm{d}x}\right|_{+} - \left.\frac{\mathrm{d}\psi}{\mathrm{d}x}\right|_{-}\right) = -\frac{2mV_0}{\hbar^2}\psi(0). \tag{5.17}$$

We learn that the delta function potential causes the first derivative of the wavefunction to be discontinuous. Nonetheless, the wavefunction $\psi(x)$ remains continuous at the Dirac delta function boundary.

We can then propose the wavefunction solutions to be for a positive $\kappa > 0$

$$\psi(x) = \begin{cases} \psi_- = Ae^{+\kappa x}, & x < 0, \\ \psi_+ = Ae^{-\kappa x}, & x > 0. \end{cases} \tag{5.18}$$

Wavefunction normalizability requires the asymptotic behaviour to be exponentially suppressed as $x \to +\infty$ and $x \to -\infty$. The wavefunction remains continuous in the limit of approaching the origin

$$\lim_{\varepsilon \to 0} \psi_+(\varepsilon) = \lim_{\varepsilon \to 0} \psi_-(\varepsilon) = A. \tag{5.19}$$

However, the first derivative is discontinuous

$$\lim_{\varepsilon \to 0} \left(\left. \frac{d\psi_+}{dx} \right|_\varepsilon - \left. \frac{d\psi_0}{dx} \right|_\varepsilon \right) = -A\kappa \lim_{\varepsilon \to 0} \left(e^{-\kappa\varepsilon} - e^{+\kappa\varepsilon} \right) = -2A\kappa. \tag{5.20}$$

This equals the boundary condition of equation (5.17) so we find

$$\kappa = \frac{mV_0}{\hbar^2}. \tag{5.21}$$

Therefore, the particle can have one single non-trivial bound state with kinetic energy

$$E = -\frac{\hbar^2\kappa^2}{2m} = -\frac{mV_0^2}{2\hbar^2}. \tag{5.22}$$

Figure 17 illustrates the bound state wavefunction of the delta function potential. This potential may seem like a contrived toy, but the ground-state wavefunction of hydrogen has precisely the $\sim e^{-kr}$ shape, where the proton is like a point at $r \to 0$. So hold this thought because we see this again in section 19.

HW3: Schrödinger equation (set Sep 17, due Sep 24)

1. Write down the time-dependent Schrödinger equation (TDSE) and explain each term.
   (a) Why is the TDSE first order in time while Newton's law $f = m\ddot{x}$ is second order?
   (b) What does the Hamiltonian $\hat{H}$ represent physically? Show that the wavefunction $\Psi(x,t) = Ae^{-i(\omega t - kx)}$ satisfies the TDSE. What is the physical significance of ω and k, $\hat{H}$, and how they are related to differential operators?
   (c) Explain the condition(s) required and show the derivation from the TDSE to reach the time-independent counterpart. What meant physically by *stationary state*?
   (d) Integrate the TDSE to derive how an initial wavefunction $\phi(t=0) = \phi_0$ evolves with time. For a wavefunction initially in state $\phi(t=0) = e^{i\pi/2}$ with energy E, what is the wavefunction at time $t = T$?

2. By taking the time derivative of the probability density $\rho = \Phi^*\Phi$ for the wavefunction $\Phi(z,t)$ and applying Schrödinger's equation, show that it satisfies
$$\frac{\partial \rho}{\partial t} + \frac{\partial j}{\partial z} = 0. \tag{5.23}$$
Find the expression for $j(z,t)$. What is the physical significance of this equation? Evaluate $j(z)$ for the plane wave states $\psi = e^{\pm ikz}$ and how is this related to velocity?

3. A free particle moves on a circle with radius R obeying Schrödinger's equation with $V(x) = 0$ and periodic boundary conditions $\psi(x) = \psi(x + 2\pi R)$. Show that the momentum p_n is integer n quantized and derive an expression for the energy E_n.

4. Write down the Schrödinger equation for wavefunction $\psi(x)$ in the potential
$$V(x) = \begin{cases} 0, & |x| < \frac{a}{2}, \\ \infty & \text{otherwise.} \end{cases} \tag{5.24}$$
   (a) What is the boundary condition at $x = \pm a/2$? Show that the energy eigenvalues E_n are integer n quantized and derive the wavefunctions $\psi_n(x)$ for $n = 1,2,3$.
   (b) Determine the normalization constant for the wavefunctions in the ground $\psi_1(x)$, first $\psi_2(x)$ and second $\psi_3(x)$ excited states.
   (c) Sketch the forms of the wavefunction $\psi_n(x)$ and probability distributions $|\psi_n(x)|^2$.
   (d) Show that the wavefunctions $\psi_1(x)$ and $\psi_2(x)$ are orthogonal
$$\int_{-a/2}^{a/2} \psi_1(x)\psi_2(x) = 0. \tag{5.25}$$

6 Bound states

We now consider potentials $V(x)$ that are less extreme and have finite values. The finite-well potential shown in figure 18 is defined by

$$V(x) = \begin{cases} -V_0, & |x| < a, \\ 0, & |x| > a. \end{cases} \tag{6.1}$$

This is similar to the infinite-well potential but with finite depth, or the delta function but with finite width and depth. We present a detailed analysis of this potential and its solutions.

6.1 Asymptotic behaviour

The finite well potential is *localized* such that it asymptotes to zero $V(x) \to 0$ in the far-away limit $x \to \pm\infty$, so the Schrödinger equation becomes free:

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + k^2\psi = 0, \quad k^2 = \frac{2mE}{\hbar^2}. \tag{6.2}$$

There are two classes of solutions:

- **Scattering states**. For particles with energy $E > 0$, we have with plane-wave general solutions for real k

$$\psi(x) = A\mathrm{e}^{\mathrm{i}kx} + B\mathrm{e}^{-\mathrm{i}kx}, \qquad E = \frac{\hbar^2 k^2}{2m} > 0. \tag{6.3}$$

- **Bound states**. For particles with energy $E < 0$, we have general solutions taking an exponential form with real κ

$$\psi(x) = A\mathrm{e}^{\kappa x} + B\mathrm{e}^{-\kappa x}, \qquad E = -\frac{\hbar^2 \kappa^2}{2m} < 0 \tag{6.4}$$

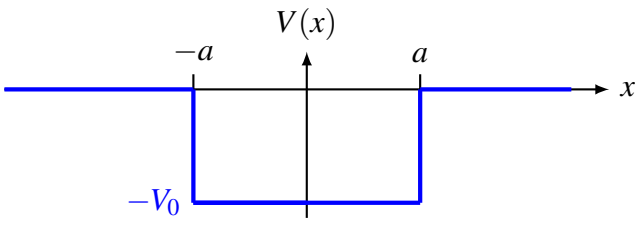


Figure 18: Finite well potential.

for real values of κ. This is equivalent to $k^2 = -\kappa^2$ with imaginary k. For bound states, the wavefunction is localized in space and the normalization condition ensures it must vanish as $x \to \pm\infty$ with the following asymptotic behaviour

$$\psi(x) \sim \begin{cases} e^{\kappa x}, & x \to -\infty, \\ e^{-\kappa x}, & x \to +\infty. \end{cases} \tag{6.5}$$

In both cases, A and B are integration constants fixed by boundary conditions.

6.2 Boundary conditions

Near the finite potential, the Schrödinger equation rearranges into

$$\frac{d^2\psi}{dx^2} + \frac{2m[E - V(x)]}{\hbar^2}\psi = 0. \tag{6.6}$$

As $|\Psi(x,t)|^2$ is a finite probability density, the wavefunction $\psi(x)$ must also be finite as per the normalizability condition. To derive boundary conditions, we integrate the Schrödinger equation (6.6) over a small interval $[a-\varepsilon, a+\varepsilon]$ at $x = a$ across a change in potential

$$\int_{a-\varepsilon}^{a+\varepsilon} \frac{d^2\psi}{dx^2} dx = -\frac{2m}{\hbar^2} \int_{a-\varepsilon}^{a+\varepsilon} [E - V(x)]\psi(x) dx. \tag{6.7}$$

Performing the integral, this becomes a constraint on the first derivative

$$\lim_{\varepsilon \to 0} \left(\frac{d\psi}{dx}\bigg|_{a+\varepsilon} - \frac{d\psi}{dx}\bigg|_{a-\varepsilon} \right) = -\frac{2m}{\hbar^2} \int_{a-\varepsilon}^{a+\varepsilon} [E - V(x)]\psi(x) dx. \tag{6.8}$$

If $V(x = a)$ is finite, the right-hand side integrates to a finite number multiplied by ε, which goes to zero in the limit $\varepsilon \to 0$. This implies that even if $V(x)$ has a discontinuity at $x = a$, its first spatial derivative $\partial\psi/\partial x$ and its integral ψ must be continuous functions of x. To summarize, the wavefunction is continuous across a boundary at $x = a$ between two potentials

$$\psi_1(a) = \psi_2(a), \tag{6.9}$$

where ψ_1 is the wavefunction at $x < a$ and ψ_2 is the wavefunction at $x > a$. The wavefunction's first derivative evaluated at $x = a$ is continuous

$$\frac{d\psi_1}{dx}\bigg|_a = \frac{d\psi_2}{dx}\bigg|_a. \tag{6.10}$$

6.3 Parity

We often work with potentials that are symmetric $V(x) = V(-x)$ about the origin, as in the finite well (6.1). This motivates the concept of parity symmetry to ease life. Inverting coordinates through the origin $x \to -x$ is called a parity transformation. If $V(x) = V(-x)$, then the Hamiltonian is also invariant upon parity inversion $\hat{H}(x) = \hat{H}(-x)$. This means wavefunctions solving the Schrödinger equation for a particular energy E must be either:

- even functions that are symmetric in $x \to -x$, e.g. $\psi(x) = x^2$ or $\psi(x) = \cos(x)$;
- odd functions that are antisymmetric in $x \to -x$, e.g. $\psi(x) = x^3$ or $\psi(x) = \sin(x)$.

The wavefunctions have well-defined parity i.e. are not a mixture of odd and even functions.

Mathematically, we define the parity operator $\hat{P}$ that flips the coordinate sign $x \to -x$

$$\hat{P}\psi(x) = \psi(-x). \tag{6.11}$$

Acting twice, we find $\hat{P}^2$ is the identity i.e. it returns the wavefunction to its original state

$$\hat{P}^2\psi(x) = \hat{P}\psi(-x) = \psi(x). \tag{6.12}$$

We wish to determine the eigenvalue λ of $\hat{P}$ that satisfies the eigenvalue equation

$$\hat{P}\psi(x) = \lambda\,\psi(x). \tag{6.13}$$

If we act on the wavefunction twice, we obtain

$$\hat{P}^2\psi(x) = \hat{P}(\lambda\,\psi(x)) = \lambda\,(\hat{P}\psi(x)) = \lambda^2\psi(x). \tag{6.14}$$

Therefore, $\lambda^2 = 1$ and we find two eigenvalues $\lambda = \pm 1$. The $\lambda = +1$ eigenvalue corresponds to even-parity wavefunctions

$$\psi(x) = \psi(-x), \quad \text{even parity (symmetric)}. \tag{6.15}$$

The $\lambda = -1$ eigenvalue corresponds to odd-parity wavefunctions

$$\psi(x) = -\psi(-x), \quad \text{odd parity (antisymmetric)}. \tag{6.16}$$

Some New York history, parity was once thought to be a symmetry of nature, heuristically "the laws of physics do not distinguish left- and right- handed mirror reflections". Electromagnetism, gravity, and the strong force are all seen to conserve parity, but the weak force governing radioactive decay does not. This shocking discovery was published in 1957 using cobalt-60 decays by Chien-Shiung Wu of Columbia University [28], where she led a team of Ernest Ambler, Raymond Hayward, Dale Hoppes, and Ralph Hudson at the National Bureau Standards. Published concurrently was its observation in pion decays by Richard Garwin, Leon Lederman, and Marcel Weinrich using the Nevis Cyclotron Laboratory [29], by Ardsley-on-Hudson just 45 minutes on the Metro North train from Grand Central.

6.4 Finite well potential

We are finally equipped to solve the finite-well potential (6.1). The Schrödinger equations for the different regions are

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + \kappa^2\psi = 0, \quad \kappa^2 = -\frac{2mE}{\hbar^2}, \qquad |x| > a, \tag{6.17}$$

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + k^2\psi = 0, \quad k^2 = \frac{2m(E+V_0)}{\hbar^2}, \quad |x| < a. \tag{6.18}$$

In the second equation for the $|x| < a$ region, we can either have

- $E < -V_0$: kinetic energies lower than the potential with exponential $\sim \mathrm{e}^{\pm\kappa x}$ solutions.
- $E > -V_0$: kinetic energies above than the potential with plane-wave $\sim \mathrm{e}^{\pm \mathrm{i}kx}$ solutions.

We focus on the second case, where bound-state energies satisfy $-V_0 < E < 0$.

The general wavefunctions are then either exponential solutions for $|x| > a$ or plane-wave solutions for $|x| < a$ in the three regions:

$$\psi(x) = \begin{cases} \psi_1(x) = A\mathrm{e}^{\kappa x}, & x < -a, \\ \psi_2(x) = B\cos(kx) + C\sin(kx), & -a < x < a, \\ \psi_3(x) = D\mathrm{e}^{-\kappa x}, & x > a. \end{cases} \tag{6.19}$$

The solutions must be either even parity $\psi^+(x) = \psi^+(-x)$ or odd parity $\psi^-(x) = -\psi^-(-x)$, so we work directly with $\cos(kx)$ and $\sin(kx)$ instead of starting with the plane-wave form $\psi_2(x) = B'\mathrm{e}^{\mathrm{i}kx} + C'\mathrm{e}^{-\mathrm{i}kx}$. Consider each parity in turn:

- **Even parity solutions**. Even solutions $\psi(x) = \psi(-x)$ require that $A = D$ because $\psi_1(a) = \psi_3(-a)$ and $C = 0$ because $\psi_2(a) = \psi_2(-a)$. We then obtain

$$\psi^+(x) = \begin{cases} \psi_1(x) = A\mathrm{e}^{\kappa x}, & x < -a, \\ \psi_2(x) = B\cos(kx), & -a < x < a, \\ \psi_3(x) = A\mathrm{e}^{-\kappa x}, & x > a. \end{cases} \tag{6.20}$$

  Applying the boundary conditions that the wavefunction and its first derivative at $x = -a$ be continuous gives

$$\psi_1(-a) = \psi_2(-a) \quad \Rightarrow \quad A\mathrm{e}^{-\kappa a} = B\cos(ka), \tag{6.21}$$

$$\left.\frac{\mathrm{d}\psi_1}{\mathrm{d}x}\right|_{-a} = \left.\frac{\mathrm{d}\psi_2}{\mathrm{d}x}\right|_{-a} \quad \Rightarrow \quad \kappa A\mathrm{e}^{-\kappa a} = kB\sin(ka). \tag{6.22}$$

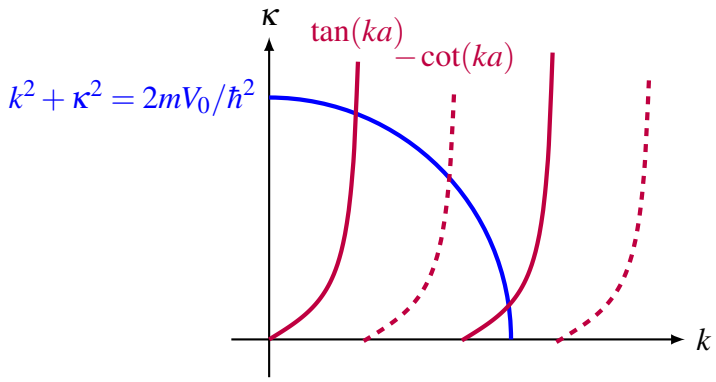


Figure 19: Finite-well solutions in k-κ plane for the $k>0, \kappa>0$ region. Displayed are the energy constraint $k^2+\kappa^2 = 2mV_0/\hbar^2$ (blue solid line), even-parity condition $\kappa = k\tan(ka)$ (purple solid line), and the odd-parity condition $\kappa = -k\cot(ka)$ (purple dashed line).

There is no need to apply boundary conditions at $x=+a$ because it is related by parity $\psi(x)=\psi(-x)$. Dividing these two equations provides a relation between k and κ

$$k\tan(ka) = \kappa. \tag{6.23}$$

- **Odd parity solutions**. An analogous calculation (homework exercise) for the odd-parity solutions gives the relation between k and κ as

$$k\cot(ka) = -\kappa. \tag{6.24}$$

We summarize the two conditions on k and κ that arise from boundary conditions as

$$\kappa = \begin{cases} +k\tan(ka), & \text{even-parity solutions (6.23),} \\ -k\cot(ka), & \text{odd-parity solutions (6.24) .} \end{cases} \tag{6.25}$$

These are transcendental equations, which have no analytic solution. We instead understand them graphically by sketching them in the k-κ plane shown in figure 19. We can also relate k and κ by their definitions in equations (6.17) and (6.18) to eliminate E:

$$k^2+\kappa^2 = \frac{2mV_0}{\hbar^2}. \tag{6.26}$$

This is a circle with radius $2mV_0/\hbar^2$. Solutions occur where this circle intersects with the lines $k\tan(ka)$ or $-k\cot(ka)$. The $\tan(ka)$ and $-\cot(ka)$ functions are periodic in 2π, so we expect discrete solutions approximately every

$$k_n = \frac{2\pi n}{a}, \quad n=1,2 \tag{6.27}$$

7 Gaussian wavepacket

7.1 Fourier modes

Recall the sinusoidal solutions for the infinite well in section 5.1. By the superposition principle, we can describe any continuous function as a sum over all the normal modes:

$$\psi(x) = \sum_{n=1}^{\infty} c_n \left[\sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right) \right]. \tag{7.1}$$

This states that *any* wiggle of a guitar string can be written as a sum of the harmonic modes weighted by a number c_n. We can write the wavefunction as a superposition of any complete basis set of functions $\phi_n(x)$

$$\psi(x) = \sum_{n=1}^{\infty} c_n \phi_n(x). \tag{7.2}$$

We just used the particle-in-a-box eigenmodes $\phi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right)$ as an example. This is a *Fourier expansion*. The functions $\phi_n(x)$ for a complete orthonormal basis from which we can construct any continuous function.

$$\int_0^L \phi_m(x)\phi_n(x)\mathrm{d}x = \delta_{mn} = \begin{cases} 0, & m \neq n, \\ 1, & m = n. \end{cases} \tag{7.3}$$

We can turn the sum to an integral to get the discrete Fourier series. Figure 20 shows an example of this construction with five sinusoidal plane waves made using `matplotlib`.

In the continuum limit, this proceeds analogously. We can consider free plane wave with well-defined energy $E_k = \hbar^2 k^2/2m$ with eigenstate $\phi(k)$

$$\phi(k) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{\mathrm{i}kx}. \tag{7.4}$$

These states are orthogonal such that they satisfy the Dirac delta function

$$\int_{-\infty}^{\infty} \phi^*(k')\phi(k)\mathrm{d}x = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}(k-k')x}\mathrm{d}x = \delta(k-k') = \begin{cases} 0, & k \neq k', \\ 1, & k = k'. \end{cases} . \tag{7.5}$$

We can expand any function into a basis using $\phi(k)$ and turn the discrete sum into an integral in the with summation index replaced with $\mathrm{d}k$

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{\psi}(k) \mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}k. \tag{7.6}$$

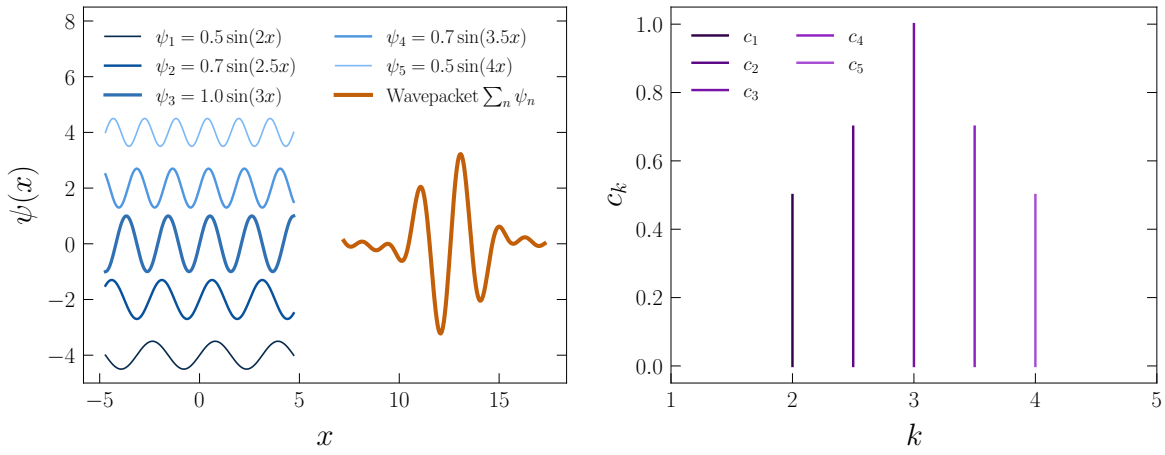


Figure 20: Left: constructing a localized wavepacket (thick orange) in `matplotlib` by summing five plane waves $\psi_n(x) = c_k \sin(k_n x)$ (blue). Right: the Fourier modes in k space.

Here, we promoted the discrete coefficients c_n labelled by integers n into a continuous function $\tilde{\psi}(k)$ labelled by real numbers k. This is a Fourier transform, whose inverse is

$$\tilde{\psi}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(x) \mathrm{e}^{-\mathrm{i}kx} \, \mathrm{d}k. \tag{7.7}$$

7.2 Gaussian wavepacket

We return to the free particle. The superposition principle implies linear combinations of plane waves $\mathrm{e}^{\mathrm{i}kx}$ are also solutions to the Schrödinger equation. While plane waves are non-normalizable, we can leverage this principle to build a Gaussian wavepacket that *is* normalizable. This is the closest description in quantum mechanics of a quasi-localized free particle wavefunction. We can centre around a particular momentum $p_0 = \hbar k_0$ with width σ

$$\tilde{\psi}(k) = \exp\left[-\frac{(k-k_0)^2}{2\sigma^2}\right]. \tag{7.8}$$

Figure 20 simplistically shows this construction for five discrete values of k, which you can imagine taking the continuum limit. We construct a spatial-dependent wavefunction as a sum over plane-wave states $\mathrm{e}^{\mathrm{i}kx}$ weighted by the Gaussian in momentum space

$$\psi(x) = \int_{-\infty}^{\infty} \tilde{\psi}(k) \mathrm{e}^{\mathrm{i}kx} \, \mathrm{d}k = \int_{-\infty}^{\infty} \exp\left[-\frac{(k-k_0)^2}{2\sigma^2} + \mathrm{i}kx\right] \mathrm{d}k. \tag{7.9}$$

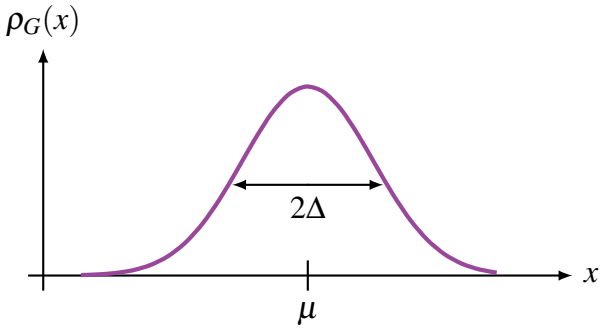


Figure 21: Standard form of a Gaussian normal probability distribution $\rho_G(x)$.

We could consider the full time-dependent wavefunction $\psi(x,t)$ by including the $\mathrm{e}^{-\mathrm{i}E_k t/\hbar} = \mathrm{e}^{-\mathrm{i}\hbar k^2 t/2m}$ piece. We can organize the argument of the exponent into a quadratic form

$$[\ldots] = -\frac{k^2 - 2kk_0 + k_0^2}{2\sigma^2} + \mathrm{i}kx = -\frac{1}{2\sigma^2}k^2 + \left(\frac{k_0}{\sigma^2} + \mathrm{i}x\right)k - \frac{k_0^2}{2\sigma^2}. \tag{7.10}$$

Next complete the square using a formula that we learnt in high school algebra

$$-ak^2 + bk + c = -a\left(k - \frac{b}{2a}\right)^2 + \left(c + \frac{b^2}{4a}\right), \tag{7.11}$$

where we identify

$$a = \frac{1}{2\sigma^2}, \quad b = \frac{k_0}{\sigma^2} + \mathrm{i}x, \quad c = -\frac{k_0^2}{2\sigma^2}. \tag{7.12}$$

Then we factorize out the k-independent pieces to write the wavefunction as

$$\psi(x,t) = \exp\left(c + \frac{b^2}{4a}\right)\int_{-\infty}^{\infty}\exp\left[-a\left(k - \frac{b}{2a}\right)^2\right]\mathrm{d}k. \tag{7.13}$$

We recognize the standard Gaussian integral

$$\int_{-\infty}^{\infty}\mathrm{e}^{-a(k-\beta)^2}\,\mathrm{d}k = \sqrt{\frac{\pi}{a}}. \tag{7.14}$$

to find the general wavefunction is

$$\psi(x) = \mathrm{e}^{-k_0^2/2\sigma^2}\sqrt{\frac{\pi}{a}}\exp\left[-\frac{1}{4a}\left(x - \frac{\mathrm{i}k_0}{\sigma^2}\right)^2\right]. \tag{7.15}$$

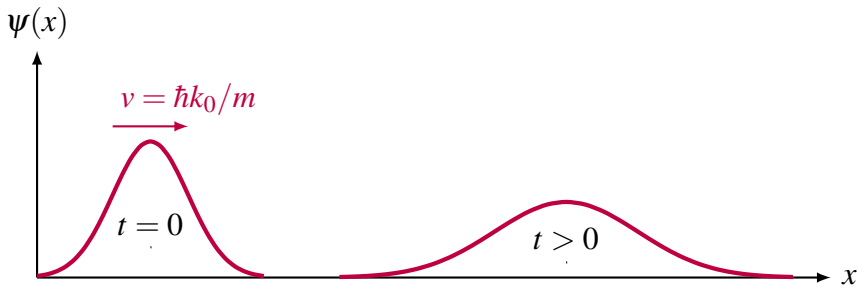


Figure 22: Sketch of Gaussian wavepacket moving in free space with speed $v = \hbar k_0/m$, whose dispersion (width) and amplitude (height) broadens with time.

7.3 Uncertainty principle

Recall the standard Gaussian (normal) probability distribution from statistics textbooks is

$$\rho_G(x) = \frac{1}{\sqrt{2\pi\Delta}} \exp\left(-\frac{(x-\mu)^2}{2\Delta^2}\right), \tag{7.16}$$

with mean μ and standard deviation Δ. Figure 21 sketches this. Rather than explicitly computing the variances, we shall compare with this standard form to "read off" the variance. The Gaussian wavepacket has a statistical variance given by the probability density

$$\rho(x) = |\psi(x)|^2, \quad \rho(k) = |\tilde{\psi}(k)|^2. \tag{7.17}$$

We then read off the standard deviation in position as

$$2\Delta x^2 = 2a^2(t=0) = \frac{1}{\sigma^2} \quad \Rightarrow \quad \Delta x = \frac{1}{\sqrt{2}\sigma} \tag{7.18}$$

The standard deviation in wavenumber is similarly

$$2\Delta k^2 = \sigma^2 \quad \Rightarrow \quad \Delta k = \frac{\sigma}{\sqrt{2}} \tag{7.19}$$

Given $p = \hbar k$, we find the product of position and momentum uncertainties of a quantum particle cannot both be known to arbitrary precision, a statement of the *uncertainty principle*

$$\Delta x \Delta p = \frac{\hbar}{2}. \tag{7.20}$$

7.4 Time evolution

When accounting for the full time-dependent solutions, a becomes a time-dependent function

$$a \to a(t) = \frac{1}{2}\left(\frac{1}{\sigma^2} + \frac{i\hbar t}{m}\right). \tag{7.21}$$

To understand the probability distribution, we must use $P(x,t) = \Psi^*(x,t)\Psi(x,t) = |\Psi(x,t)|^2$.

$$P(x,t) = |\Psi(x,t)|^2 = \mathrm{e}^{-k_0^2/\sigma^2} \frac{\pi}{|a(t)|} \exp\left[-\frac{1}{4a(t)} \left(x - \frac{\mathrm{i}k_0}{\sigma^2} \right)^2 - \frac{1}{4a^*(t)} \left(x + \frac{\mathrm{i}k_0}{\sigma^2} \right)^2 \right]. \tag{7.22}$$

Some algebraic gymnastics turns this into a more physically insightful form

$$P(x,t) = \mathrm{e}^{-k_0^2/\sigma^2} \frac{\pi}{|a(t)|} \exp\left[-\frac{1}{2\sigma^2 |a(t)|^2} (x - vt)^2 \right]. \tag{7.23}$$

First, the probability distribution has the form $f(x-vt)$, which is a wave travelling in the positive x direction with speed

$$v = \frac{\hbar k_0}{m}. \tag{7.24}$$

Second, the width of the probability distribution grows as a function of time

$$|a(t)|^2 = \frac{1}{4} \left(\frac{1}{\sigma^2} + \frac{\hbar^2 t^2}{m^2} \right). \tag{7.25}$$

Figure 22 shows a sketch of a Gaussian wavepacket moving in space. As time increases, the width broadens such that it satisfies the inequality

$$\Delta x \Delta p \geq \frac{\hbar}{2}. \tag{7.26}$$

The expression (7.20) at $t = 0$ therefore denotes the minimum uncertainty.

This uncertainty principle is also generic to classical waves. It appears in quantum mechanics due to wave-particle duality with profound consequences. The more precise a quantum particle's position, the less precision we have on its momentum and vice versa. The uncertainty principle is a fundamental feature of quantum theory and *not* due to instrument precision. This strikingly contrasts with a classical particle, where only experiment not theory restricts the precision of position and momentum. This is our first look at the Heisenberg uncertainty principle, which section 13.1 proves in generality.

HW4: Bound states and wavepacket (set Sep 24, due Oct 1)

1. Consider the finite-well potential in section 6.4 with potential

$$V(x) = \begin{cases} -V_0, & |x| < a, \\ 0, & |x| > a. \end{cases} \tag{7.27}$$

   (a) Explain the concept of parity, why the Hamiltonian with the above $V(x)$ has parity symmetry, and what this implies for wavefunction $\psi(x)$ solutions.

   (b) Consider bound states such that $-V_0 < E < 0$. Write down the time-independent Schrödinger equation for the three regions: $x < -a, -a < x < a, x > a$.

   (c) What functions (e.g. exponential, trigonometric functions) do the odd-parity wavefunction solutions take in the three regions $x < -a, -a < x < a, x > a$? What expressions do the kinetic energy take in each region? Sketch the even and odd parity wavefunctions with the lowest energy states in relation to the potential (7.27).

   (d) Show the steps to derive the condition that odd-parity wavefunctions must satisfy:

$$k\cot(ka) = -\kappa. \tag{7.28}$$

   State the expressions for k and κ. Find the quadratic relation between k, κ, V_0 by eliminating E from Schrödinger's equation. Sketch this quadratic relation and $k\cot(ka) = -\kappa$, then explain qualitatively how to find the solutions graphically.

2. Consider a particle wavefunction $\psi(x)$ taking the form of a Gaussian wavepacket

$$\psi(x) = \frac{1}{(2\pi\sigma)^{1/4}} \exp\left(-\frac{x^2}{4\sigma^2}\right). \tag{7.29}$$

   (a) What is the wavefunction normalization condition? Show that this wavefunction is normalized with the help of the standard Gaussian integral identity

$$\int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{e}^{-(a^2x^2+bx)} = \frac{\sqrt{\pi}}{a} \mathrm{e}^{b^2/4a^2}. \tag{7.30}$$

   (b) Explain qualitatively how the Fourier transform from position $\psi(x)$ to momentum $\psi(p)$ space, $\psi(p) = \frac{1}{\sqrt{h}} \int \mathrm{d}x\, \mathrm{e}^{-\mathrm{i}px/\hbar} \psi(x)$, is analogous to a discrete Fourier series. What is the physical interpretation in terms of quantum superposition?

   (c) Perform the Fourier transform integral to show that the wavefunction $\psi(p)$ and its probability distribution $|\psi(p)|^2$ are also Gaussians.

   (d) How are the standard deviations of the wavepacket in position and momentum spaces related mathematically? Comment on its physical significance.

8 Scattering and tunnelling

In classical mechanics, when we throw a basketball at a brick wall or the cliff edge of the Hudson River Palisades, Newton's laws makes the ball either bounce right back or go over it depending on how much kinetic energy we impart. The basketball certainly does not do both at once. In quantum mechanics with Schrödinger's equation, when we throw objects with wave properties onto analogous barriers, the scattered particles seem to do both!

We can consider $\Psi(x,t) = \phi(t)\psi(x)$ as a plane wave

$$\Psi(x,t) = A\mathrm{e}^{-\mathrm{i}(\omega t - kx)}. \tag{8.1}$$

Given this is formally not normalizable, we instead think of this as describing a continuous particle beam with a probability density

$$\rho(x,t) = |\Psi(x,t)|^2 = |A|^2. \tag{8.2}$$

Then the probability current is given by equation (4.15)

$$j(x,t) = -\frac{\mathrm{i}\hbar}{2m}\left(\Psi^*\frac{\partial\Psi}{\partial x} - \Psi\frac{\partial\Psi^*}{\partial x}\right) = -\frac{\mathrm{i}\hbar}{2m}\left(|A|^2 2\mathrm{i}k\right) = |A|^2\frac{p}{m} = \rho v. \tag{8.3}$$

We interpret this as a beam with flux equal to its particle density ρ multiplied by velocity v.

8.1 Scattering off cliff edge

Consider the cliff edge modelled with a Heaviside step potential

$$V(x) = \begin{cases} 0, & x < 0, \\ V_0, & x > 0. \end{cases} \tag{8.4}$$

Mathematically, this is the integral of the Dirac delta function. We inject a quantum particle with unit normalization. The wavefunction takes the following form in the two regions

$$\psi(x) = \begin{cases} \psi_- = \mathrm{e}^{\mathrm{i}kx} + r\mathrm{e}^{-\mathrm{i}kx}, & x < 0, \\ \psi_+ = t\mathrm{e}^{\mathrm{i}Kx}, & x > 0. \end{cases} \tag{8.5}$$

The scattered waves have coefficients r,t to be determined. Figure 23 shows a sketch of this problem. The two wavenumbers are given by Schrödinger's equation

$$k = \sqrt{\frac{2mE}{\hbar^2}}, \quad K = \sqrt{\frac{2m(E-V_0)}{\hbar^2}}. \tag{8.6}$$

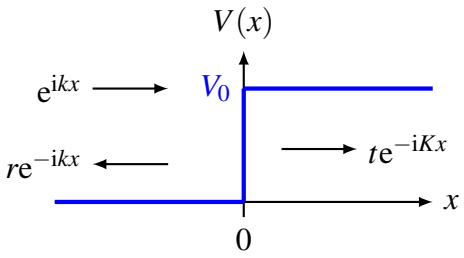


Figure 23: Scattering off a cliff-edge step potential.

We consider the case $E > V_0$. Applying the continuity of wavefunction and its first derivative using equations (8.5) imposes these conditions

$$\psi_-(x=0) = \psi_+(x=0) \quad \Rightarrow \quad 1 = r + t, \tag{8.7}$$

$$\left.\frac{\partial \psi_-}{\partial x}\right|_{x=0} = \left.\frac{\partial \psi_+}{\partial x}\right|_{x=0} \quad \Rightarrow \quad \mathrm{i}k(1-r) = \mathrm{i}Kt. \tag{8.8}$$

We can then solve for the unknown coefficients

$$r = \frac{k-K}{k+K}, \quad t = \frac{2k}{k+K}, \tag{8.9}$$

which we call the reflection r and transmission t amplitudes.

Given we physically observe $|\Psi(x,t)|^2$, we should relate this to probability currents $j(x,t)$ by applying equation (8.3). The incident probability current is then

$$j_{\text{incident}} = \frac{\hbar k}{m}. \tag{8.10}$$

Similarly, the probability currents reflected and transmitted by the cliff edge are

$$j_{\text{reflected}} = r\frac{\hbar k}{m} = \frac{\hbar k}{m}\left(\frac{k-K}{k+K}\right), \tag{8.11}$$

$$j_{\text{transmitted}} = t\frac{\hbar K}{m} = \frac{\hbar K}{m}\left(\frac{2k}{k+K}\right). \tag{8.12}$$

We conventionally define the reflectivity R and transmittivity T coefficients as the reflected and transmitted probability current densities (4.15) relative to the incident current

$$R = \left|\frac{j_{\text{reflected}}}{j_{\text{incident}}}\right|, \quad T = \left|\frac{j_{\text{transmitted}}}{j_{\text{incident}}}\right|. \tag{8.13}$$

Implementing the fluxes, we find the reflectivity R and transmittivity T are

$$R = \left(\frac{k-K}{k+K}\right)^2, \quad T = \frac{4kK}{(k+K)^2}. \tag{8.14}$$

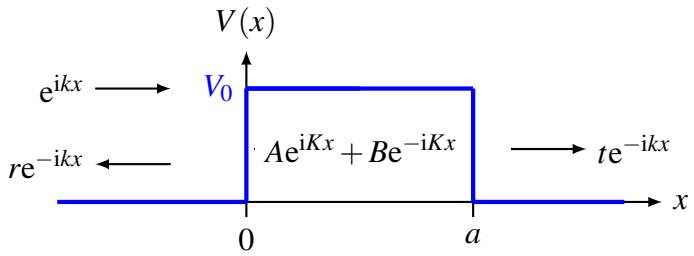


Figure 24: Scattering off a thick-wall finite potential.

By conservation of probability, we must have $R+T=1$. Both $r\neq 0$ and $t\neq 0$ are non-zero, which indicates the particle has both reflected and transmitted wavefunction! This implies non-zero probabilities of being in both $x<0$ and $x>0$ regions after scattering.

Contrast this with classical physics. If we throw a ball with insufficient vertical kinetic energy to go over the cliff with height h, $\frac{1}{2}mv^2 < mgh$, it will only bounce back. If we throw the ball with enough kinetic energy $\frac{1}{2}mv^2 > mgh$ to go above the cliff, it will only go over. The quantum wavefunction $\psi(x)$ *both* bounces back and goes over the cliff after scattering. Of course, once we apply Born's rule $|\psi(x)|^2$, we only observe the electron in one location.

8.2 Tunnelling through a wall

Now consider throwing a quantum particle at a thick wall. Figure 24 sketches a thick-wall potential with fixed width a and finite height V_0:

$$V(x)=\begin{cases} V_0, & 0<x<a,\\ 0, & \text{otherwise.}\end{cases} \tag{8.15}$$

We throw an electron with initial wavefunction $\mathrm{e}^{\mathrm{i}kx}$ and seek the resulting wavefunctions in these three regions

$$\psi(x)=\begin{cases} \psi_1(x)=\mathrm{e}^{\mathrm{i}kx}+r\mathrm{e}^{-\mathrm{i}kx} & x<0,\\ \psi_2(x)=A\mathrm{e}^{\mathrm{i}Kx}+B\mathrm{e}^{-\mathrm{i}Kx} & 0<x<a,\\ \psi_3(x)=t\mathrm{e}^{\mathrm{i}kx} & x>a.\end{cases} \tag{8.16}$$

The two values of the wavenumber correspond to

$$k=\sqrt{\frac{2mE}{\hbar^2}},\quad K=\sqrt{\frac{2m(E-V_0)}{\hbar^2}}. \tag{8.17}$$

We can see there are two cases for K in $0<x<a$:

- $E > V_0$ kinetic energy greater than potential barrier, implying K is real and corresponds to free plane-wave states;
- $E < V_0$ kinetic energy less than potential barrier, implying K is imaginary so our trial solutions can be written exponential states

$$\psi_2(x) = Ae^{\kappa x} + Be^{-\kappa x}, \quad \kappa = -iK. \tag{8.18}$$

Consider the case of sub-critical energy where the kinetic energy is less than the potential barrier $E < V_0$. We use the exponential rather than oscillatory form in equation (8.18). Next apply the boundary conditions that the wavefunction and its first derivative are both continuous at $x = 0$ and $x = a$. We find the following equations at $x = 0$

$$\psi_1(0) = \psi_2(0) \quad \Rightarrow \quad A + B = 1 + r, \tag{8.19}$$

$$\left.\frac{\partial \psi_1}{\partial x}\right|_{x=0} = \left.\frac{\partial \psi_2}{\partial x}\right|_{x=0} \quad \Rightarrow \quad \frac{\kappa}{k}(A - B) = 1 - r. \tag{8.20}$$

Next apply the boundary conditions to $x = a$

$$\psi_2(a) = \psi_3(a) \quad \Rightarrow \quad te^{ika} = Ae^{\kappa a} + Be^{-\kappa a}, \tag{8.21}$$

$$\left.\frac{\partial \psi_2}{\partial x}\right|_{x=a} = \left.\frac{\partial \psi_3}{\partial x}\right|_{x=a} \quad \Rightarrow \quad \frac{ik}{\kappa} te^{ika} = Ae^{\kappa a} - Be^{-\kappa a}. \tag{8.22}$$

This comprises a set of four linear simultaneous equations for four unknowns r, t, A, B. Perfect, we can solve for all the unknown coefficients. We can add equations (8.19) and (8.20) to eliminate r, then add equations (8.21) and (8.22) to eliminate B. Use the remaining two equations to eliminate A. Applying identities for hyperbolic functions $\cosh(x) = (e^x + e^{-x})/2, \sinh(x) = (e^x - e^{-x})/2$ and some algebra, we obtain the transmission coefficient

$$t = \frac{2e^{-2ika}}{2\cosh(\kappa a) - i\left(\frac{k}{\kappa} - \frac{\kappa}{k}\right)\sinh(\kappa a)}. \tag{8.23}$$

We take the complex conjugate and modulus square to obtain the transmittivity

$$T = t^* t = \frac{1}{1 + \frac{1}{4}\left(\frac{k^2 + \kappa^2}{k\kappa}\right)^2 \sinh^2(\kappa a)}. \tag{8.24}$$

This uses the hyperbolic identity $\cosh^2(x) - \sinh^2(x) = 1$.

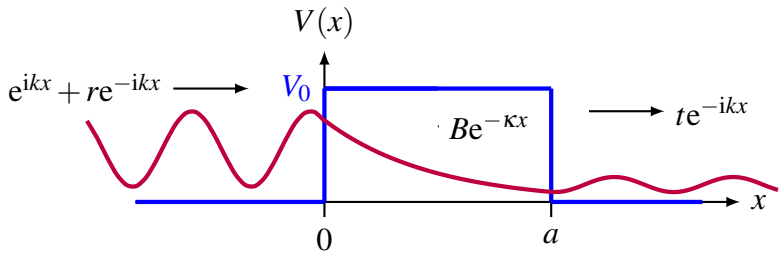


Figure 25: Sketch of wavefunctions (red) for thick-wall scattering with $E < V_0$.

Let us define two dimensionless parameters, namely the energy fraction f and the normalized potential width g

$$f = \frac{E}{V_0}, \quad g = \frac{2mV_0a^2}{\hbar^2}. \tag{8.25}$$

Using this, we can rewrite the transmittivity as

$$T = \left(1 + \frac{\sinh^2[\sqrt{g^2(1-f)}]}{4f(1-f)}\right)^{-1}. \tag{8.26}$$

In the weak transmission limit, the potential is wide and high $g \gg 1$ compared to the kinetic energy $f \ll 1$. As $x \to \infty$, $\sinh(x) \to e^x$ so the transmittivity becomes an exponential decay

$$T \approx 16fe^{-2g}. \tag{8.27}$$

This non-zero transmittivity is called tunnelling.

Figure 25 shows a sketch of the wavefunction tunnelling through the thick wall. There is non-zero probability of finding the particle in the classically forbidden region. The probability of tunnelling decreases with increasing V_0 and a, while the reflectivity is $R = 1 - T$. Tunnelling finds a brilliant application in *scanning tunnelling microscopy* invented by Gerd Binnig and Heinrich Rohrer at IBM Zürich in 1981 [30]. This technique uses quantum tunnelling to resolve features with 0.1 nm resolution, enabling the remarkable "imaging" of atoms (figure 26)

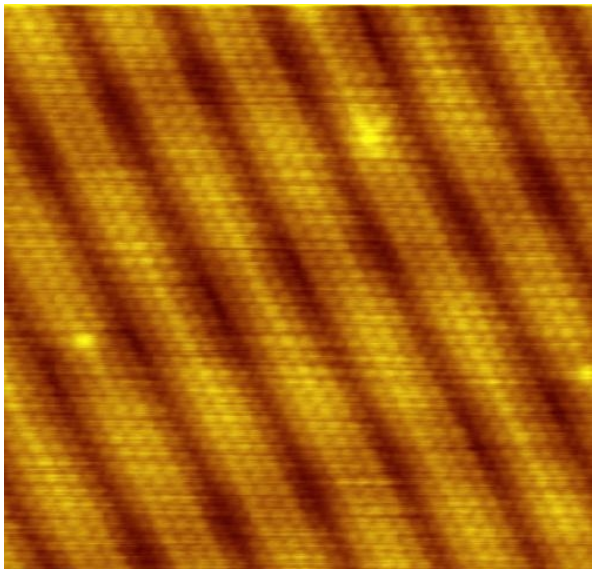


Figure 26: Lattice of individual gold atoms reconstructed by scanning tunnelling microscopy. Image: Erwin Rossen.

III Operator formalism

9 States and Hilbert space

We now generalize what we have studied to the theoretical formalism of quantum mechanics. This part of the course has an unavoidably mathematical flavour at the level of Mathematical Physics (PHYS-UA 106) and Linear Algebra (MATH-UA 140). We endeavour to motivate and illustrate mathematical ideas relevant to quantum physics.

9.1 Dirac notation

Dirac notation denotes the state of quantum system as a vector $|\psi\rangle$. This angle bracket notation is called a *ket vector*. This is a coordinate-free notation, just like $\mathbf{r}$ represents a position vector without explicitly specifying Cartesian, spherical or other coordinates. While a position vector $\mathbf{r}$ inhabits a real vector space, a quantum state $|\psi\rangle$ inhabits a complex vector space called a **Hilbert space**. For an N-dimensional Hilbert space, we can expand $|\psi\rangle$ as a column vector with complex scalars a_i and basis state vectors $|e_i\rangle$

$$|\psi\rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix} = \sum_n a_n |e_n\rangle \tag{9.1}$$

We have seen this structure of expanding vectors in terms of basis vectors: (i) position vectors $\mathbf{r}$ as sum over Cartesian basis vectors $\mathbf{e}_i \in \{\mathbf{i},\mathbf{j},\mathbf{k}\}$ in a Euclidean real vector space, (ii) writing $\psi(x)$ as a sum over sinusoidal functions as a basis of Fourier modes $\{\phi_n\}$ in equation (7.1)

$$\mathbf{r} = \sum_{i=1}^{3} x_i \mathbf{e}_i = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}, \tag{9.2}$$

$$\psi(x) = \sum_{n=1}^{\infty} c_n \phi_n = \sum_{n=1}^{\infty} c_n \left[\sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right) \right]. \tag{9.3}$$

Note the similarity in mathematical structure. The functions $\{\phi_n(x)\}$ form a basis but in an abstract Hilbert space that only lacks the geometric picture of a Cartesian basis $\{\mathbf{i},\mathbf{j},\mathbf{k}\}$. What matters is that $|e_i\rangle$ obeys the rules of vector spaces, not if they have geometric interpretations.

Given we have complex vectors, we also define the dual (adjoint or Hermitian conjugate) of a state as a *bra vector* $\langle\psi|$, which occupies the *adjoint space*. In vector notation, the adjoint (†) takes the complex conjugate ($*$) and transposes (T) a column vector into a row vector

$$\langle\psi| \equiv |\psi\rangle^\dagger = |\psi\rangle^{*\mathrm{T}} = \begin{pmatrix} a_1^* & a_2^* & \dots & a_N^* \end{pmatrix} = \sum_n a_n^* \langle e_n|. \tag{9.4}$$

This forms the *Dirac bra-ket notation*, where the inner product is written as

$$\langle \text{bra}|\text{ket}\rangle. \tag{9.5}$$

Paul Dirac introduced this notation in 1939 [31], who systematized the formalism of linear operators for quantum mechanics in 1930 [32].

9.2 Inner product

Recall the length of a vector $\mathbf{r}$ on a plane is the scalar (or dot) product or with itself

$$\text{length}^2 = |\mathbf{r}|^2 = \mathbf{r}\cdot\mathbf{r} = \mathbf{r}^\mathrm{T}\mathbf{r} = \sum_i r_i r_i = x^2 + y^2. \tag{9.6}$$

We also represent complex numbers in a two-dimensional complex plane $c = a + \mathrm{i}b$ and define the modulus via multiplication with its complex conjugate

$$\text{modulus}^2 = |c|^2 = c^* c = (a - \mathrm{i}b)(a + \mathrm{i}b) = |a|^2 + |b|^2. \tag{9.7}$$

We generalize these magnitudes to N-dimensional complex vectors as the dual vector (transposed row vector plus complex conjugate) multiplied into a column vector

$$\langle\psi|\psi\rangle = \sum_i a_i^* a_i \langle e_i|e_i\rangle = \begin{pmatrix} a_1^* & a_2^* & \dots & a_N^* \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N^* \end{pmatrix} = a_1^* a_1 + a_2^* a_2 + \dots a_N^* a_N. \tag{9.8}$$

We call this the inner product of a vector space, which we apply to quantum mechanics:

- We find one of the complex probability amplitudes a_i by performing an inner product of the state $|\psi\rangle$ with one of the basis vectors $|e_i\rangle$ as shown in figure 27

$$a_i = \langle e_i|\psi\rangle. \tag{9.9}$$

- Like Cartesian basis vectors ($\mathbf{i}\cdot\mathbf{i}=1, \mathbf{i}\cdot\mathbf{j}=0$), the basis state vectors are orthonormal

$$\langle e_i|e_j\rangle = \delta_{ij} = \begin{cases} 1 & i=j, \\ 0 & i\neq j. \end{cases} \tag{9.10}$$

- Wavefunction normalizability requires that the inner product be finite

$$|\langle\psi|\psi\rangle|^2 < \infty. \tag{9.11}$$

- The inner product is non-negative i.e. positive definite

$$\langle\psi|\psi\rangle \geq 0. \tag{9.12}$$

This generalizes to infinite-dimensional continuous functions in a Hilbert space. We already saw this with wavefunction normalization by performing the integral

$$|\psi|^2 = \langle\psi|\psi\rangle = \int_{-\infty}^{\infty} \psi^*(x)\psi(x)\,\mathrm{d}x. \tag{9.13}$$

This is an infinite-dimensional dot product where we replaced the finite labels of a vector $i=1,2,\ldots,N$ with the infinite set of real numbers $x\in\mathbf{R}$ that the wavefunction is defined over $\psi(x)$. We have already expanded functions as a Fourier series, where we perform integrals with sinusoidal functions $\phi_n(x)=\sin(n\pi x/L)$ of the infinite square well

$$|\phi_n|^2 = \langle\phi_n|\phi_n\rangle = \int_0^L \frac{2}{L}\sin^2\left(\frac{n\pi x}{L}\right)\,\mathrm{d}x. \tag{9.14}$$

We can extend these concepts to two different state vectors. Recall for two different vectors $\mathbf{a},\mathbf{b}$, we can find the "overlap" via the scalar dot product

$$\mathbf{a}\cdot\mathbf{b} = |a||b|\cos\theta = \sum_i a_i b_i. \tag{9.15}$$

Geometrically in real space, this is the projection of vector $\mathbf{a}$ onto $\mathbf{b}$. If $\mathbf{a}$ is at a right angle to $\mathbf{b}$, then $\theta=\pi/2$ and there is no overlap so the dot product is zero. If $\mathbf{a}$ is parallel to $\mathbf{b}$, then $\theta=0$ and the overlap is maximal so the dot product is the product of their lengths $|\mathbf{a}||\mathbf{b}|$. This statement is formalized as the Cauchy-Schwartz inequality

$$\mathbf{a}\cdot\mathbf{b} \leq |\mathbf{a}||\mathbf{b}|. \tag{9.16}$$

This is none other than a fancy statement of the triangle inequality from elementary geometry.

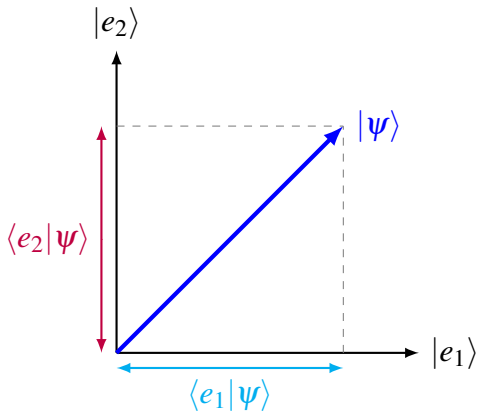


Figure 27: A state vector wavefunction $|\psi\rangle$ projected into basis vectors $|e_i\rangle$ in a two-dimensional Hilbert space.

We generalize these concepts to states in a Hilbert space. Consider another state vector

$$|\phi\rangle = \sum_i b_i |e_i\rangle. \tag{9.17}$$

We can perform the inner product with $|\psi\rangle$ in an N-dimensional Hilbert space using Dirac notation. This is multiplying a row vector with the complex conjugate of a column vector

$$\langle\psi|\phi\rangle = \sum_i a_i^* b_i \langle e_i|e_i\rangle = \begin{pmatrix} a_1^* & a_2^* & \dots & a_N^* \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N^* \end{pmatrix} = a_1^* b_1 + a_2^* b_2 + \dots a_N^* b_N. \tag{9.18}$$

For infinite-dimensional wavefunctions $\psi(x)$, the sum turns into an integral over x

$$\langle\psi|\phi\rangle = \int_{-\infty}^{\infty} \psi^*(x)\phi(x)\,\mathrm{d}x. \tag{9.19}$$

These inner products are generally complex numbers denoting a probability amplitude. They have a few key properties:

- The inner product is linear: if $|\phi\rangle = \alpha|\phi_1\rangle + \beta|\phi_2\rangle$ for α, β complex numbers, then

$$\langle\psi|\phi\rangle = \langle\psi|\alpha\phi_1 + \beta\phi_2\rangle = \alpha\langle\psi|\phi_1\rangle + \beta\langle\psi|\phi_2\rangle. \tag{9.20}$$

- Complex conjugation flips the bra-ket

$$\langle\psi|\phi\rangle^* = \langle\phi|\psi\rangle. \tag{9.21}$$

- We generalize the Cauchy-Schwartz inequality to state vectors as

$$\langle\psi|\phi\rangle \leq \sqrt{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}. \tag{9.22}$$

- The orthogonality of two states is stated via the Kronecker delta

$$\langle\psi_i|\psi_j\rangle = \delta_{ij}. \tag{9.23}$$

9.3 Position representation

We introduced the spatial wavefunction early in this course to pedagogically build intuition. However, we were unwittingly carrying around the baggage of the infinite-dimensional position representation of a state vector

$$|\psi\rangle = \int_{-\infty}^{\infty} \psi(x)|x\rangle \quad \longleftrightarrow \quad |\psi\rangle = \sum_i a_i|e_i\rangle. \tag{9.24}$$

We describe a spatial wavefunction $\psi(x)$ in Dirac formalism:

- Every infinitesimal spatial position $|x\rangle$ forms a basis set of states analogous to $|e_i\rangle$. Every value of x on the real number line labels a different basis vector.
- The integral $\int \mathrm{d}x$ over all position basis states $|x\rangle$ weighted by the wavefunction $\psi(x)$ is analogous to the expansion as a sum $\sum_i$ over basis vectors.
- Each infinitesimal value of x is associated with probability amplitude $\psi(x)$ of finding the particle there, analogous to the complex elements a_i. The spatial wavefunction $\psi(x)$ is the probability amplitude and given by the inner product

$$\langle x|\psi\rangle = \psi(x) \quad \longleftrightarrow \quad \langle e_i|\psi\rangle = a_i. \tag{9.25}$$

Orthogonality is given by the Dirac delta function analogous to the Kronecker delta

$$\langle x'|x\rangle = \delta(x-x') = \begin{cases} 1, & x = x', \\ 0, & x \neq x'. \end{cases} \quad \longleftrightarrow \quad \langle e_i|e_j\rangle = \delta_{ij}. \tag{9.26}$$

This has the desired property of picking out a wavefunction at a specific position

$$\langle x'|\psi\rangle = \int_{-\infty}^{\infty} \langle x'|x\rangle\psi(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} \delta(x-x')\psi(x)\,\mathrm{d}x = \psi(x'). \tag{9.27}$$

This distribution has the property that its integral over all space is unity

$$\int_{-\infty}^{\infty} \delta(x-x')\,\mathrm{d}x = 1 \tag{9.28}$$

HW5: Scattering and Dirac notation (set Oct 1, due Oct 8)

1. Consider the scattering problem through a thick-wall potential

$$V(x) = \begin{cases} V_0, & 0 < x < a, \\ 0, & \text{otherwise.} \end{cases} \tag{9.29}$$

   (a) Write down the Schrödinger equation and the functional forms of the wavefunctions in the three regions $x < 0, 0 < x < a, x > a$ for (i) $E < V_0$ and (ii) $E > V_0$.

   (b) For $E > V_0$, apply appropriate boundary conditions and the definitions of probability currents j to derive the transmittivity $T = |j_{\text{transmitted}}/j_{\text{incident}}|$

$$T = \frac{1}{1 + \frac{1}{4}\left(\frac{k}{K} - \frac{K}{k}\right)^2 \sin^2(Ka)}. \tag{9.30}$$

   (c) What is the condition on K for $T = 1$ and thus reflectivity $R = 0$ to vanish?

2. A vector space V has position vectors $\mathbf{r} = (x, y, z)^{\mathrm{T}}$. How do we write $\mathbf{r}$ in terms of basis vectors. How do we write a scalar product with another vector $\mathbf{r} \cdot \mathbf{r}'$ in Dirac notation? What is meant by a Hilbert space and adjoint space?

3. Given a state $|\psi\rangle = \mathrm{e}^{\mathrm{i}\pi/6}|a\rangle + (\mathrm{i}/2)|b\rangle$, what is its dual $\langle\psi|$?

4. How do we expand a state $|\psi\rangle$ with basis states $|n\rangle$ and coefficients a_n? In quantum mechanics, what property must $\langle\psi|\psi\rangle$ satisfy and what is the significance of a_n?

5. How do we mathematically define the inner product $\langle\psi|\phi\rangle$ for finite-dimensional vectors and infinite-dimensional functions?

6. Like the position representation, we can write a state $|\psi\rangle$ in momentum representation

$$|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}p\, \psi(p)|p\rangle. \tag{9.31}$$

   (a) Explain each object $|\psi\rangle$, $\int_{-\infty}^{\infty}\mathrm{d}p$, $\psi(p)$ and $|p\rangle$ in the language of probability amplitudes and basis vectors of a Hilbert space.

   (b) How do we extract the wavefunction in momentum space $\psi(p)$ from state $|\psi\rangle$?

   (c) How do we write the orthogonality condition for the momentum basis states $|p_1\rangle$ and $|p_2\rangle$?

10 Operators and observables

Classical mechanics specifies the particle *state* with its position x and velocity $\dot{x}$. The quantities $(x,\dot{x})$ are also *observables* that we can measure with rulers and clocks, from which we build fancier observables like energy $\frac{1}{2}m\dot{x}^2$. Classically, we did not pedantically distinguish states and observables. By contrast, a quantum state $|\psi\rangle$ is *not directly observable*

$$\text{Classical state and observable} : (x,\dot{x}), \quad \text{Quantum state} : |\psi\rangle. \tag{10.1}$$

So how do we define experimental observables of a quantum state? The answer is **operators**.

10.1 Matrix and differential operators

There are two types, or formally *representations*, of operators useful for quantum mechanics:

- **Matrices operating on vectors**: *discrete, finite N-dimensional Hilbert space*. You have seen matrices acting on vectors in algebra classes. For a position vector $\mathbf{r}$ living in real space, we rotate it to a different position $\mathbf{r}'$ in the same space via a rotation matrix $\hat{\mathbf{R}}$

$$\mathbf{r}' = \hat{\mathbf{R}}\cdot\mathbf{r} \quad \leftrightarrow \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{10.2}$$

  The rotation matrix $\hat{\mathbf{R}}$ *operates* on the vector $\mathbf{r}$. In general, an $n\times n$ matrix $\hat{\mathbf{M}}$ acts on an n-dimensional vector $\mathbf{x}' = \hat{\mathbf{M}}\cdot\mathbf{x}$. If the vector $\mathbf{x}' = \lambda_i\mathbf{x}$ is unchanged up to n multiplicative scalars $\{\lambda_i\}$, then we have an eigenvalue equation with eigenvector $\mathbf{x}$

$$\hat{\mathbf{M}}\cdot\mathbf{x} = \lambda_i\mathbf{x}. \tag{10.3}$$

- **Differential operators on functions**: *continuous, infinite-dimensional Hilbert space*. We have seen differential operators acting on exponential states

$$\frac{\partial}{\partial x}\mathrm{e}^{\lambda x} = \lambda\mathrm{e}^{\lambda x}. \tag{10.4}$$

  Here, the first-order derivative operator acts on the function $\phi(x) = \mathrm{e}^{\lambda x}$ to return the same function (so it is an eigenfunction) with continuous set of eigenvalues λ

$$\hat{D}\phi(x) = \lambda\phi(x), \quad \hat{D} = \frac{\partial}{\partial x}. \tag{10.5}$$

  This bears uncanny resemblance to the matrix eigenvalue equation (10.3). This shared structure is no accident. There is a deep connection between differential operators acting on infinite-dimensional functions (functional analysis) and matrices acting on N-dimensional vectors (linear algebra), which underpins quantum mechanics.

Observable	Operator $\hat{Q}$	Position representation	Eigenvalue q_n	Eigenstate $\lvert\phi_n\rangle$
Energy	$\hat{H}$	$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+V$	E_n	V dependent
Momentum	$\hat{p}$	$-\mathrm{i}\hbar\frac{\partial}{\partial x}$	$\hbar k$	$\mathrm{e}^{\mathrm{i}kx}$
Position	$\hat{x}$	x	x_i	$\delta(x-x_i)$

Table 1: Physical observables in quantum operator $\hat{Q}|\phi_n\rangle = q_n|\phi_n\rangle$ formalism.

We generalize these ideas for quantum operators. An operator $\hat{Q}$ transforms a quantum state $|\psi\rangle$ into another quantum state $|\psi'\rangle$ in the same Hilbert space

$$|\psi'\rangle = \hat{Q}|\psi\rangle. \tag{10.6}$$

Quantum operators $\hat{Q},\hat{R}$ are linear and associative, meaning they satisfy

$$\hat{Q}(\alpha|\psi\rangle+\beta|\phi\rangle) = \alpha\hat{Q}|\psi\rangle+\beta\hat{Q}|\phi\rangle \qquad \text{(linear)}, \tag{10.7}$$

$$\hat{Q}\hat{R}|\psi\rangle = \hat{Q}(\hat{R}|\psi\rangle) = (\hat{Q}\hat{R})|\psi\rangle \quad \text{(associative)}, \tag{10.8}$$

where α,β are complex numbers. If an operator $\hat{Q}$ acts on a state $|\phi_n\rangle$ by multiplicative rescaling, this is an eigenvalue equation

$$\hat{Q}|\phi_n\rangle = q_n|\phi_n\rangle. \tag{10.9}$$

The eigenvectors $|\phi_n\rangle$ with $n = 1,2,3\ldots$ are **eigenstates** of the operator $\hat{Q}$. In matrix language, the eigenstates $|\phi_n\rangle$ diagonalize $\hat{Q}$. The corresponding scalar numbers q_n are **eigenvalues** of $\hat{Q}$. The set of eigenvalues $\{q_n\}$ is called the spectrum of the $\hat{Q}$ operator. We translate the mathematical nomenclature of linear algebra into physics

$$\text{Hermitian operator } \hat{Q} \quad \longleftrightarrow \quad \text{Observable} \tag{10.10}$$

$$\text{Eigenvalue spectrum } \{q_n\} \quad \longleftrightarrow \quad \text{Measurable values} \tag{10.11}$$

$$\text{Eigenvectors } |\phi_n\rangle \quad \longleftrightarrow \quad \text{States of system.} \tag{10.12}$$

This dictionary is the heart of quantum mechanics. Quantum mechanics uses a subclass of operators that are *Hermitian*, which guarantee experimentally measured eigenvalues are real numbers. We define Hermitian and prove this fundamental result momentarily.

10.2 Physical observables

Table 1 describes familiar physical observables in this quantum formalism. First, the Hamiltonian operator $\hat{H}$ has an energy eigenvalue E_n for each eigenstate ϕ_n.

$$\hat{H}\phi_n = E_n\phi_n. \tag{10.13}$$

We extensively studied the different forms of ϕ_n for different $V(x)$ in *Wave mechanics*.

Next, the momentum operator $\hat{p} = -\mathrm{i}\hbar\partial/\partial x$ has eigenvalue p

$$\hat{p}\phi = p\phi, \quad \Rightarrow \quad -\mathrm{i}\hbar\frac{\partial\phi}{\partial x} = p\phi. \tag{10.14}$$

This is why we put hats over $\hat{p}$: it prevents such equations descending into tautologies $p\phi = p\phi$. The hat $\hat{p}$ reminds us this is a differential equation, solved by integrating to obtain plane-wave eigenstates $\phi_k = \mathrm{e}^{\mathrm{i}kx}$ and eigenvalues $\hbar k$:

$$\hat{p}\mathrm{e}^{\mathrm{i}kx} = \hbar k\mathrm{e}^{\mathrm{i}kx}. \tag{10.15}$$

As k is a real number, there are an infinite number of eigenvalues. Each specific value of k defines a different eigenstate ϕ_k in an infinite-dimensional Hilbert space.

Finally, the position operator $\hat{x} = x$ acts on a wavefunction by multiplication. A particle can sit at any specific position x_i on the infinite real number line $x \in \mathbf{R}$. This specific position is the eigenvalue x_i corresponding to the eigenvalue equation

$$\hat{x}\psi(x_i) = x_i\psi(x_i), \quad \text{where} \quad \psi(x_i) = \delta(x - x_i). \tag{10.16}$$

The eigenstate $\psi(x_i)$ for eigenvalue x_i is non-zero only when $x = x_i$, else $\psi(x \neq x_i)$ vanishes. We describe this mathematically with the delta function $\delta(x - x_i)$. Integrating both sides checks this solves the eigenvalue equation $\hat{x}\psi(x_i) = x\psi(x_i)$

$$\text{left-hand side :} \quad \hat{x}\psi(x_i) = \int_{-\infty}^{\infty} x\delta(x - x_i)\,\mathrm{d}x = x_i\psi(x_i), \tag{10.17}$$

$$\text{right-hand side :} \quad x_i\psi(x_i) = x_i\int_{-\infty}^{\infty} \delta(x - x_i)\,\mathrm{d}x = x_i\psi(x_i). \tag{10.18}$$

This seems like mathematical pedantry to describe particle position as an eigenvalue equation, but only because the position operator acts by simple multiplication $\hat{x} = x$.

10.3 Hermitian operators

Quantum physics uses a special subset of operators that are *Hermitian*. An N-dimensional matrix $\hat{\mathbf{M}}$ equal to its adjoint is defined as *Hermitian*

$$\hat{\mathbf{M}}^\dagger = \hat{\mathbf{M}}. \tag{10.19}$$

The adjoint † takes the transpose and complex conjugation $\hat{\mathbf{M}}^\dagger = \left(\hat{\mathbf{M}}^*\right)^{\mathrm{T}}$, for example

$$\hat{\mathbf{M}} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \hat{\mathbf{M}}^\dagger = \left(\hat{\mathbf{M}}^*\right)^{\mathrm{T}} = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix}. \tag{10.20}$$

We are familiar with sandwiching a matrix between two vectors $\mathbf{x}, \mathbf{y}$ to obtain a scalar

$$\mathbf{y}^\dagger \cdot \hat{\mathbf{M}} \cdot \mathbf{x} = \begin{pmatrix} y_1^* & y_2^* \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1^* & y_2^* \end{pmatrix} \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix} = \sum_{i=1}^{N} \sum_{j=1}^{N} y_i^* M_{ij} x_j. \tag{10.21}$$

The dots $\cdot$ denote a sums over the N components of the vectors and matrices labelled by indices $\{i, j\}$. Here, M_{ij} is one component of the matrix $\mathbf{M}$ called the *matrix element*.

We extend these ideas using Dirac notation to general quantum operators $\hat{Q}$. Acting on state vectors $|\psi\rangle, |\phi\rangle$, we can write analogously

$$\mathbf{y}^\dagger \cdot \hat{\mathbf{M}} \cdot \mathbf{x} \to \langle \phi | \hat{Q} | \psi \rangle. \tag{10.22}$$

Expanding a state via basis vectors $|\psi\rangle = \sum_i a_i |i\rangle, |\phi\rangle = \sum_j b_j |j\rangle$, the matrix element of $\hat{Q}$ is

$$\langle i | \hat{Q} | j \rangle = \hat{Q}_{ij}. \tag{10.23}$$

Taking the adjoint is complex conjugation and transpose by swapping the indices

$$\hat{Q}^\dagger_{ij} = (\hat{Q}^{\mathrm{T}}_{ij})^* = \hat{Q}^*_{ji}. \tag{10.24}$$

We prove the adjoint of an operator product $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger \hat{A}^\dagger$ by manipulating matrix elements:

$$(\hat{A}\hat{B})^\dagger_{ij} = (\hat{A}\hat{B})^*_{ji} = \sum_k \hat{A}^*_{jk} \hat{B}^*_{ki} = \sum_k \hat{A}^\dagger_{kj} \hat{B}^\dagger_{ik} = \sum_k \hat{B}^\dagger_{ik} \hat{A}^\dagger_{kj} = (\hat{B}\hat{A})^\dagger_{ij}. \tag{10.25}$$

The sum is a trick of inserting an identity matrix $\hat{I}$ called *resolution of identity*

$$\hat{I} = \sum_k |k\rangle\langle k| = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & 1 & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}. \tag{10.26}$$

We generalize from a sum over finite-dimensional vectors and matrices to an integral over continuous variables for infinite-dimensional wavefunctions in a Hilbert space

$$\langle \phi | \hat{Q} | \psi \rangle = \int_{-\infty}^{\infty} \phi^*(x) [\hat{Q}\psi(x)] \, \mathrm{d}x. \tag{10.27}$$

The definition of a Hermitian operator in coordinate-independent Dirac notation is

$$\langle \phi | \hat{Q} | \psi \rangle = \langle \psi | \hat{Q} | \phi \rangle^*. \tag{10.28}$$

Writing the integrals, an operator $\hat{Q}$ acting on wavefunctions $\psi(x), \phi(x)$ is Hermitian if

$$\int_{-\infty}^{\infty} \phi^*(x) [\hat{Q}\psi(x)] dx = \left\{ \int_{-\infty}^{\infty} \psi^*(x) [\hat{Q}\phi(x)] \, \mathrm{d}x \right\}^*. \tag{10.29}$$

10.4 Properties of Hermitian operators

Let $|\phi_n\rangle$ be a normalized eigenstate of Hermitian operator $\hat{Q}$ corresponding to eigenvalues q_n

$$\hat{Q}|\phi_n\rangle = q_n|\phi_n\rangle. \tag{10.30}$$

Multiply both sides from the left by the dual eigenstate with label $\langle\phi_m|$

$$\langle\phi_m|\hat{Q}|\phi_n\rangle = \langle\phi_m|q_n|\phi_n\rangle = q_n\langle\phi_m|\phi_n\rangle \tag{10.31}$$

Repeat this for the Hermitian conjugate of the eigenvalue equation and apply $\hat{Q}^\dagger = \hat{Q}$:

$$\langle\phi_m|\hat{Q}^\dagger = q_m^*\langle\phi_m| \quad \Rightarrow \quad \langle\phi_m|\hat{Q}|\phi_n\rangle = q_m^*\langle\phi_m|\phi_n\rangle. \tag{10.32}$$

Subtracting (10.32) from (10.31) gives

$$\Rightarrow (q_n - q_m^*)\langle\phi_m|\phi_n\rangle = 0 \tag{10.33}$$

From this equation, we identify important properties of Hermitian operators:

- **Eigenvalues are real**. For equation (10.33) to be true when $m = n$, q_n must be real:

  $$q_n = q_n^*. \tag{10.34}$$

  Therefore, Hermitian operators have real eigenvalues. Charles Hermite proved this key result for Hermitian matrices in 1855. This property is fundamental to quantum physics. After all, we measure real numbers in experimental observables.

- **Eigenstates are mutually orthogonal**. The second important property of Hermitian operators is that the states are orthogonal:

  $$\langle\phi_m|\phi_n\rangle = \delta_{nm} = \begin{cases} 1, & m = n, \\ 0, & m \neq n. \end{cases} \tag{10.35}$$

- **Eigenstates are complete**. We state without proof that the eigenstates $|\phi_n\rangle$ of $\hat{Q}$ form a *complete basis set*. Completeness means the eigenstates $|\phi_n\rangle$ form an orthonormal basis for all $1 \leq n \leq N$ of the Hermitian operator $\hat{Q}$ living in an N-dimensional Hilbert space. The proof exists in mathematical physics textbooks as the spectral theorem. This guarantees that we can expand any wavefunction in the eigenbasis states.

11 Measurement

11.1 Quantum prescription

Measurement is central to empirical science. We state its formalism in quantum mechanics.

- **Before measurement: wavefunction in superposition state**. Before measurement, a quantum state $|\psi\rangle$ is initially in a linear superposition of eigenstates $|\phi_n\rangle$

$$|\psi\rangle = \sum_n a_n |\phi_n\rangle = a_1|\phi_1\rangle + a_2|\phi_2\rangle + \dots a_N|\phi_N\rangle, \tag{11.1}$$

  where a_n are complex numbers denoting the *probability amplitudes* for observing each eigenstate $|\phi_n\rangle$. The eigenbasis $\{|\phi_n\rangle\}$ spans all possible states we can observe for the Hermitian operator $\hat{Q}$ in the N-dimensional Hilbert space. We can evaluate them by the overlap of eigenstate $\langle\phi_n|$ with the superposition state $|\psi\rangle$

$$a_n = \langle\phi_n|\psi\rangle. \tag{11.2}$$

- **Operator expectation values**. Before we perform a measurement in the laboratory, we can only predict the expectation value $\langle\hat{Q}\rangle$ (i.e. the mean average) of operator $\hat{Q}$

$$\langle\hat{Q}\rangle = \langle\psi|\hat{Q}|\psi\rangle. \tag{11.3}$$

  When a system is in a superposition state, the expectation value of $\hat{Q}$ is the sum of its eigenvalues q_n, satisfying $\hat{Q}|\phi_n\rangle = q_n|\phi_n\rangle$, weighted by the probabilities $|a_n|^2$

$$\langle\hat{Q}\rangle = \sum_n |a_n|^2 q_n = |a_1|^2 q_1 + |a_2|^2 q_2 + \cdots + |a_N|^2 q_N. \tag{11.4}$$

  The normalization condition $\langle\psi|\psi\rangle = 1$ guarantees sum of all probabilities is unity

$$\sum_n |a_n|^2 = 1. \tag{11.5}$$

- **Perform measurement: wavefunction collapses**. When we perform a measurement, the system wavefunction $|\psi\rangle$ collapses into one of its eigenstates $|\phi_n\rangle$

$$|\psi\rangle \xrightarrow[\text{measurement}]{} |\phi_n\rangle. \tag{11.6}$$

  This collapse is fundamentally non-deterministic and discontinuous. It is a radical departure from the continuous clockwork determinism of Newton's laws. The probability of measuring state $|\phi_n\rangle$ given it was prepared in state $|\psi\rangle$ follows the Born rule

$$P(\phi_n|\psi) = |\langle\phi_n|\psi\rangle|^2 = |a_n|^2. \tag{11.7}$$

Quantum mechanics cannot predict with certainty to which eigenstate $|\phi_n\rangle$ the wavefunction collapses. The theory can only predict the probabilities. This is *not* one of those "trust me" situations where your teacher is hiding the *Grand Complicated but Complete Theory of Nature* from you for simplified pedagogical purposes. This really is how the theory is formulated. We therefore state it again for emphasis:

Wavefunction collapse is a non-deterministic feature of quantum theory.

- **After measurement: single eigenstate**. The system is in one well-defined eigenstate $|\phi_n\rangle$. Repeat measurements of $\hat{Q}$ report the same expectation value (the eigenvalue q_n)

$$\langle \hat{Q} \rangle = \langle \phi_n | \hat{Q} | \phi_n \rangle = \langle \phi_n | q_n | \phi_n \rangle = q_n \langle \phi_n | \phi_n \rangle = q_n. \tag{11.8}$$

  In the laboratory, measuring the eigenvalue q_n tells us the system is in eigenstate $|\phi_n\rangle$.

Wavefunction collapse example Let us illustrate quantum measurement step by step with an example: radioactive decay of a cobalt-60 nucleus. The nucleus is prepared in a superposition state $|\psi\rangle$ that can be in either the decayed $|a\rangle$ or not decayed $|b\rangle$ state

$$|\psi\rangle = \sqrt{\frac{1}{3}}|a\rangle + \sqrt{\frac{2}{3}}|b\rangle. \tag{11.9}$$

We now apply the quantum prescription of measurement to the nucleus:

1. **Inner product with eigenstate**. To find the probability amplitude of measuring state as "decayed", we multiply from the left with $\langle a|$

$$\langle a|\psi\rangle = \sqrt{\frac{1}{3}}\langle a|a\rangle + \sqrt{\frac{2}{3}}\langle a|b\rangle. \tag{11.10}$$

2. **Orthonormality to amplitude**. Eigenstates are orthonormal, which implies that $\langle a|a\rangle = 1$ and $\langle a|b\rangle = 0$. The probability amplitude therefore becomes

$$\langle a|\psi\rangle = \sqrt{\frac{1}{3}}. \tag{11.11}$$

3. **Born rule for probability**. From the probability amplitude, we can find the probability of finding the state in state $|a\rangle$ given it was initially prepared in state $|\psi\rangle$

$$P(a|\psi) = |\langle a|\psi\rangle|^2 = \frac{1}{3}. \tag{11.12}$$

4. **Measured state**. After measurement, the wavefunction $|\psi\rangle$ collapsed (represented by $\rightarrow$) to the measured state $|a\rangle$ with probability $|\langle a|\psi\rangle|^2$

$$|\psi\rangle \rightarrow |a\rangle, \quad P(a) = \frac{1}{3}. \tag{11.13}$$

The same steps determine the probability that the wavefunction collapses into state $|b\rangle$ is

$$|\psi\rangle \rightarrow |b\rangle, \quad P(b|\psi) = |\langle b|\psi\rangle|^2 = \frac{2}{3}. \tag{11.14}$$

That's it! Nothing in the theory tells us to which state $|a\rangle$ or $|b\rangle$ the wavefunction collapses. It is non-deterministic. The theory only predicts the probabilities of measuring each state.

11.2 Commutators

In quantum mechanics, the algebra of operators tells us if we can measure both observables simultaneously. The commutator of two operators A and B is defined as

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}. \tag{11.15}$$

The order of multiplying two matrices is generally important. Here are some useful properties they obey. Their reverse is simply the same with a minus sign

$$[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}]. \tag{11.16}$$

Commutators are linear, such that

$$[\hat{A}, \hat{B} + \hat{C}] = [\hat{A}, \hat{B}] + [\hat{A}, \hat{C}] \tag{11.17}$$

The commutator for the product of operators $\hat{A}\hat{B}$ with a third operator $\hat{C}$ is

$$[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B}. \tag{11.18}$$

You can prove this by writing out the commutators explicitly. A helpful way to remember this commutator identity is to notice how this is similar to the product rule when differentiating

$$\frac{\partial(AB)}{\partial C} = A\frac{\partial B}{\partial C} + \frac{\partial A}{\partial C}B. \tag{11.19}$$

When we first encounter commutators of operators in (11.15), it can be helpful to remember that they can represent differential operators that act on a wavefunction. It can be safer to write this out explicitly when evaluating commutators

$$[\hat{A}, \hat{B}]\psi(x) \equiv (\hat{A}\hat{B} - \hat{B}\hat{A})\psi(x). \tag{11.20}$$

11.3 Compatibility

If two operators $\hat{A}$ and $\hat{B}$ commute, $[\hat{A},\hat{B}]=0$, they are *compatible*. Compatibility means a system can be in an eigenstate of both operators simultaneously. First consider an state $|\phi_n\rangle$ that is an eigenstate of both $\hat{A}$ with eigenvalue a_n and an eigenstate of $\hat{B}$ with eigenvalue b_n. This implies that the product of operators gives

$$\hat{A}\hat{B}|\phi_n\rangle = \hat{A}(b_n|\phi_n\rangle) = b_n(\hat{A}|\phi_n\rangle) = b_n a_n|\phi_n\rangle. \tag{11.21}$$

Similarly, we can consider the operators in reversed order

$$\hat{B}\hat{A}|\phi_n\rangle = a_n b_n|\phi_n\rangle \tag{11.22}$$

Subtracting $\hat{B}\hat{A}|\phi_n\rangle$ from $\hat{A}\hat{B}|\phi_n\rangle$ gives

$$[\hat{A},\hat{B}]|\phi_n\rangle = (\hat{A}\hat{B}-\hat{B}\hat{A})|\phi_n\rangle = (b_n a_n - a_n b_n)|\phi_n\rangle. \tag{11.23}$$

As eigenvalues a_n, b_n are real numbers, they commute $b_n a_n = a_n b_n$ and this evaluates to zero

$$[\hat{A},\hat{B}]|\phi_n\rangle = 0. \tag{11.24}$$

This is important for physics. If two operators $\hat{A}$ and $\hat{B}$ commute, $[\hat{A},\hat{B}]=0$, there exists the same eigenstates $|\phi_n\rangle$ labelled by their eigenvalues that satisfy the two eigenvalue equations

$$\hat{A}|\phi_n\rangle = a_n|\phi_n\rangle, \quad \hat{B}|\phi_n\rangle = b_n|\phi_n\rangle. \tag{11.25}$$

We can observe the eigenvalues $\{a_n, b_n\}$ simultaneously. This cannot be true if $[\hat{A},\hat{B}]\neq 0$.

11.4 Canonical commutation relation

Let us illustrate the commutation relations of the position $\hat{x}$ and momentum $\hat{p} = -\mathrm{i}\hbar\frac{\partial}{\partial x}$ operators. First proceed with sanity checks. For $[\hat{x},\hat{x}]$ acting on any wavefunction $\psi(x)$, we find $[\hat{x},\hat{x}]\psi(x) = (xx-xx)\psi(x) = 0$. If that was too easy, check two dimensions: for $\hat{x}$ and $\hat{y}$ in the two orthogonal axes of two-dimensional space, their commutator is

$$[\hat{x},\hat{y}]\psi(x,y) = (xy-yx)\psi(x,y) = 0. \tag{11.26}$$

Sure enough, this commutes because elementary school taught us the order we multiply two numbers does not matter: $4\times 5 = 5\times 4$. We can measure each independent coordinate simultaneously. As $[\hat{p}_x,\hat{p}_x]=0$ is similarly easy, let's check $[\hat{p}_x,\hat{p}_y]$ in two different dimensions:

$$[\hat{p}_x,\hat{p}_y]\psi(x,y) = (\hat{p}_x\hat{p}_y - \hat{p}_y\hat{p}_x)\psi(x,y) = (-\mathrm{i}\hbar)^2\left\{\frac{\partial}{\partial x}\left(\frac{\partial\psi}{\partial y}\right) - \frac{\partial}{\partial y}\left(\frac{\partial\psi}{\partial x}\right)\right\} = 0. \tag{11.27}$$

This vanishes because the order of partial derivatives does not matter. These results agree with classical intuition and generalize to the three spatial dimensions of our universe.

We next check the commutator of position and momentum

$$[\hat{x}, \hat{p}]\,\psi(x) = -\mathrm{i}\hbar \left(x\frac{\partial}{\partial x} - \frac{\partial}{\partial x}x \right) \psi(x). \tag{11.28}$$

The first term only differentiates $\psi(x)$. However, we must take care with the second term, which differentiates the product $x\psi(x)$. We see why it is important to act commutators onto a wavefunction. We must use the product rule, which gives a non-zero piece

$$[\hat{x}, \hat{p}]\,\psi(x) = \mathrm{i}\hbar\,\psi(x) \left(\frac{\partial}{\partial x}x \right) = \mathrm{i}\hbar\psi(x) \quad \text{for all } \psi(x). \tag{11.29}$$

This implies the position and momentum operators do not commute

$$[\hat{x}, \hat{p}] = \mathrm{i}\hbar. \tag{11.30}$$

This commutator enjoys the grandiose name of the **canonical commutation relation**. Werner Heisenberg formalized this deep result in 1927 [33]. As we shall prove in section 13.1, it implies the uncertainty principle. Unlike classical mechanics, a quantum particle's position and momentum cannot simultaneously be well-defined. As $\hat{x}$ and $\hat{p}$ are not compatible, there does not exist a complete set of states that are eigenstates of both $\hat{x}$ and $\hat{p}$. Mathematically, the commutator arises from the product rule. Physically, the momentum derivative $\hat{p} = -\mathrm{i}\hbar\partial/\partial x$ is from the de Broglie relation $p = h/\lambda$ that we are forced to accept by the Davisson-Germer experiment. The quantum leap is treating particles as waves.

HW6: Operators and measurement (set Oct 8, due Oct 15)

1. What is the definition of a *Hermitian operator* and why is this important for physics? If $\hat{X}$ and $\hat{P}$ are Hermitian operators, show that $\mathrm{i}[\hat{X},\hat{P}]$ is also Hermitian.

2. What is the condition for an operator $\hat{Q}$ acting on wavefunctions $\psi(x),\phi(x)$ to be Hermitian in Dirac and integral notation? Why do $\psi(x)$ and $\phi(x)$ vanish as $x\to\infty$? Prove via integration by parts that the momentum operator $\hat{p}=-\mathrm{i}\hbar\frac{\partial}{\partial x}$ is Hermitian.

3. Consider the state $|\psi\rangle = a|A\rangle + b|B\rangle$.

   (a) What is the amplitude $\langle A|\psi\rangle$ and explain your answer? Upon measurement, what is the probability that the state is in $|A\rangle$ and $|B\rangle$?

   (b) What is the probability of measuring state $|A\rangle$ if (i) $a=\mathrm{e}^{\mathrm{i}\pi}$, (ii) $b=\mathrm{i}/2$?

4. Given the adjoint of an operator product is $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$, show that $(\hat{A}\hat{B}\hat{C})^\dagger = \hat{C}^\dagger\hat{B}^\dagger\hat{A}^\dagger$. By explicitly writing out the left and right hand sides, prove the following identities

$$[\hat{A},\hat{B}+\hat{C}] = [\hat{A},\hat{B}] + [\hat{A},\hat{C}], \tag{11.31}$$

$$[\hat{A}\hat{B},\hat{C}] = \hat{A}[\hat{B},\hat{C}] + [\hat{A},\hat{C}]\hat{B}. \tag{11.32}$$

5. Consider two observables of Hermitian operators $\hat{A}$ and $\hat{B}$ with eigenvalue equations

$$\hat{A}|a_n\rangle = a_n|a_n\rangle, \quad \hat{B}|b_n\rangle = b_n|b_n\rangle. \tag{11.33}$$

   (a) Given $\hat{A}$ is Hermitian, show that its (i) eigenvalues must be real, (ii) eigenstates are mutually orthogonal. If the $[\hat{A},\hat{B}]=0$ commute, show that they are simultaneous eigenstates of both $\hat{A}$ and $\hat{B}$. Why are these results important for physics?

   (b) Express the state $|\psi\rangle$ as a expansion of eigenstates $|a_n\rangle$ weighted by amplitudes $\gamma_n = \langle a_n|\psi\rangle$. Using this, derive the expectation value of $\hat{A}$ when acting on $|\psi\rangle$.

   (c) Consider $[\hat{A},\hat{B}]\neq 0$ on state $|\psi\rangle$. Let us measure one observable first, then the other immediately after without resetting the system. Show that the probability of measuring $\hat{A}$ first then $\hat{B}$ is $P = |\langle b_n|a_n\rangle|^2|\langle a_n|\psi\rangle|^2$. Find the probability of measuring $\hat{B}$ first then $\hat{A}$ and show that it differs. Discuss the physical significance.

6. By writing out the differential operators and acting on a general wavefunction $\psi(\mathbf{x})$, prove the canonical commutation relation for any number of dimensions $i,j=1,2,\ldots,N$

$$[\hat{x}_i,\hat{p}_j] = \mathrm{i}\hbar\delta_{ij}, \tag{11.34}$$

   where the Kronecker delta is $\delta_{ij}=0$ if $i\neq j$ and $\delta_{ij}=1$ if $i=j$. What is the physical significance of this commutator for $i=j$ and $i\neq j$? Why does $[\hat{x}_i,\hat{x}_j]=[\hat{p}_i,\hat{p}_j]=0$?

12 Harmonic oscillator

We illustrate the operator formalism with an important example: the simple harmonic oscillator. The classical Hooke's law describes the quadratic potential energy $V(x) = \frac{1}{2}kx^2$ stored in a spring. The spring constant $k = m\omega^2$ is related to the particle mass m and oscillation angular frequency ω. We extend this to a quantum particle with the Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2. \tag{12.1}$$

We use this to introduce the *algebraic* method of solving Schrödinger's equation using ladder operators. This approach finds important applications, from describing electrons bound inside materials to electromagnetic waves in quantum field theory.

12.1 Ladder operators

The Hamiltonian is quadratic in $\hat{x}$ and $\hat{p}$. This quadratic form motivates a "complete the square" trick $f^2 + g^2 = (f - \mathrm{i}g)(f + \mathrm{i}g)$ to factorize the Hamiltonian into a pair of operators

$$\hat{A} = \frac{1}{\sqrt{2m\hbar\omega}}(m\omega x + \mathrm{i}\hat{p}), \tag{12.2}$$

$$\hat{A}^\dagger = \frac{1}{\sqrt{2m\hbar\omega}}(m\omega x - \mathrm{i}\hat{p}). \tag{12.3}$$

These are called *ladder operators*: $\hat{A}$ is the lowering (or annihilation) operator and its Hermitian adjoint $\hat{A}^\dagger$ is the raising (or creation) operator. The product of these operators is

$$\hat{A}^\dagger\hat{A} = \frac{1}{2m\hbar\omega}(m\omega\hat{x} - \mathrm{i}\hat{p})(m\omega\hat{x} + \mathrm{i}\hat{p}). \tag{12.4}$$

Because position and momentum operators do not commute $\hat{x}\hat{p} \neq \hat{p}\hat{x}$, be careful to keep the order of multiplication strict when expanding brackets

$$\hat{A}^\dagger\hat{A} = \frac{1}{2m\hbar\omega}\{(m\omega x)^2 + \mathrm{i}m\omega\underbrace{(\hat{x}\hat{p} - \hat{p}\hat{x})}_{[\hat{x},\hat{p}]=\mathrm{i}\hbar} + \hat{p}^2\} = \frac{1}{\hbar\omega}\hat{H} - \frac{1}{2}. \tag{12.5}$$

This enables us to write the Hamiltonian in two equivalent forms

$$\hat{H} = \left(\hat{A}^\dagger\hat{A} + \frac{1}{2}\right)\hbar\omega, \quad \hat{H} = \left(\hat{A}\hat{A}^\dagger - \frac{1}{2}\right)\hbar\omega. \tag{12.6}$$

Adding and subtracting these two forms for $\hat{H}$ gives the anti-commutation $\{\hat{A},\hat{A}^\dagger\}$ and commutation $[\hat{A},\hat{A}^\dagger]$ relations for the lowering and raising operators

$$\{\hat{A},\hat{A}^\dagger\} \equiv \hat{A}\hat{A}^\dagger + \hat{A}^\dagger\hat{A} = \frac{2\hat{H}}{\hbar\omega}, \tag{12.7}$$

$$[\hat{A},\hat{A}^\dagger] \equiv \hat{A}\hat{A}^\dagger - \hat{A}^\dagger\hat{A} = 1. \tag{12.8}$$

12.2 Energy eigenstates

Let us write $|E_n\rangle$ as the eigenstates of the Hamiltonian $\hat{H}$ with energy eigenvalue E_n

$$\hat{H}|E_n\rangle = E_n|E_n\rangle. \tag{12.9}$$

When we operate a raising operator $\hat{A}^\dagger$ on an energy eigenstate $|E_n\rangle$, we get a different state $|\psi\rangle = \hat{A}^\dagger|E_n\rangle$. What are the energy eigenvalues for $|\psi\rangle$? To find out, we operate the Hamiltonian on this new state $|\psi\rangle$ and use an algebraic trick $0 = -\hat{A}^\dagger\hat{H} + \hat{A}^\dagger\hat{H}$ to write

$$\hat{H}|\psi\rangle = \hat{H}(\hat{A}^\dagger|E_n\rangle) = (\hat{H}\hat{A}^\dagger - \hat{A}^\dagger\hat{H} + \hat{A}^\dagger\hat{H})|E_n\rangle \tag{12.10}$$

$$= ([\hat{H},\hat{A}^\dagger] + \hat{A}^\dagger\hat{H})|E_n\rangle. \tag{12.11}$$

To evaluate the commutator $[\hat{H},\hat{A}^\dagger]$, we apply the rules of commutator multiplication (11.18) $[\hat{A}\hat{B},\hat{C}] = \hat{A}[\hat{B},\hat{C}] + [\hat{A},\hat{C}]\hat{B}$ to find

$$[\hat{H},\hat{A}^\dagger] = \hbar\omega[\hat{A}^\dagger\hat{A},\hat{A}^\dagger] = \hbar\omega\hat{A}^\dagger\underbrace{[\hat{A},\hat{A}^\dagger]}_{=1} + \hbar\omega\underbrace{[\hat{A}^\dagger,\hat{A}^\dagger]}_{=0}\hat{A} = \hbar\omega\hat{A}^\dagger, \tag{12.12}$$

where we used equation (12.8) $[\hat{A},\hat{A}^\dagger] = 1$ to evaluate the commutator. As an exercise, you can show that $[\hat{H},\hat{A}] = -\hbar\omega\hat{A}$. Using $\hat{H}|E_n\rangle = E_n|E_n\rangle$, equation (12.11) simplifies to

$$\hat{H}|\psi\rangle = (E_n + \hbar\omega)|\psi\rangle, \quad |\psi\rangle = \hat{A}^\dagger|E_n\rangle. \tag{12.13}$$

This shows that the state $|\psi\rangle = \hat{A}^\dagger|E_n\rangle$ is also an energy eigenstate of $\hat{H}$. However, the eigenvalue is one discrete unit of $\hbar\omega$ higher than the original energy E_n. The action of $\hat{A}^\dagger$ operating on $|E_n\rangle$ brings us into a higher energy eigenstate $|E_n + \hbar\omega\rangle$ by one unit of $\hbar\omega$. Similarly, the lowering operator reduces an energy eigenstate $|\phi\rangle = \hat{A}|E_n\rangle$ by one unit of $\hbar\omega$

$$\hat{H}|\phi\rangle = (E_n - \hbar\omega)|\phi\rangle, \quad |\phi\rangle = \hat{A}|E_n\rangle, \tag{12.14}$$

To summarize, $\hat{A}^\dagger$ raises an energy state by $\hbar\omega$ while $\hat{A}$ lowers an energy state by $\hbar\omega$:

$$\hat{A}^\dagger|E_n\rangle \propto |E_n + \hbar\omega\rangle, \quad \hat{A}|E_n\rangle \propto |E_n - \hbar\omega\rangle. \tag{12.15}$$

In the language of differential equations, quantized energies arise from boundary conditions. In the language of algebra, quantized energies arise from non-commuting operators.

12.3 Spectrum and wavefunctions

Repeatedly applying $A^\dagger$ constructs an infinite ladder of discrete energy eigenstates. Figure 28 shows this evenly spaced ladder of energies. The states are uniformly separated by $\hbar\omega$ intervals. We can evaluate $\langle E_n|\hat{H}|E_n\rangle$ to establish that the energies E_n must be non-negative:

$$E_n = \langle E_n|\hat{H}|E_n\rangle = \frac{1}{2m\omega}\left(|\hat{p}|E_n\rangle|^2 + m^2\omega^2|\hat{x}|E_n\rangle|^2\right) \geq 0. \tag{12.16}$$

Energy positivity implies there must exist a minimum energy E_0. The lowering operator cannot further reduce energies and instead makes the state $|E_0\rangle$ vanish

$$\hat{A}|E_0\rangle = 0. \tag{12.17}$$

We evaluate the minimum energy by operating from the left with $\langle E_0|\hat{A}^\dagger$ to give

$$0 = \langle E_0|\hat{A}^\dagger\hat{A}|E_0\rangle = \langle E_0|\left(\frac{\hat{H}}{\hbar\omega} - \frac{1}{2}\right)|E_0\rangle = \frac{E_0}{\hbar\omega} - \frac{1}{2}, \tag{12.18}$$

where we used the Hamiltonian (12.6) in the final equality. The minimum energy is non-zero

$$E_0 = \frac{1}{2}\hbar\omega. \tag{12.19}$$

This minimum-energy state is called the ground state. Therefore, we establish the energy spectrum of the quantum harmonic oscillator is

$$E_n = \left(n + \tfrac{1}{2}\right)\hbar\omega, \quad n = 0, 1, 2, \ldots \tag{12.20}$$

The energies are an infinite ladder of positive values separated by integer n units of $\hbar\omega$ with a minimum energy of $\hbar\omega/2$.

We next derive the ground state wavefunction in the x representation $\langle x|E_0\rangle = \psi_0(x)$. We start with the condition that the lowering operator makes the ground state vanish $\hat{A}|E_0\rangle = 0$ of equation (12.17). Multiply from the left with $\langle x|$ gives $\langle x|\hat{A}|E_0\rangle = 0$. Recalling the operators are $\hat{x} = x, \hat{p} = -\mathrm{i}\hbar\partial/\partial x$, this becomes the first-order differential equation

$$\langle x|\hat{A}|E_0\rangle = 0 \quad \Rightarrow \quad \frac{1}{\sqrt{2m\hbar\omega}}\left(m\omega x\langle x|E_0\rangle + \hbar\frac{\mathrm{d}\langle x|E_0\rangle}{\mathrm{d}x}\right) = 0. \tag{12.21}$$

Solving this for $\langle x|E_0\rangle$ and normalizing, we find the ground state eigenfunction is a Gaussian

$$\langle x|E_0\rangle = \psi_0(x) = \frac{1}{(2\pi\ell)^{1/4}}\exp\left(-\frac{x^2}{4\ell^2}\right), \quad \text{where} \quad \ell \equiv \sqrt{\frac{\hbar}{2m\omega}}. \tag{12.22}$$

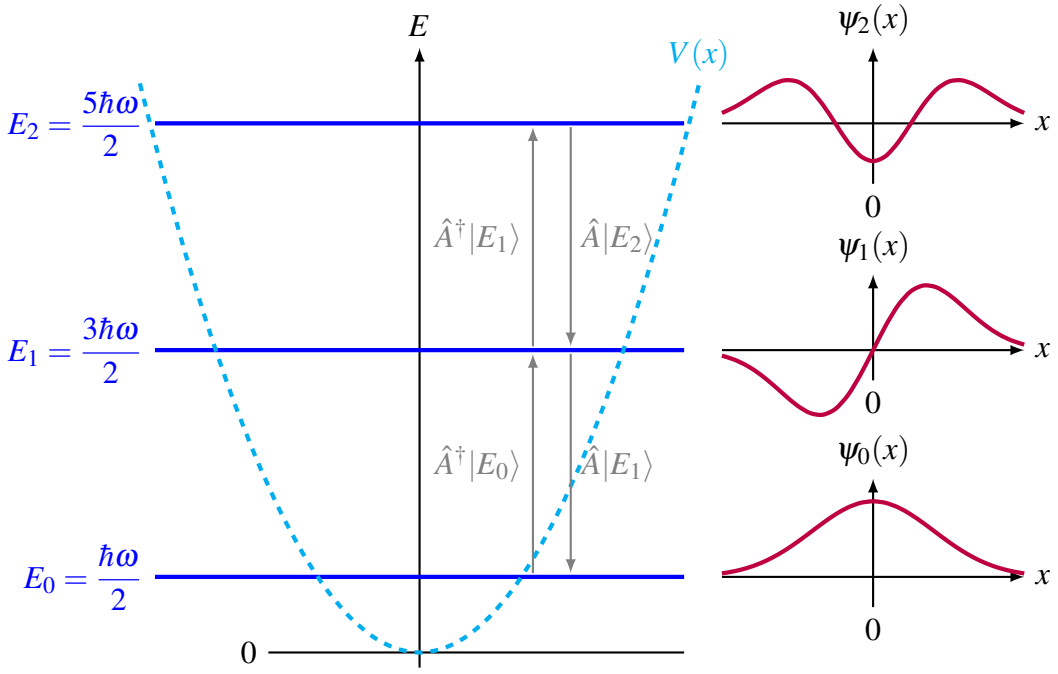


Figure 28: Sketch of harmonic oscillator solutions showing the first few energy levels and corresponding wavefunctions.

We obtain the first excited state $\langle x|E_1\rangle = \psi_1(x)$ using the raising operator $\hat{A}^\dagger|E_0\rangle$ gives

$$\langle x|E_1\rangle = \psi_1(x) \propto x \exp\left(-\frac{x^2}{4\ell^2}\right). \tag{12.23}$$

We note $\langle x|E_0\rangle$ has even parity and $\langle x|E_1\rangle$ has odd parity, which is a rule that continues for higher n. Figure 28 sketches the first few eigenstates called Hermite polynomials.

12.4 Zero-point energy

We wish to evaluate the uncertainty relation by calculating the position $\sigma_x^2 = \langle x^2\rangle - \langle x\rangle^2$ and momentum $\sigma_p^2 = \langle p^2\rangle - \langle p\rangle^2$ variances defined in (3.17). As an exercise, first recast the position and momentum operator in terms of ladder operators from equation (12.3)

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}(\hat{A} + \hat{A}^\dagger), \quad \hat{p} = \mathrm{i}\sqrt{\frac{m\omega\hbar}{2}}(\hat{A}^\dagger - \hat{A}). \tag{12.24}$$

From this, we can evaluate the mean position $\langle x\rangle = \langle E_n|\hat{x}|E_n\rangle$ is given by

$$\langle x\rangle = \sqrt{\frac{\hbar}{2m\omega}}\langle E_n|(\hat{A} + \hat{A}^\dagger)|E_n\rangle = \sqrt{\frac{\hbar}{2m\omega}}\left(\langle E_n|E_{n+1}\rangle + \langle E_{n-1}|E_n\rangle\right) = 0, \tag{12.25}$$

where the lowering and raising operators cause $\langle E_n|E_{n+1}\rangle = 0$ terms to vanish by orthogonality of states. The mean of the position squared $\langle x^2\rangle = \langle E_n|\hat{x}^2|E_n\rangle$ quantifies the spread in position and is non-vanishing

$$\langle x^2\rangle = \frac{\hbar}{2m\omega}\langle E_n|(\hat{A}+\hat{A}^\dagger)^2|E_n\rangle = \frac{\hbar}{2m\omega}\langle E_n|\left(\hat{A}\hat{A}+\hat{A}\hat{A}^\dagger+\hat{A}^\dagger\hat{A}+\hat{A}^\dagger\hat{A}^\dagger\right)|E_n\rangle. \tag{12.26}$$

The $\hat{A}\hat{A}$ and $\hat{A}^\dagger\hat{A}^\dagger$ make orthogonal states $\langle E_n|E_{n+1}\rangle = 0$ that vanish. Meanwhile, the anti-commutation relation (12.7) implies $\hat{A}\hat{A}^\dagger + \hat{A}^\dagger\hat{A} = 2\hat{H}/\hbar\omega$ so the position variance is

$$\sigma_x^2 = \langle x^2\rangle - \langle x\rangle^2 = \frac{\hbar}{2m\omega}\frac{2}{\hbar\omega}\langle E_n|\hat{H}|E_n\rangle - 0 = \frac{\hbar}{m\omega}\left(n+\frac{1}{2}\right). \tag{12.27}$$

You can show the mean momentum also vanishes $\langle p\rangle = 0$. The momentum uncertainty is

$$\sigma_p^2 = \langle p^2\rangle - \langle p\rangle^2 = \langle E_n|\hat{p}^2|E_n\rangle = -\frac{m\omega\hbar}{2}\langle E_n|(\hat{A}-\hat{A}^\dagger)^2|E_n\rangle = -m\omega\hbar\left(n+\frac{1}{2}\right). \tag{12.28}$$

Taking the magnitude of the product $|\sigma_x^2\sigma_p^2|$ gives the uncertainty relation

$$\sigma_x\sigma_p = \hbar\left(n+\frac{1}{2}\right) \geq \frac{\hbar}{2}. \tag{12.29}$$

The product $\sigma_x\sigma_p$ is bounded from below by $\hbar/2$. This inequality satisfies the same uncertainty principle that we encountered for the Gaussian wavepacket from (7.26). This is perhaps unsurprising given both wavefunctions are Gaussians.

Notably, we see the ground state $n = 0$ of the harmonic oscillator minimizes the uncertainty relation at the quantum limit $\hbar/2$. The uncertainty principle requires that the zero-point energy is non-zero for states bound inside a potential. This physically arises because a particle is localized in space when trapped inside a potential, so there must be a minimum momentum p and thus kinetic energy $p^2/2m$.

13 Uncertainty and conservation

13.1 Heisenberg uncertainty principle

We have seen how the position-momentum dispersions of the Gaussian wavepacket (7.26) and harmonic oscillator (12.29) satisfy the inequality $\sigma_x\sigma_p \geq \hbar/2$. We now prove this uncertainty principle in generality for non-commuting operators. The variances of operators $\hat{x}, \hat{p}$ are defined by its mean square deviations from their expectation values:

$$\sigma_x^2 = \langle\psi|(\hat{x} - \langle\hat{x}\rangle)^2|\psi\rangle = \langle\psi|\hat{X}^2|\psi\rangle, \tag{13.1}$$
$$\sigma_p^2 = \langle\psi|(\hat{p} - \langle\hat{p}\rangle)^2|\psi\rangle = \langle\psi|\hat{P}^2|\psi\rangle. \tag{13.2}$$

To simplify notation, we defined the parentheses as a shift by the mean $\hat{X} = \hat{x} - \langle\hat{x}\rangle, \hat{P} = \hat{p} - \langle\hat{p}\rangle$. The mean position $\langle\hat{x}\rangle$ and momentum $\langle\hat{p}\rangle$ are just numbers so they commute. The commutators are therefore equal

$$[\hat{X}, \hat{P}] = [\hat{x}, \hat{p}]. \tag{13.3}$$

To proceed, we construct the state $|\phi\rangle$ with a complex parameter λ as a linear superposition

$$|\phi\rangle = \hat{X}|\psi\rangle + \mathrm{i}\lambda\hat{P}|\psi\rangle. \tag{13.4}$$

From this state, we can evaluate its modulus square

$$\langle\phi|\phi\rangle = \left(\langle\psi|\hat{X} - \mathrm{i}\lambda\langle\psi|\hat{P}\right)\left(\hat{X}|\psi\rangle + \mathrm{i}\lambda\hat{P}|\psi\rangle\right) \tag{13.5}$$
$$= \langle\psi|\hat{X}^2|\psi\rangle + \lambda^2\langle\psi|\hat{P}^2|\psi\rangle + \mathrm{i}\lambda\langle\psi|\hat{X}\hat{P}|\psi\rangle - \mathrm{i}\lambda\langle\psi|\hat{P}\hat{X}|\psi\rangle, \tag{13.6}$$

where $\hat{X}^\dagger = \hat{X}$ and $\hat{P}^\dagger = \hat{P}$ are Hermitian. Recognizing the variances from equations (13.1) and (13.2) and the definition of the commutator, we write this as

$$\langle\phi|\phi\rangle = \sigma_x^2 + \lambda^2\sigma_p^2 + \mathrm{i}\lambda\langle\psi|[\hat{X}, \hat{P}]|\psi\rangle \geq 0. \tag{13.7}$$

The modulus square of quantum states is positive definite $\langle\phi|\phi\rangle \geq 0$. This inequality is saturated at the minimum of $\langle\phi|\phi\rangle$. We find this by differentiating with respect to λ

$$2\lambda_{\min}\sigma_p^2 + \mathrm{i}\langle[\hat{X}, \hat{P}]\rangle = 0, \quad \Rightarrow \quad \lambda_{\min} = -\mathrm{i}\frac{\langle[\hat{X}, \hat{P}]\rangle}{2\sigma_p^2}. \tag{13.8}$$

Substituting $\lambda_{\min}$ into the inequality (13.7) gives

$$\sigma_x^2\sigma_p^2 \geq -\frac{1}{4}\langle[\hat{X}, \hat{P}]\rangle^2. \tag{13.9}$$

Applying the commutator (13.3) and the square root gives a statement about the magnitudes

$$\sigma_x \sigma_p \geq \left| \frac{\mathrm{i}}{2} \langle [\hat{x}, \hat{p}] \rangle \right|. \tag{13.10}$$

Applying the canonical commutation relation $[\hat{x}, \hat{p}] = \mathrm{i}\hbar$, we obtain the famous **Heisenberg uncertainty principle**

$$\sigma_x \sigma_p \geq \frac{\hbar}{2}. \tag{13.11}$$

If we perform a measurement on position, the state collapses to an eigenstate of $\hat{x}$, but exists in a superposition of momentum eigenstates. The more precisely we know the position of a particle, the less precisely we know its momentum.

The derivation of the inequality (13.10) generalizes for any Hermitian operators $\hat{A}, \hat{B}$. The uncertainties of two operators $\hat{A}$ and $\hat{B}$ satisfy this *generalized uncertainty relation*

$$\sigma_A \sigma_B \geq \left| \frac{\mathrm{i}}{2} \langle [\hat{A}, \hat{B}] \rangle \right|. \tag{13.12}$$

For $[\hat{A}, \hat{B}] \neq 0$, the more precision we have for one observable, the less precise the other is. In classical physics, uncertainties on observables arise from the experimental apparatus. For example, the smallest divisions of a ruler being a millimetre mean we can only measure the length of a stationary table to be $x \pm \sigma_x = 212.027 \pm 0.001$ m. It is our experimental capability and not a feature of Newton's laws that obstructs measuring its length to arbitrary precision. By contrast, the uncertainty principle is intrinsic to quantum theory.

13.2 Unitary evolution

In quantum mechanics, time only appears as a parameter in Schrödinger's equation. We restate this for *any* state $|\psi\rangle$ in Dirac notation, not just the spatial wavefunction $\psi(x)$ of (4.1)

$$\mathrm{i}\hbar \frac{\partial |\psi\rangle}{\partial t} = \hat{H} |\psi\rangle. \tag{13.13}$$

For a time-independent Hamiltonian $\hat{H}$, we can express the time evolution of a state $|\psi(t)\rangle$ as a unitary operator $\hat{U} = \exp\left(-\mathrm{i}\hat{H}t/\hbar\right)$

$$|\psi(t)\rangle = \hat{U} |\psi(0)\rangle = \exp\left(-\frac{\mathrm{i}\hat{H}t}{\hbar}\right) |\psi(0)\rangle. \tag{13.14}$$

The exponential of an operator is formally defined via a Taylor series

$$\hat{U} = \exp\left(-\frac{\mathrm{i}\hat{H}t}{\hbar}\right) = 1 - \frac{\mathrm{i}\hat{H}t}{\hbar} + \frac{1}{2}\left(-\frac{\mathrm{i}\hat{H}t}{\hbar}\right)^2 + \ldots \tag{13.15}$$

An matrix is unitary if its inverse $\hat{U}^{-1}$ is the Hermitian conjugate

$$\hat{U}^{-1} = \hat{U}^\dagger, \quad \Rightarrow \quad \hat{U}^\dagger \hat{U} = \mathbb{I}. \tag{13.16}$$

A unitary transformation leaves the inner product invariant

$$\langle \phi' | \psi' \rangle = \langle \hat{U}\phi | \hat{U}\psi \rangle = \langle \phi | \hat{U}^\dagger \hat{U} | \psi \rangle = \langle \phi | \psi \rangle. \tag{13.17}$$

Energy eigenstates $|E_n\rangle$ only undergo phase changes with time at rate $E_n/\hbar$

$$|E_n(t)\rangle = |E_n(0)\rangle \exp\left(-\frac{\mathrm{i}E_n t}{\hbar}\right). \tag{13.18}$$

These are stationary states that we met back in section 4.3. For a general initial state $|\psi(0)\rangle$ at $t = 0$, we can write its time evolved state $|\psi(t)\rangle$ in the energy eigenbasis as

$$|\psi(t)\rangle = \sum_n \exp\left(-\frac{\mathrm{i}E_n t}{\hbar}\right) |E_n\rangle\langle E_n|\psi(0)\rangle. \tag{13.19}$$

13.3 Ehrenfest's theorem

How does the expectation value of a quantum observable $\hat{Q}$ change with time? We can evaluate the time derivative with an operator that may depend on time

$$\frac{\mathrm{d}\langle \hat{Q} \rangle}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\langle \psi | \hat{Q} | \psi \rangle. \tag{13.20}$$

Consider the case where $\hat{Q}$ is time independent. Applying the product rule, this becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \psi | \hat{Q} | \psi \rangle = \left(\frac{\partial \langle \psi |}{\partial t}\right) \hat{Q} | \psi \rangle + \langle \psi | \underbrace{\left(\frac{\partial \hat{Q}}{\partial t}\right)}_{=0} | \psi \rangle + \langle \psi | \hat{Q} \left(\frac{\partial | \psi \rangle}{\partial t}\right). \tag{13.21}$$

Now apply the time-dependent Schrödinger equation $\mathrm{i}\hbar\partial|\psi\rangle/\partial t = \hat{H}|\psi\rangle$ to write this as

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \psi | \hat{Q} | \psi \rangle = \frac{1}{\mathrm{i}\hbar}\langle \psi | \hat{Q}\hat{H} | \psi \rangle - \frac{1}{\mathrm{i}\hbar}\langle \psi | \hat{H}\hat{Q} | \psi \rangle, \tag{13.22}$$

where we note the Hamiltonian is Hermitian $\hat{H}^\dagger = \hat{H}$. With the commutator $[\hat{Q}, \hat{H}] = \hat{Q}\hat{H} - \hat{H}\hat{Q}$, we find the rate of change of the expectation of a time-independent operator $\hat{Q}$ is

$$\frac{\mathrm{d}\langle \hat{Q} \rangle}{\mathrm{d}t} = \frac{1}{\mathrm{i}\hbar}\langle [\hat{Q}, \hat{H}] \rangle. \tag{13.23}$$

This is called **Ehrenfest's theorem** (for time-independent operators), which implies:

- The expectation value of $\langle\hat{Q}\rangle$ is therefore a *conserved quantity* or constant of motion, $\langle\hat{Q}\rangle = \text{constant}$, if it commutes with the Hamiltonian $[\hat{Q},\hat{H}] = 0$.
- If the eigenvalues q_i of operator $\hat{Q}$ and $[\hat{Q},\hat{H}] = 0$, then q_i does not change with time and are called *good quantum numbers* that label states.

Let us examine two important physical examples of conservation laws:

- **Hamiltonian operator** $\hat{Q} = \hat{H}$. The Hamiltonian commutes with itself. The expectation value of the Hamiltonian is the average energy $\langle\hat{H}\rangle = \langle E|\hat{H}|E\rangle = \langle E\rangle$ so
$$\frac{\mathrm{d}\langle E\rangle}{\mathrm{d}t} = 0 \quad \Rightarrow \quad \langle E\rangle = \text{constant}. \tag{13.24}$$
This expresses the conservation of average energy for time-independent Hamiltonians.
- **Momentum operator** $\hat{Q} = \hat{p}$. Applying Ehrenfest's theorem gives
$$\frac{\mathrm{d}\langle\hat{p}\rangle}{\mathrm{d}t} = \frac{1}{\mathrm{i}\hbar}\langle[\hat{p},\hat{H}]\rangle = \frac{1}{\mathrm{i}\hbar}\langle[\hat{p},V(\hat{x})]\rangle, \tag{13.25}$$
where the Hamiltonian is $\hat{H} = \hat{p}^2/2m + V(\hat{x})$ and $\hat{p}$ commutes with itself. We then need to evaluate the commutator of $\hat{p}$ with a function of the position operator $V(\hat{x})$. Assume that $V(\hat{x})$ is smooth and continuous such that we can Taylor expand
$$V(x) = V(0) + x\left.\frac{\mathrm{d}V}{\mathrm{d}x}\right|_{x=0} + \frac{x^2}{2}\left.\frac{\mathrm{d}^2V}{\mathrm{d}x^2}\right|_{x=0} + \ldots \tag{13.26}$$
We use the simplifying notation $V' = \left.\frac{\mathrm{d}V}{\mathrm{d}x}\right|_{x=0}$ for derivatives to write
$$[\hat{p},V(\hat{x})] = [\hat{p},\hat{x}]\left(V' + V''\hat{x} + V'''\frac{\hat{x}^2}{2} + \ldots\right) = [\hat{p},\hat{x}]\frac{\mathrm{d}V}{\mathrm{d}\hat{x}} = -\mathrm{i}\hbar\frac{\mathrm{d}V}{\mathrm{d}\hat{x}}. \tag{13.27}$$
We recast (13.25) noting the rate of change of momentum as acceleration times mass
$$\frac{\mathrm{d}\langle\hat{p}\rangle}{\mathrm{d}t} = -\left\langle\frac{\mathrm{d}V}{\mathrm{d}\hat{x}}\right\rangle \quad \longleftrightarrow \quad m\frac{\mathrm{d}^2\langle x\rangle}{\mathrm{d}t^2} = \langle F\rangle, \tag{13.28}$$
where the derivative of the potential is force F. We have recovered Newton's second law for expectation values! This expresses momentum conservation in the absence of external forces.

The classical limit of quantum mechanics is sometimes called the correspondence principle. In the limit that we take expectation values with averaged measurements, we recover the dynamics of classical physics. Conserved quantities are related to Noether's theorem in analytical mechanics, which states that every continuous symmetry of a system has an associated conserved quantity, discovered by Emmy Noether in 1918. Energy and momentum conservation is due to a symmetry in time and space translations, respectively.

HW7: Harmonic oscillator and conservation (set Oct 22, due Oct 29)

1. The harmonic oscillator Hamiltonian is $\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2$ and ladder operators are

$$\hat{A} = \frac{1}{\sqrt{2m\hbar\omega}}(m\omega x + \mathrm{i}\hat{p}), \quad \hat{A}^\dagger = \frac{1}{\sqrt{2m\hbar\omega}}(m\omega x - \mathrm{i}\hat{p}). \tag{13.29}$$

   (a) Taking the product $\hat{A}\hat{A}^\dagger$, show that the $\hat{H}$ can be written as $\hat{H} = \left(\hat{A}\hat{A}^\dagger - \frac{1}{2}\right)\hbar\omega$.

   (b) Show that the commutator $\hat{H}$ with the lowering operator is $[\hat{H}, \hat{A}] = -\hbar\omega\hat{A}$.

   (c) Use the previous result to show that $\hat{A}$ acting on $|E_n\rangle$ yields the eigenvalue equation $\hat{H}(\hat{A}|E_n\rangle) = (E_n - \hbar\omega)(\hat{A}|E_n\rangle)$. Comment on the physical significance.

   (d) Consider the operator $\hat{N} = \hat{A}^\dagger\hat{A}$ with $\hat{N}|n\rangle = n|n\rangle$ so we can label energy states by integers n e.g. $|n\rangle = |E_n\rangle, |n \pm 1\rangle = |E_n \pm \hbar\omega\rangle$. By taking the modulus squares $|\hat{A}|n\rangle|^2 = \langle n|\hat{A}^\dagger\hat{A}|n\rangle$ and $[\hat{A}, \hat{A}^\dagger] = 1$, show that $\alpha = \sqrt{n}$ in the eigenvalue equation

$$\hat{A}|n\rangle = \alpha|n-1\rangle. \tag{13.30}$$

   (e) Derive the ground state wavefunction $\langle x|E_0\rangle$ by solving a differential equation, and show that the excited state $\psi_n(x) = \langle x|E_n\rangle$ is $\langle x|E_1\rangle = Ax\mathrm{e}^{-x^2/4\ell^2}$, What is the constant A? Explain why the wavefunctions have well-defined parity.

2. For two non-commuting observables $[\hat{A}, \hat{B}] \neq 0$, what inequality must the product of their uncertainties satisfy? What is the physical significance of this? What is this inequality when considering position $\hat{x}$ and momentum $\hat{p}$? Explain physically why the zero-point energy of the harmonic oscillator is non-zero.

3. Consider an operator $\hat{Q}$ that is time dependent $\partial\hat{Q}/\partial t \neq 0$.

   (a) Taking the time derivative of its expectation value $\langle\hat{Q}\rangle = \langle\psi|\hat{Q}|\psi\rangle$ and using the Schrödinger equation, present the derivation for Ehrenfest's theorem

$$\frac{\mathrm{d}\langle\hat{Q}\rangle}{\mathrm{d}t} = \frac{\mathrm{i}}{\hbar}\langle\psi|[\hat{H}, \hat{Q}]|\psi\rangle + \left\langle\frac{\partial\hat{Q}}{\partial t}\right\rangle. \tag{13.31}$$

   (b) If an operator $\hat{A}$ is time independent $\partial\hat{A}/\partial t = 0$ and commutes with the Hamiltonian, what is the physical significance of this equation?

   (c) If the system is in an energy eigenstate $|E\rangle$ such that $\hat{H}|E\rangle = E|E\rangle$, show that $\langle\hat{A}\rangle$ is conserved even if it does not commute with the Hamiltonian $[\hat{A}, \hat{H}] \neq 0$.

   (d) For $\hat{Q} = \hat{H}$, what physically is conserved?

14 Mid-course review

We are approaching half way through this course and it is apt to review. We first introduced the foundational experiments motivating the theoretical concepts of the wavefunction and Schrödinger equation. We then studied one-dimensional solutions of Schrödinger equation with appropriate boundary conditions and various potentials $V(x)$. After building this intuition, we formalized these in the language of operators for quantum mechanics.

14.1 Experimental foundations

There are numerous foundational experiments in the late nineteenth and early twentieth centuries that motivate quantum physics.

- Electron diffraction experiments: shows electrons are wave behaviour in the Davisson-Germer and Thomson-Reid experiments. This expands the discovery of electron as a particle by J. J. Thomson. Particles have a de Broglie wavelength inversely proportional to its momentum $\lambda = h/p$.
- Photoelectric effect and Compton X-ray scattering: shows photons have particle-like behaviour. This builds on the wave behaviour of light established in the nineteenth century by Young, Fraunhofer. The energy carried by light is proportional to its frequency $E = hf$, which is the Planck-Einstein relation.
- Radioactivity studied by Becquerel, the Curies showed an intrinsically probabilistic effect and evidence for high-energy scales of the nucleus inside atoms.
- Geiger-Marsden-Rutherford gold foil experiments directly probed the existence of atomic nucleus.
- Discrete line spectra measured by spectroscopists then empirically described by Balmer and Rydberg formulas indirectly before the semi-classical Bohr model. Franck-Hertz experiment: electron scattering experiments directly probed discrete energy levels of atomic bound states and the transitions are probabilistic.

14.2 Formalism and postulates

We can restate what we have learnt as the key postulates of quantum mechanics:

1. **Wavefunction postulate**. Quantum states are described by a wavefunction state vector $|\psi\rangle$ that is square integrable for normalizability.

2. **Operator postulate**. Observables are eigenvalues q_n of Hermitian operators $\hat{Q}$ according to the eigenvalue equation $\hat{Q}|\phi_n\rangle = q_n|\phi_n\rangle$ for eigenstate $|\phi_n\rangle$.

3. **Measurement postulate**. Given an initial state $|\psi\rangle$ can be expanded in terms of an eigenbasis $\sum_n q_n|\phi_n\rangle$, the probability amplitude that the state is in $|\phi_n\rangle$ upon measurement is given by the Born rule

$$P(\phi_n|\psi) = |\langle\phi_n|\psi\rangle|^2. \tag{14.1}$$

4. **Time evolution postulate**: the wavefunction $|\psi\rangle$ evolves smoothly in time according to Schrödinger's equation according to the Hamiltonian $\hat{H}$

$$\mathrm{i}\hbar\frac{\partial|\psi\rangle}{\partial t} = \hat{H}|\psi\rangle. \tag{14.2}$$

   If $|\psi(t)\rangle$ is an energy E eigenstate, then the wavefunction evolves by a phase factor given by the energy $|\psi(t)\rangle = |\psi(0)\rangle \mathrm{e}^{-\mathrm{i}Et/\hbar}$.

This lecture plans to review key concepts and worked examples from previous classes on the one-dimensional solutions to Schrödinger equation and operator formalism with Dirac notation.

15 Interpretations (optional reading)

This class is when mid-term exam 1 takes place.

> These notes instead provide some brief qualitative discussion as an invitation to the philosophical issues of quantum mechanics. *The material in section 15 is optional reading and not part of this course*.

These notes serve to introduce phrases as invitations to further reading. Philosophers will have to forgive the somewhat slippery use of language. For those interested in precise and detailed discussion, NYU has a renowned philosophy program that offers relevant classes[5]:

- Philosophy of Science (PHIL-UA 90). This course is regularly offered, which discusses empiricism and the scientific method of knowledge acquisition.
- Philosophy of Physics (PHIL-UA 94). This course is periodically offered and sometimes discusses philosophical issues of quantum mechanics.

Metaphysics Heuristically, metaphysics studies states of reality (ontology) and knowledge (epistemology). The metaphysical status of the wavefunction is debated.

- **Ontology**: do wavefunctions *exist in external reality* as a physical entity? Ontic positions prefer the wavefunction as representing real states existing in nature.
- **Epistemology**: do wavefunctions reflect an *observer's knowledge* of a physical entity? Epistemic positions prefer the wavefunction as representing knowledge about states.

Metaphysical debates are not unique to quantum mechanics. The history of science is filled with arguments about whether atoms and magnetic fields are real objects in nature or just mathematical conveniences to describe chemical reactions and compass needles.

The Schrödinger cat thought experiment, conceived in 1935, illustrates how quantum superposition and measurement defies our classical intuition with macroscopic objects. A cat is sleeping inside a box with a radioactive cobalt-60 atom. If the nucleus emits an electron, this fires a Geiger counter that triggers an alarm clock and wakes up the cat. The cat being awake or asleep depends on the quantum mechanical superposition state of the nucleus being decayed or not. When does the wavefunction collapse occur? Is it opening the box or the moment an observer inspects its interior that triggers wavefunction collapse?

[5] https://as.nyu.edu/departments/philosophy/undergraduate/course-offerings.html

Hidden variables and determinism Quantum mechanics has a deterministic and a non-deterministic postulate. The continuous time evolution of the Schrödinger equation *is* deterministic. By contrast, the wavefunction collapse is discontinuous and non-deterministic. The latter non-determinism receives objections that claim quantum theory as being *incomplete*. At the Institute of Advanced Studies in Princeton, Albert Einstein, Boris Podolsky, and Nathan Rosen critiqued this with a though experiment in 1935 [34], which we revisit in section 22.4. They objected to the instantaneous collapse of the wavefunction that seems to non-locally determine another observer's measured state.

Hidden-variable theories propose that there exists information, which fully determined the states from the start. Observers simply lack access to that knowledge. This restores probabilistic uncertainty to the same epistemological status as classical probabilities of coin flips. Hidden-variables theories return predictions to the clockwork determinism of classical physics. In the mid-1930s, John von Neumann attempted to refute hidden-variable theories, but Grete Hermann pointed out a loophole. It took until 1964 for John Bell to publish his eponymous inequality that shows it is possible to distinguish quantum measurements from local hidden-variable theories [35], which section 23 briefly discusses. Hidden-variable theories that assume locality are experimentally disfavoured in quantum systems.

Many worlds The many-worlds interpretation proposed by Hugh Everett treats every outcome of the wavefunction as real. It rejects the notion that there is only one outcome in a measurement. It is a strong ontology, where every outcome of the superposition states is real. The outcomes correspond to multiple worlds, each hosting one of the possible eigenstates that really do exist. Information from these different universes are inaccessible but are interpreted to really exist. This poses epistemological challenges about our knowledge of their existence and empiricism of observing states in an inaccessible worlds.

Objective observers and empiricism Empiricism underpins science, which gained prominence in European intellectual thought at the turn of the seventeenth century. Notable advocates in the English renaissance include Francis Bacon on philosophy, William Gilbert on magnetism, and William Harvey on physiology. It separates the observer as independent from the observed. Knowledge is acquired by an observer reporting their experience of phenomena in objective, dispassionate accounts who is sceptical of their own subjective biases. In classical empiricism, we can in principle measure e.g. the position and velocity of a football without disturbing those observables. Another observer can reproduce the measurements.

Quantum mechanics challenges this epistemology. For a radioactive nucleus in a superposition state, one observer can observe the "decayed" outcome while another can observe

the other outcome "not decayed". Quantum outcomes depend on who looks. There appears to be *subjectivity* in measurements, which questions the notion of an objective observer of classical empiricism. We can even ask what is an observer? When an ionization chamber detects an electron, this generates a voltage pulse to a voltmeter that an experimentalist then observes before writing up a report to share with other scientists. Which of these steps is the measurement that triggers the wavefunction collapse? Is it the instrument or the human? Microscopic quantum observables transition to macroscopic classical observer, and quantum decoherence attempts to address this.

Copenhagen and instrumentalism The Copenhagen interpretation is held widely among physicists [36]. It generally interprets an electron wavefunction as being a mathematical tool and has no reality until an electron is measured. This epistemic interpretation of the wavefunction captures observer ignorance while making very accurate predictions of probabilities. There is no deeper explanation for the transition from superposition to collapsed state. Some find this position unsatisfying. But it is no surprise that many practising physicists adopt a "keep calm and don't worry" pragmatism to get on with their exquisite experiments.

Philosophy of science discusses criteria for accepting scientific theories. Various schools of thought argue theories should have elements of simple axioms, predictability, falsifiability, causal mechanism, reproducibility, and wide-ranging empirical evidence. But perhaps those criteria are too strict, built on a classical intuition of objective empiricism. The measurement postulate questions criteria such as predictability and causal mechanism.

Quantum mechanics suggests a revised empiricism with relaxed notions of predictability (only probabilities are predictable) and observer independence (individual outcomes depend on observer but not in aggregate). Relaxing these requirements succeeds empirically where classical physics fails. Indeed science has repeatedly humbled us. The Copernican principle challenged the geocentrism assumption as too strong. Relativity argued the absolute space-time assumption as too strong and introduces observer dependence of clocks and rulers. Perhaps causal determinism was too strict an assumption.

A better question is: once we relax determinism and accept probabilistic outcomes, why does quantum mechanics describe nature so remarkably well? Instrumentalism judges theories by their ability to describe experimental results. Quantum mechanics has spectacular empirical support. It predicts the electron magnetic dipole agreeing with measurements to 13 decimal places, underpins modern computers, and saves lives with medical imaging. This success elevates quantum mechanics into the prevailing scientific paradigm. It intimates that nature is fundamentally probabilistic. Classical determinism appears emergent.

IV Spherical systems

16 Orbital angluar momentum

Our quantum endgame is to describe atoms. We must therefore solve Schrödinger's equation in three dimensions by extending the position $x \to \mathbf{r}$ and momentum to vectors with the gradient operator $\boldsymbol{\nabla}$

$$p = -\mathrm{i}\hbar\frac{\partial}{\partial x} \to \mathbf{p} = -\mathrm{i}\hbar\boldsymbol{\nabla}, \quad \boldsymbol{\nabla} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}. \tag{16.1}$$

The double partial derivative becomes the Laplacian operator $\nabla^2 = \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}$

$$\frac{\partial^2}{\partial x^2} \to \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \tag{16.2}$$

This enables us to write the Hamiltonian of the time-independent Schrödinger equation as

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}). \tag{16.3}$$

16.1 Central potentials

In this course, we only consider central potentials $V(\mathbf{r}) = V(r), r = |\mathbf{r}|$ with spherical symmetry. Such potentials are familiar from classical orbital mechanics with the gravitational potential (see PHYS-UA 120 Dynamics)

$$V_{\mathrm{grav}}(r) = \frac{GM_1M_2}{r^2}. \tag{16.4}$$

The classical angular momentum $\mathbf{L}$ is defined by the cross product of position $\mathbf{r}$ and linear momentum $\mathbf{p} = m\mathbf{v}$ vectors (figure 29)

$$\mathbf{L} = \mathbf{r}\times\mathbf{p} = \mathbf{r}\times(m\mathbf{v}). \tag{16.5}$$

Crucially, central potentials imply the angular momentum is a conserved quantity

$$\frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t} = m(\dot{\mathbf{r}}\times\dot{\mathbf{r}} + \mathbf{r}\times\ddot{\mathbf{r}}) = \mathbf{0}. \tag{16.6}$$

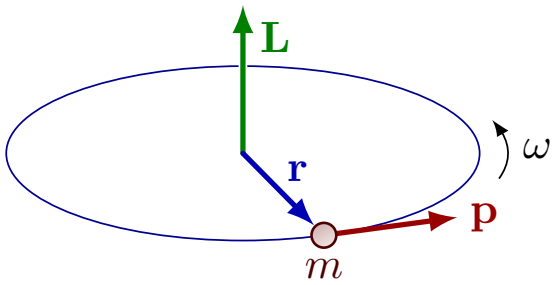


Figure 29: Visualization of orbital angular momentum **L**. Figure: Izaak Neutelings.

This arises because the force $m\ddot{\mathbf{r}}$ is parallel to $\mathbf{r}$ and perpendicular to linear momentum $\mathbf{p} = m\dot{\mathbf{r}}$. For Earth with mass M_E orbiting the Sun, the classical Hamiltonian $H = T + V$ is

$$H = \underbrace{\frac{1}{2}M_E\dot{r}^2}_{T_{\text{rad}}(r)} + \underbrace{\frac{L^2}{2M_E r} + V_{\text{grav}}(r)}_{V_{\text{eff}}(r)}. \tag{16.7}$$

This treats Earth moving only radially with kinetic energy $T(\dot{r})$ in an effective potential $V_{\text{eff}}(r)$, which has an angular momentum barrier $L^2/2M_E r$ and gravitational potential $V_{\text{grav}}(r)$. Spherical symmetry both simplifies the problem and illuminates the physics.

The quantum mechanics of an electron in a hydrogen proceeds similarly, where we model the proton as a static Coulomb potential

$$V(r) = -\frac{e^2}{4\pi\varepsilon_0 r}. \tag{16.8}$$

The next few sections develop the analogous angular momentum methods to solve the Schrödinger equation $\hat{H}\psi = E\psi$ for an electron wavefunction $\psi(\mathbf{r})$ and energies E in three dimensions

$$\left[-\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\varepsilon_0 r}\right]\psi(\mathbf{r}) = E\psi(\mathbf{r}). \tag{16.9}$$

16.2 Angular momentum commutators

We turn angular momentum into a quantum observable using operators with derivatives

$$\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}} = -i\hbar\hat{\mathbf{r}} \times \boldsymbol{\nabla} \tag{16.10}$$

The cross product is written explicitly as a determinant

$$\hat{\mathbf{L}} = -\mathrm{i}\hbar \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x & y & z \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \end{vmatrix}. \tag{16.11}$$

We can write out each component of $\hat{\mathbf{L}}$ explicitly to look at:

$$\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \tag{16.12}$$
$$\hat{L}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z, \tag{16.13}$$
$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x. \tag{16.14}$$

In differential operator form, these look like

$$\hat{L}_x = -\mathrm{i}\hbar \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y} \right), \tag{16.15}$$
$$\hat{L}_y = -\mathrm{i}\hbar \left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z} \right), \tag{16.16}$$
$$\hat{L}_z = -\mathrm{i}\hbar \left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} \right). \tag{16.17}$$

Due to the commutator $[\hat{x}_i, \hat{p}_j] = \mathrm{i}\hbar\delta_{ij}$, these angular momentum operators do not commute

$$[\hat{L}_x, \hat{L}_y] = \mathrm{i}\hbar\hat{L}_z, \quad [\hat{L}_y, \hat{L}_z] = \mathrm{i}\hbar\hat{L}_x, \quad [\hat{L}_z, \hat{L}_x] = \mathrm{i}\hbar\hat{L}_y. \tag{16.18}$$

If we stare at these results closely, we find the $\hat{L}_i$ operator by cycling through the coordinates $x \to y \to z \to x$. So we can write (16.17) and (16.18) compactly as

$$\hat{L}_i = -\mathrm{i}\hbar\varepsilon_{ijk} x_j \frac{\partial}{\partial x_k}, \tag{16.19}$$
$$[\hat{L}_i, \hat{L}_j] = \mathrm{i}\hbar\varepsilon_{ijk}\hat{L}_z. \tag{16.20}$$

The ε_{ijk} object is called the Levi-Civita symbol, which is defined by

$$\varepsilon_{ijk} = \begin{cases} +1, & \text{even permutations of } (i,j,k), \\ -1, & \text{odd permutations of } (i,j,k), \\ 0, & \text{otherwise.} \end{cases} \tag{16.21}$$

For example, ε_{ijk} is positive unity for $(i,j,k) = (1,2,3)$ and $(2,3,1)$ because it requires two swaps to make them the same. The non-commutation of $\hat{L}_i$ has serious physical consequences. If we measured $\hat{L}_z$, then the angular momenta projected in the x direction and the y direction are not simultaneous eigenstates, i.e. $\hat{L}_x$ and $\hat{L}_y$ are not well defined.

16.3 Total magnitude observable

Another important operator is the total magnitude square of the angular momentum

$$\hat{L}^2 = \hat{\mathbf{L}} \cdot \hat{\mathbf{L}} = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2. \tag{16.22}$$

Crucially, this operator *does* commute with each of the angular momentum components:

$$[\hat{L}^2, \hat{L}_i] = 0. \tag{16.23}$$

A state can therefore be a simultaneous eigenstate of both total square $\hat{L^2}$ and azimuthal angular momentum $\hat{L}_z$. This is so important, we prove $[\hat{L}^2, \hat{L}_x] = 0$ by writing out $\hat{L}^2$ in full

$$[\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2, \hat{L}_x] = [\hat{L}_x^2, \hat{L}_x] + [\hat{L}_y^2, \hat{L}_x] + [\hat{L}_z^2, \hat{L}_x]. \tag{16.24}$$

Invoke the operator identity (11.18) to write

$$[\hat{A}^2, \hat{B}] = \hat{A}[\hat{A}, \hat{B}] - [\hat{A}, \hat{B}]\hat{A}, \tag{16.25}$$

which one can prove by expanding out each commutator. Each angular momentum operator commutes with itself

$$[\hat{L}_i, \hat{L}_i] = 0 \quad \Rightarrow \quad [\hat{L}_x^2, \hat{L}_x] = 0. \tag{16.26}$$

For the next commutator $[\hat{L}_y^2, \hat{L}_x]$, we explicitly expand this out to find

$$[\hat{L}_y^2, \hat{L}_x] = \hat{L}_y[\hat{L}_y, \hat{L}_x] - [\hat{L}_y, \hat{L}_x]\hat{L}_y = -\mathrm{i}\hbar(\hat{L}_y\hat{L}_z + \hat{L}_z\hat{L}_y), \tag{16.27}$$

where we evaluate each commutator using equation (16.18). Repeating for the final commutator involving $\hat{L}_z^2$, we find it exactly cancels $[\hat{L}_y^2, \hat{L}_x]$

$$[\hat{L}_z^2, \hat{L}_x] = \hat{L}_z[\hat{L}_z, \hat{L}_x] - [\hat{L}_z, \hat{L}_x]\hat{L}_z = +\mathrm{i}\hbar(\hat{L}_z\hat{L}_y + \hat{L}_y\hat{L}_z). \tag{16.28}$$

This proves the claim (16.23), where the other components follow by cyclicity. Given we can only measure $\hat{L}^2$ and $\hat{L}_z$ simultaneously, this implies the common eigenstates $|\lambda, m\rangle$

$$\hat{L}^2|\lambda, m\rangle = \hbar^2\lambda|\lambda, m\rangle, \tag{16.29}$$

$$\hat{L}_z|\lambda, m\rangle = \hbar m|\lambda, m\rangle. \tag{16.30}$$

Next we need to find the angular momentum eigenvalues λ, m.

17 Quantized angular momentum

17.1 Ladder operators

We can use algebra to determine the eigenvalues λ, m without any calculus. We follow the harmonic oscillator (section 12) and define angular momentum ladder operators

$$\hat{L}_\pm = \hat{L}_x \pm \mathrm{i}\hat{L}_y. \tag{17.1}$$

Consider the commutation of the raising operator $\hat{L}_+$ with $\hat{L}_z$

$$[\hat{L}_z, \hat{L}_+] = [\hat{L}_z, \hat{L}_x] + \mathrm{i}[\hat{L}_z, \hat{L}_y] = \mathrm{i}\hbar\hat{L}_y + \hbar\hat{L}_x = +\hbar\hat{L}_+. \tag{17.2}$$

Because $\hat{L}_+$ are linear combinations of $\hat{L}_x$ and $\hat{L}_y$, the ladder operators commute with $\hat{L}^2$

$$[\hat{L}^2, \hat{L}_+] = 0. \tag{17.3}$$

So $\hat{L}^2$ and $\hat{L}_+$ share simultaneous eigenstates. Operating $\hat{L}_z$ on the $\left(\hat{L}_+|\lambda, m\rangle\right)$ state gives

$$\hat{L}_z\left(\hat{L}_+|\lambda, m\rangle\right) = (\hat{L}_+\hat{L}_z + [\hat{L}_z, \hat{L}_+])|\lambda, m\rangle = \hbar(m+1)\left(\hat{L}_+|\lambda, m\rangle\right). \tag{17.4}$$

We see that $(\hat{L}_+|\lambda, m\rangle)$ is also an eigenstate of the $\hat{L}_z$ operator, except the eigenvalues differ by one unit of $\hbar$. This means that $\hat{L}_+$ acting on $|\lambda, m\rangle$ is proportional to the state $|\lambda, m+1\rangle$

$$\hat{L}_+|\lambda, m\rangle = \alpha_+\hbar|\lambda, m+1\rangle. \tag{17.5}$$

The proportionality constant α_+ is the eigenvalue we wish to evaluate. To do this, we take the modulus square of both sides of the equation to give

$$\hbar^2|\alpha_+|^2 = \langle\lambda, m|\hat{L}_+^\dagger\hat{L}_+|\lambda, m\rangle. \tag{17.6}$$

Now we perform algebra on the combination $\hat{L}_+^\dagger\hat{L}_+$. Take care to track signs when expanding the parentheses as the order of $\hat{L}_i$ operators matters, we expand this out

$$\begin{aligned}
\hat{L}_+^\dagger\hat{L}_+ &= (\hat{L}_x - \mathrm{i}\hat{L}_y)(\hat{L}_x + \mathrm{i}\hat{L}_y) && (17.7)\\
&= \hat{L}_x\hat{L}_x + \hat{L}_y\hat{L}_y + \mathrm{i}\hat{L}_x\hat{L}_y - \mathrm{i}\hat{L}_y\hat{L}_x && (17.8)\\
&= \hat{L}^2 - \hat{L}_z^2 + \mathrm{i}[\hat{L}_x, \hat{L}_y] && (17.9)\\
&= \hat{L}^2 - \hat{L}_z^2 - \hbar\hat{L}_z. && (17.10)
\end{aligned}$$

Applying this to equation (17.6) and the eigenvalue equations for $\hat{L}^2$ and $\hat{L}_z$ (16.30) gives

$$\begin{aligned}
\hbar^2|\alpha_+|^2 &= \langle\lambda, m|(\hat{L}^2 - \hat{L}_z^2 - \hbar\hat{L}_z)|\lambda, m\rangle && (17.11)\\
&= \hbar^2(\lambda - m^2 - m)\langle\lambda, m|\lambda, m\rangle. && (17.12)
\end{aligned}$$

The eigenvalues of $\hat{L}_+$ and analogously for $\hat{L}_-$ are

$$\alpha_\pm = \sqrt{\lambda - m(m\pm 1)}. \tag{17.13}$$

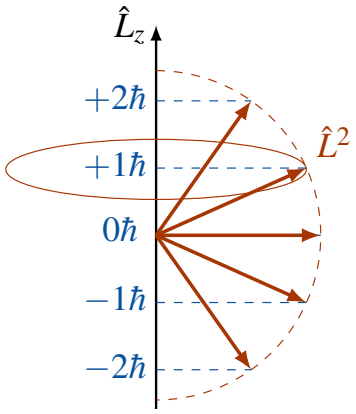


Figure 30: Quantized total angular momentum $\hat{L}^2$ and its azimuthal projection $\hat{L}_z$.

17.2 Angular momentum eigenvalues

We can now determine the eigenvalues of $\hat{L}_z$ and $\hat{L}^2$ to solve equations (16.30). The projection of a vector onto the z component cannot be larger than its magnitude. So there must exist a maximum state where the raising operation vanishes $\hat{L}_+|\lambda, m_{\max}\rangle = 0$. This implies $\alpha_+ = 0$ when m reaches a maximum value $m_{\max}$

$$\lambda - m_{\max}(m_{\max}+1) = 0. \tag{17.14}$$

Similarly, repeatedly applying $\hat{L}_-$ must reach the ladder's bottom and cause the minimum state to vanish $\hat{L}_-|\lambda, m_{\min}\rangle = 0$. This implies $\alpha_- = 0$ when m reaches a minimum value

$$\lambda - m_{\min}(m_{\min}-1) = 0. \tag{17.15}$$

Eliminating α from these two equations and solving the quadratic equation in $m_{\min}$ gives

$$m_{\min} = \frac{1}{2}[1 \pm (2m_{\max}+1)]. \tag{17.16}$$

Only the negative solution makes sense because a minimum must of course be lower than a maximum $m_{\min} < m_{\max}$. This implies $m_{\min} = -m_{\max}$. Defining $\ell = m_{\max}$, we find

$$\lambda = \ell(\ell+1), \quad \text{where } -\ell \le m \le \ell. \tag{17.17}$$

We can only apply $\hat{L}_+$ an integer number of times to get from $m = -\ell$ to $m = +\ell$ because it makes no sense to operate $\hat{L}_+$ on a state a fraction number of times. Therefore 2ℓ must be an integer, so ℓ takes half-integer values

$$\ell = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots \tag{17.18}$$

Depending on the value of ℓ, there are a total of $(2\ell+1)$ values of m shown in figure 30:

$$m = -\ell, -\ell+1, \ldots, 0, \ell-1, \ell. \tag{17.19}$$

To summarize, the ladder operators change the m quantum number by integers

$$\hat{L}_\pm |\ell, m\rangle = \sqrt{\ell(\ell+1) - m(m\pm 1)}\,|\ell, m\pm 1\rangle. \tag{17.20}$$

The general simultaneous eigenstates and eigevalues for the $\hat{L}^2$ and $\hat{L}_z$ operators are

$$\hat{L}^2 |\ell, m\rangle = \hbar^2 \ell(\ell+1)|\ell, m\rangle, \quad \ell = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots \tag{17.21}$$

$$\hat{L}_z |\ell, m\rangle = \hbar m |\ell, m\rangle, \qquad m = -\ell, -\ell+1, \ldots, 0, \ell-1, \ell. \tag{17.22}$$

We find the eigenvalues are allowed to be half-integer quantized in units of $\hbar$. We can trace this back to the commutation relations of $[\hat{L}_x, \hat{L}_y] = \mathrm{i}\hbar \hat{L}_z$, which prevents us from simultaneously measuring the other two projections of angular momentum $\hat{L}_x, \hat{L}_y$.

17.3 Spherical coordinates

Given central potentials have spherical symmetry, it is natural to transform $\hat{L}_i$ (16.17) from Cartesian to spherical coordinates via trigonometry (figure 31)

$$x = r\sin\theta\cos\phi, \tag{17.23}$$

$$y = r\sin\theta\sin\phi, \tag{17.24}$$

$$z = r\cos\theta. \tag{17.25}$$

Like how we map Earth's surface, the polar angle swivels from $0 \le \theta < \pi$ from the north to south pole along lines of longitude. The azimuthal angle swivels about the polar axis from $0 \le \phi < 2\pi$ along lines of latitude. From the momentum operator $\hat{\mathbf{p}} = -\mathrm{i}\hbar\boldsymbol{\nabla}$, we convert derivatives from Cartesian (x, y, z) to spherical (r, θ, ϕ) coordinates via the chain rule

$$\frac{\partial}{\partial x} = \frac{\partial r}{\partial x}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial x}\frac{\partial}{\partial \theta} + \frac{\partial \phi}{\partial x}\frac{\partial}{\partial \phi}, \tag{17.26}$$

$$= \sin\theta\cos\phi\frac{\partial}{\partial r} + \cos\theta\cos\phi\frac{1}{r}\frac{\partial}{\partial\theta} - \frac{\sin\phi}{\sin\theta}\frac{1}{r}\frac{\partial}{\partial\phi}. \tag{17.27}$$

Similarly doing so for $\partial/\partial y$ and $\partial/\partial z$, we find the angular momentum operators become

$$\hat{L}_x = \mathrm{i}\hbar\left(\sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right), \tag{17.28}$$

$$\hat{L}_y = \mathrm{i}\hbar\left(-\cos\phi\frac{\partial}{\partial\theta} + \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right), \tag{17.29}$$

$$\hat{L}_z = -\mathrm{i}\hbar\frac{\partial}{\partial\phi}. \tag{17.30}$$

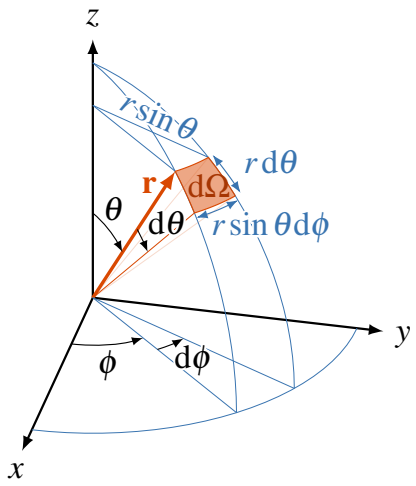


Figure 31: Spherical coordinates (r,θ,ϕ) in terms of Cartesian coordinates (x,y,z).

This allows us to write the magnitude square of the angular momentum operator as

$$\hat{L}^2 = \hat{\mathbf{L}}\cdot\hat{\mathbf{L}} = -\hbar^2\left[\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]. \tag{17.31}$$

Note the similarity with the Laplacian operator in spherical coordinates

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]. \tag{17.32}$$

The square brackets is exactly $-\hat{L}^2/\hbar^2$, enabling us to write the Schrödinger equation as

$$-\frac{\hbar^2}{2m_e}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{\hat{L}^2}{\hbar^2 r^2}\right]\psi + V(r)\psi = E\psi. \tag{17.33}$$

We next solve the angular and radial parts of this differential equations in turn.

HW8: Angular momentum (set Oct 29, due Nov 5)

1. The projection of the orbital angular momentum operator along the x axis is

$$\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y. \tag{17.34}$$

   (a) By using the operator identity $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$ and canonical commutation relation $[\hat{x}_i, \hat{p}_j] = \mathrm{i}\hbar\delta_{ij}$, show that $\hat{L}_x$ is Hermitian. Why does this imply that $\hat{L}^2$ is Hermitian?

   (b) By writing $[\hat{L}_x, \hat{L}_y]$ as $[\hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \hat{z}\hat{p}_x - \hat{x}\hat{p}_z]$, prove the commutator

$$[\hat{L}_x, \hat{L}_y] = \mathrm{i}\hbar\hat{L}_z. \tag{17.35}$$

   Discuss its physical significance. Hint: the $\mathrm{i}\hbar$ arises from the commutator $[\hat{z}, \hat{p}_z]$.

   (c) By explicitly expanding the left- and right-hand side, prove the commutator identity $[\hat{A}^2, \hat{B}] = \hat{A}[\hat{A}, \hat{B}] + [\hat{A}, \hat{B}]\hat{A}$. By writing out the definition of $\hat{L}^2$, use this to prove that $[\hat{L}^2, \hat{L}_y] = 0$. What is the physical significance of this?

2. Write down the definition of the angular momentum lowering operator $\hat{L}_-$.

   (a) Prove that $\hat{L}_-$ satisfies the commutator $[\hat{L}_z, \hat{L}_-] = -\hbar\hat{L}_-$.

   (b) Why does the lowering operator commute with the $\hat{L}^2$ i.e. $[\hat{L}^2, \hat{L}_-] = 0$?

   (c) Show that $|\psi\rangle = \hat{L}_-|\lambda, m\rangle$ is an eigenstate of $\hat{L}_z$ and find its eigenvalue.

   (d) Write down the eigenvalue equations for $\hat{L}^2$ and $\hat{L}_z$. Show that

$$\hat{L}_-^\dagger\hat{L}_- = \hat{L}^2 - \hat{L}_z^2 + \hbar\hat{L}_z. \tag{17.36}$$

   Use this to show the eigenvalue equation for the lowering operator is

$$\hat{L}_-|\lambda, m\rangle = \alpha_-|\lambda, m-1\rangle. \tag{17.37}$$

   Find the expression for α_-. Comment on the physical significance of this result.

3. Consider spherical coordinates $x = r\sin\theta\cos\phi, y = r\sin\theta\sin\phi, z = r\cos\theta$. Apply the chain rule to show that the $\hat{L}_z$ operator in spherical coordinates is

$$\hat{L}_z = -\mathrm{i}\hbar\frac{\partial}{\partial\phi}. \tag{17.38}$$

18 Angular Schrödinger equation

The Schrödinger equation written with radial derivatives and and angular momentum $\hat{L}^2$ as

$$-\frac{\hbar^2}{2m_e}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)-\frac{\hat{L}^2}{\hbar^2 r^2}\right]\psi+V(r)\psi=E\psi. \tag{18.1}$$

This is analogous to orbital dynamics but with an electron moving along r in an effective potential

$$V_{\text{eff}}(r)=V(r)-\frac{\hat{L}^2}{2m_e r^2}. \tag{18.2}$$

The angular momentum operators only depend on (θ,ϕ) while the Hamiltonian only depends on r, so they automatically commute

$$[\hat{H},\hat{L}^2]=0,\quad [\hat{H},\hat{L}_i]=0. \tag{18.3}$$

Quantum mechanically, this means a state can simultaneously have well-defined energy E and angular momentum ℓ,m eigenvalues. Mathematically, the solutions are separable

$$\psi(r,\theta,\phi)=R(r)Y(\theta,\phi). \tag{18.4}$$

Multiplying equation (18.1) through by $-2m_e r^2/(RY)$ gives

$$\frac{1}{R}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right)+\frac{2m_e r^2}{\hbar^2}(E-V)=\frac{\hat{L}^2 Y}{\hbar^2 Y}=\text{constant}. \tag{18.5}$$

The left-hand side depends only on r and right-hand side depends only on (θ,ϕ), so must be equal to a constant λ. We first focus on the angular solutions called spherical harmonics $Y(\theta,\phi)$, which is independent on the form of the potential $V(r)$.

18.1 Spherical harmonics

We seek simultaneous eigenfunctions for the two angular momentum eigenvalue equations

$$\hat{L}^2 Y(\theta,\phi)=\hbar^2\lambda Y(\theta,\phi), \tag{18.6}$$

$$\hat{L}_z Y(\theta,\phi)=\hbar m Y(\theta,\phi). \tag{18.7}$$

We already found quantized eigenvalues in equation (17.22) using algebra. Let us examine $\hat{L}_z Y=\hbar m Y$ using the differential operator $\hat{L}_z=-\mathrm{i}\hbar\partial/\partial\phi$ from (17.30)

$$-\mathrm{i}\hbar\frac{\partial Y}{\partial\phi}=\hbar m Y. \tag{18.8}$$

Integrating with respect to ϕ yields an exponential phase $\Phi(\phi) = e^{im\phi}$ multiplied by a θ-dependent function $\Theta(\theta)$ as an "integration constant in ϕ"

$$Y(\theta,\phi) = \Theta(\theta)e^{im\phi}. \tag{18.9}$$

The takes the form $Y(\theta,\phi) = \Theta(\theta)\Phi(\phi)$ with $\Phi(\phi) = e^{im\phi}$. Like flying around Earth's equator, the ϕ coordinate is periodic, giving the boundary condition $\Phi(\phi) = \Phi(\phi + 2\pi)$:

$$e^{im\phi} = e^{im\phi}e^{2\pi im} \quad \Rightarrow \quad e^{2\pi im} = 1. \tag{18.10}$$

This boundary condition requires m to be integer quantized

$$m = \ldots, -2, -1, 0, 1, 2, \ldots \tag{18.11}$$

18.2 Polar differential equation

Textbooks then offer two approaches to solving $\Theta(\theta)$:

1. **Ladder operator approach**. We can write the raising operator in spherical coordinates using $\hat{L}_x, \hat{L}_y$ from equations (17.28) and (17.29)

   $$\hat{L}_+ = \hat{L}_x + i\hat{L}_y = \hbar e^{i\phi}\left(\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right). \tag{18.12}$$

   When this acts on the top eigenstate, it vanishes

   $$\hat{L}_+|\ell, m_{\max} = \ell\rangle = 0 \tag{18.13}$$

   So we can write $\hat{L}_+ Y_\ell^\ell(\theta,\phi) = 0$ in spherical coordinates to give the differential equation

   $$\left(\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right)\Theta(\theta)e^{i\ell\phi} = 0. \tag{18.14}$$

   Performing the $\partial/\partial\phi$ differentiation gives a differential equation in θ

   $$\frac{\partial\Theta}{\partial\theta} = (\ell\cot\theta)\Theta. \tag{18.15}$$

   Integrating with respect to θ gives $\Theta = C(\sin\theta)^\ell$, where C is a constant. The full $Y_\ell^\ell(\theta,\phi)$ solution is then (up to a normalization constant A)

   $$Y_\ell^\ell(\theta,\phi) = Ae^{i\ell\phi}(\sin\theta)^\ell. \tag{18.16}$$

   To find the other solutions for $m = \ell - 1$, we repeatedly apply the lowering operator

   $$\hat{L}_- = \hat{L}_x - i\hat{L}_y = \hbar e^{-i\phi}\left(-\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right). \tag{18.17}$$

$Y_\ell^m(\theta,\phi)$	$m=0$	$m=\pm1$	$m=\pm2$
$\ell=0$	$\sqrt{\frac{1}{4\pi}}$		
$\ell=1$	$\sqrt{\frac{3}{4\pi}}\cos\theta$	$\mp\sqrt{\frac{3}{8\pi}}e^{\pm i\phi}\sin\theta$	
$\ell=2$	$\sqrt{\frac{5}{16\pi}}(3\cos^2\theta-1)$	$\mp\sqrt{\frac{15}{8\pi}}e^{\pm i\phi}\cos\theta\sin\theta$	$\sqrt{\frac{15}{32\pi}}e^{\pm 2i\phi}\sin^2\theta$

Table 2: The first few spherical harmonics for $\ell=0,1,2$.

2. **Legendre differential equation approach**. Other textbooks plug $Y(\theta,\phi)=\Theta(\theta)e^{im\phi}$ into the equation $\hat{L}^2Y(\theta,\phi)=\hbar^2\lambda Y(\theta,\phi)$ to give

$$-\frac{1}{\sin^2\theta}\left[\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)-m^2\right]\Theta=\lambda\Theta. \tag{18.18}$$

This is a second-order differential equation in the form of the *associated Legendre*[6] *equation*. Then open a favourite differential equations textbook to find the solutions are the *associated Legendre polynomials* with degree $|m|$

$$Y_\ell^m(\theta)=P_{\ell,m}(\cos\theta)=(1-x^2)^{|m|/2}\frac{d^{|m|}P_\ell(x)}{dx^{|m|}}. \tag{18.19}$$

Here, $P_\ell(x)$ are the *Legendre polynomials* that we can differentiate $|m|$ number of times. They are the solutions to the Legendre differential equation

$$\frac{d}{dx}\left[(1-x^2)\frac{dP_\ell}{dx}\right]+\ell(\ell+1)P_\ell(x)=0,\quad \ell=0,1,2,\ldots \tag{18.20}$$

The eigenvalue corresponding to the eigenfunction is

$$\lambda=\ell(\ell+1),\quad \text{with}\quad m=-\ell,-\ell+1,\ldots,0,\ell-1,\ell. \tag{18.21}$$

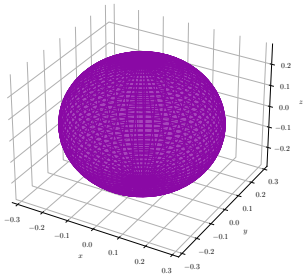
This constraint arises mathematically because we can only differentiate the $P_\ell(x)$ polynomials ℓ number of times before they vanish.

Collecting everything together, the angular momentum eigenvalue equations are

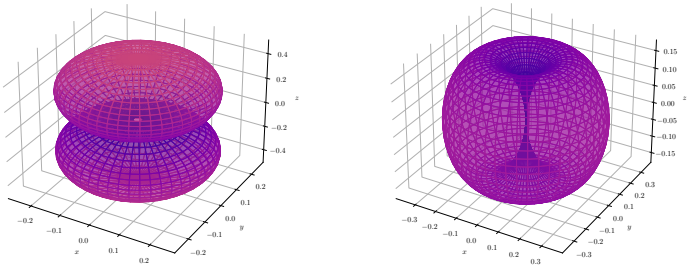
$$\hat{L}^2Y_\ell^m(\theta,\phi)=\hbar^2\ell(\ell+1)Y_\ell^m(\theta,\phi), \tag{18.22}$$
$$\hat{L}_zY_\ell^m(\theta,\phi)=\hbar mY_\ell^m(\theta,\phi). \tag{18.23}$$

[6]Named after Adrien-Marie Legendre (1752–1833) not French politician Louis Legendre (1752–1797), whose portrait was often erroneously used until 2008 in a curious case of mistaken identity.

$|Y_\ell^m(\theta,\phi)|,\ \ell=0,\ m=0$



$|Y_\ell^m(\theta,\phi)|,\ \ell=1,\ m=0$ $|Y_\ell^m(\theta,\phi)|,\ \ell=1,\ m=1$



$|Y_\ell^m(\theta,\phi)|,\ \ell=2,\ m=0$ $|Y_\ell^m(\theta,\phi)|,\ \ell=2,\ m=1$ $|Y_\ell^m(\theta,\phi)|,\ \ell=2,\ m=2$

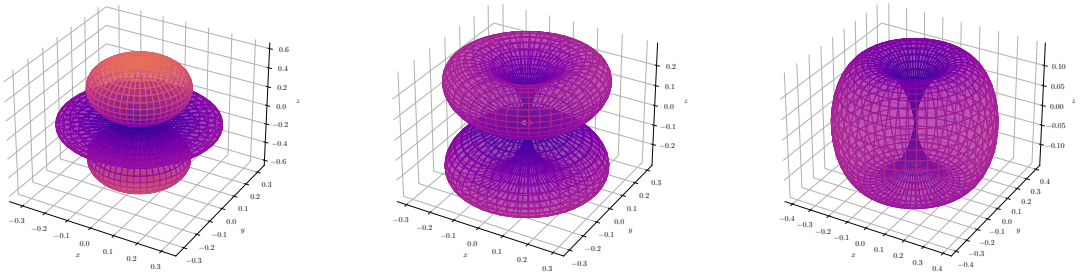


Figure 32: Surface plots made in `Python` visualizing spherical harmonics $|Y_m^\ell|$ for $\ell=0,1,2$.

The letter m is used as it is known as the "magnetic" quantum number in atomic physics. Subscripts m_ℓ can prevent confusion with electron mass m_e. The full spherical harmonics are

$$\langle\theta,\phi|\ell,m\rangle = Y_\ell^m(\theta,\phi) = N\,\mathrm{e}^{\mathrm{i}m\phi}\,P_\ell^m(\cos\theta), \tag{18.24}$$

where $P_\ell^m(x)$ are the associated Legendre polynomials. The normalization N is defined such that the spherical harmonics are normalized over a solid angle

$$\int_0^{2\pi}\int_0^{\pi} |Y_\ell^m|^2 \sin\theta \mathrm{d}\theta\mathrm{d}\phi = 1. \tag{18.25}$$

For completeness, the closed form expression for N is

$$N = (-1)^m \sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell-m)!}{(\ell+m)!}}. \tag{18.26}$$

Table 2 displays the first few spherical harmonics. Figure 32 uses the `sph_harm(m, l, phi, theta)` function of the Python `scipy.special` package to visualize surface plots of $|Y_\ell^m(\theta,\phi)|$. We use them to describe the angular dependence of hydrogen wavefunctions. Beyond atoms, $Y_\ell^m(\theta,\phi)$ describes normal modes of a balloon surface as the spherical counterpart to guitar strings. Cosmologists also use them for the cosmic microwave background.

19 Radial Schrödinger equation

We now solve the radial equation $\hat{H}R(r) = ER(r)$ by modelling the proton as a static Coulomb potential $V(r) = -e^2/4\pi\varepsilon_0 r$ in the Schrödinger equation (18.5). It is useful to recognize what physics each term represents in the resulting Hamiltonian that is only dependent on r

$$\hat{H} = \underbrace{-\frac{\hbar^2}{2m_e}\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}}{\mathrm{d}r}\right)}_{\text{radial kinetic}} + \underbrace{\frac{\ell(\ell+1)\hbar^2}{2m_e r^2}}_{\text{angular momentum}} \underbrace{-\frac{e^2}{4\pi\varepsilon_0 r}}_{\text{Coulomb}}. \tag{19.1}$$

The radial equation simplifies with a change of variables $u(r) = rR(r)$. This reduces the Hamiltonian to only one double derivative

$$-\frac{\hbar^2}{2m_e}\frac{\mathrm{d}^2u}{\mathrm{d}r^2} + \left[\frac{\ell(\ell+1)\hbar^2}{2m_e r^2} - \frac{e^2}{4\pi\varepsilon_0 r}\right]u = Eu. \tag{19.2}$$

Next introduce the wavenumber k and Bohr radius r_B as characteristic scales of the system

$$k^2 = -\frac{2m_e E}{\hbar^2}, \qquad r_B = \frac{4\pi\varepsilon_0\hbar^2}{e^2 m_e}, \tag{19.3}$$

where k is a real number because bound states have $E < 0$. The combination kr_B is dimensionless. This cleans up the notation of the radial equation

$$\frac{\mathrm{d}^2u}{\mathrm{d}r^2} - \frac{\ell(\ell+1)}{r^2}u + \frac{2}{r_B r}u = k^2u. \tag{19.4}$$

This form is helpful to examine the asymptotic behaviour of the radial wavefunctions.

19.1 Asymptotic radial behaviour

Far away limit. Let us develop intuition by taking the limit $r/r_B \to \infty$ where we are very far away from the atom to give

$$\frac{\mathrm{d}^2u}{\mathrm{d}r^2} = k^2u, \quad r/r_B \to \infty. \tag{19.5}$$

The solutions take the exponential form

$$u(r) = A\mathrm{e}^{-kr} + B\mathrm{e}^{+kr}. \tag{19.6}$$

We require the wavefunctions have finite normalization

$$\int |\Psi|^2\,\mathrm{d}^3\mathbf{r} \sim \int_0^\infty |R(r)|^2 r^2 \mathrm{d}r = \int_0^\infty |u(r)|^2\,\mathrm{d}r < \infty. \tag{19.7}$$

This normalizable criterion is a physics requirement that imposes $B = 0$, giving

$$u(r) \sim \mathrm{e}^{-kr}, \quad r \to \infty. \tag{19.8}$$

Near the origin. Now let's consider the opposite limit. As we approach the origin $r/r_B \to 0$, we have the other asymptotic limit where the angular momentum term dominates

$$\frac{\mathrm{d}^2 u}{\mathrm{d}r^2} - \frac{\ell(\ell+1)}{r^2} = 0, \quad r/r_B \to 0. \tag{19.9}$$

This has a general polynomial solution that can be checked by differentiating twice

$$u(r) = Cr^{\ell+1} + Dr^{-\ell}. \tag{19.10}$$

We require the radial wavefunction $R(r)$ be finite at the origin $r \to 0$. So the $u(r) = rR(r)$ functions must vanish and its first derivative must be finite

$$u(r \to 0) \to 0, \quad \left.\frac{\mathrm{d}u}{\mathrm{d}r}\right|_{r\to 0} < \infty. \tag{19.11}$$

The normalizability criterion means $r^{-\ell} \to \infty$ diverges for $r \to 0$, which imposes $D = 0$. Therefore we find the solutions near the origin take the form

$$u(r) \sim r^{\ell+1}, \quad r \to 0. \tag{19.12}$$

19.2 Full radial equation

We then physically motivate the full wavefunction to be a product of the asymptotic forms connected by a smooth function, $y(r)$, denoting the intermediate regime

$$u(r) = [\text{near origin}] \times [\text{intermediate}] \times [\text{far from origin}] = r^{\ell+1} \times y(r) \times \mathrm{e}^{-kr}. \tag{19.13}$$

To substitute this ansatz into equation (19.4), grab some paper and patiently differentiate

$$\frac{\mathrm{d}u}{\mathrm{d}r} = r^{\ell}\mathrm{e}^{-kr}\left[(\ell+1-kr)y + r\frac{\mathrm{d}y}{\mathrm{d}r}\right]. \tag{19.14}$$

Differentiating again and gathering the terms by derivative order gives

$$\frac{\mathrm{d}^2 u}{\mathrm{d}r^2} = r^{\ell}\mathrm{e}^{-kr}\left\{\left[\frac{\ell(\ell+1)}{r} - 2k(\ell+1) + k^2 r\right]y + 2(\ell+1-kr)\frac{\mathrm{d}y}{\mathrm{d}r} + r\frac{\mathrm{d}^2 y}{\mathrm{d}r^2}\right\}. \tag{19.15}$$

Substituting this into the Schrödinger equation (19.4), the $\ell(\ell+1)/r$ angular momentum and k^2 terms cancel. The radial Schrödinger equation reduces to (units of one over length)

$$r\frac{\mathrm{d}^2 y}{\mathrm{d}r^2} + 2(\ell+1-kr)\frac{\mathrm{d}y}{\mathrm{d}r} + 2k\left[\frac{1}{kr_B} - (\ell+1)\right]y = 0. \tag{19.16}$$

This takes the form of the *associated Laguerre differential equation*. We could now open a textbook to the associated Laguerre polynomials as solutions and move on with our day. But quantum textbooks propose a polynomial ansantz for $y(r)$ with unknown coefficients C_b

$$y(r) = \sum_{b=0}^{\infty} C_b r^b. \tag{19.17}$$

A physical asymptotic behaviour requires $C_0 \neq 0$. Differentiating twice yields

$$\frac{\mathrm{d}y}{\mathrm{d}r} = \sum_{b=0}^{\infty} bC_b r^{b-1} = \sum_{b=0}^{\infty} (b+1)C_{b+1} r^b, \tag{19.18}$$

$$\frac{\mathrm{d}^2y}{\mathrm{d}r^2} = \sum_{b=0}^{\infty} b(b-1)C_b r^{b-2} = \sum_{b=0}^{\infty} (b+1)bC_{b+1} r^{b-1}. \tag{19.19}$$

The second equalities shift the summation index $b \to b+1$, which does not change the sum (shifting the sum to start at $b=-1$, the $b+1$ makes it vanish). These expressions let us choose the one that has a uniform power of r^b when substituting into equation (19.16):

$$\sum_b (b+1)bC_{b+1}r^b + 2(\ell+1)\sum_b (b+1)C_{b+1}r^b - 2k\sum_b bC_b r^b + 2k\left[\frac{1}{kr_B} - (\ell+1)\right]\sum_b C_b r^b = 0. \tag{19.20}$$

Equating the coefficients of r^b gives us a recurrence relation between all the C_b coefficients

$$C_{b+1} = \frac{2k(b+\ell+1-1/kr_B)}{(b+1)[b+2(\ell+1)]}C_b. \tag{19.21}$$

This fully determines all the coefficients C_b and therefore the Laguerre polynomials (19.17).

19.3 Energy quantization

Consider the limit of large values of $b \to \infty$, where the recurrence relation (19.21) becomes

$$C_{b+1} \to \frac{2kb}{b(b+1)}C_b = \frac{2k}{b+1}C_b, \quad b \to \infty. \tag{19.22}$$

This allows us to relate a general coefficient C_b to the zeroth C_0 by

$$C_b = \frac{2^b k}{b!}C_0, \quad b \to \infty. \tag{19.23}$$

If this were the exact result, then the radial wavefunction becomes

$$u(r) = C_0 r^{\ell+1}\mathrm{e}^{-kr}\underbrace{\left(\sum_{b=0}^{\infty}\frac{2^b k}{b!}r^b\right)}_{\mathrm{e}^{2kr}} = C_0 r^{\ell+1}\mathrm{e}^{kr}. \tag{19.24}$$

This is divergent for $r \to \infty$ so is non-normalizable and unphysical. The only way to impose wavefunction normalization is to require the series be finite and must terminate. This means for some integer N, the coefficient and all its successors vanish

$$C_N = 0, \quad C_{N-1} \neq 0. \tag{19.25}$$

Plugging this terminal requirement into the recurrence relation equation (19.21) implies

$$N + \ell - \frac{1}{kr_B} = 0 \tag{19.26}$$

Given both ℓ and N are integers, we conventionally define the integer $n = N + \ell$ called the **principal quantum number**

$$n = \frac{1}{kr_B}. \tag{19.27}$$

Now invoking the definitions of wavenumber k and Bohr radius r_B from equation (19.3), we find the energies E_n must be integer quantized

$$E_n = -\frac{m_e}{2}\left(\frac{e^2}{4\pi\varepsilon_0\hbar}\right)^2 \frac{1}{n^2} = -\frac{E_R}{n^2} \simeq -\frac{13.6\text{ eV}}{n^2}, \quad n = 1, 2, 3, \ldots. \tag{19.28}$$

This is the Bohr energy formula for the hydrogen that we met earlier (2.13), but now correctly derived from Schrödinger's equation rather than Newton's laws. Energy quantization arises from requiring that the wavefunction not diverge for $r \to \infty$.

19.4 Fine-structure constant

Realizing $e^2/4\pi\varepsilon_0\hbar$ has dimensions of velocity, we use the speed of light c to define a dimensionless number called the electromagnetic **fine-structure constant**

$$\alpha = \frac{e^2}{4\pi\varepsilon_0\hbar c} \simeq \frac{1}{137} \approx 0.0073. \tag{19.29}$$

This number sets the fundamental strength of electromagnetism and is the same in any system of units. It is determined experimentally and currently has a relative uncertainty of 10^{-10}. The Bohr energy formula (19.28) becomes much more illuminating and easier to remember

$$E_n = -\left(\frac{1}{2}\alpha^2 m_e c^2\right)\frac{1}{n^2}. \tag{19.30}$$

It is now profoundly clear what physics sets the scale of hydrogen: the electron rest mass m_e

$$m_e c^2 \simeq 5.11 \times 10^5\text{ eV} = 511\text{ keV}. \tag{19.31}$$

Hydrogen energies are much smaller than the electron rest mass due to the two powers of α

$$\alpha^2 \simeq 5.3 \times 10^{-5}. \tag{19.32}$$

We can also rewrite the Bohr radius (19.3) more insightfully as

$$r_B = \frac{\hbar c}{m_e c^2 \alpha} = \frac{\lambda\!\!\!^{-}_C}{\alpha} \simeq 0.529 \text{ Å}, \tag{19.33}$$

where $\lambda\!\!\!^{-}_C = \lambda_C/2\pi \simeq 3.9 \times 10^{-13}$ m is the *reduced Compton wavelength* of the electron, which we met in equation (1.20). Again, this is not only easier to remember but illuminates what physics sets the size of hydrogen. It is the wavelength $\lambda\!\!\!^{-}_C$ associated with the electron rest-mass energy and is enlarged by the fundamental strength of electromagnetism α.

This analysis importantly shows the fundamental physics of hydrogen are really just two quantities: an invariant rest mass m_e and a dimensionless number α

$$\alpha \simeq \frac{1}{137}, \quad m_e \simeq 511 \text{ keV}/c^2. \tag{19.34}$$

The fundamental constants of nature $\hbar c \simeq 197$ eV nm then convert energy into a length scale $\lambda\!\!\!^{-}_C = \hbar c/(m_e c^2 \alpha)$. There is no need to remember all the factors of $e^2/4\pi\varepsilon_0$, m_e and powers of $\hbar$. We shall return to discussing hydrogen with greater detail in section 24.

HW9: Spherical Schrödinger problem (set Nov 5, due Nov 12)

1. The total angular momentum operator in spherical coordinates (r,θ,ϕ) is

$$\hat{L}^2 = -\hbar^2\left[\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]. \tag{19.35}$$

   Why do $\hat{L}^2$, $\hat{L}_z$, and the Hamiltonian $\hat{H}$ commute for central potentials $V = V(r)$? Explain the physical significance.

2. Write down the eigenvalue equations for $\hat{L}^2$ and $\hat{L}_z$. Show that the following spherical harmonics are (i) eigenstates of $\hat{L}^2$ and $\hat{L}_z$, (ii) orthogonal by integration:

$$Y_1^0 = \sqrt{\tfrac{3}{4\pi}}\cos\theta, \quad Y_1^1 = -\sqrt{\tfrac{3}{8\pi}}e^{+i\phi}\sin\theta, \quad Y_1^{-1} = \sqrt{\tfrac{3}{8\pi}}e^{-i\phi}\sin\theta. \tag{19.36}$$

3. The time-independent Schrödinger equation in spherical coordinates is

$$\left[-\frac{\hbar^2}{2m_e}\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}}{\mathrm{d}r}\right) + \frac{\hat{L}^2}{2m_e r^2} + V(r)\right]\psi(r,\theta,\phi) = E\psi(r,\theta,\phi). \tag{19.37}$$

   (a) Comment on the physical significance of each term in this equation. Why can we factorize the wavefunction into a product of radial and angular functions $\psi(r,\theta,\phi) = R(r)Y(\theta,\phi)$? Why can we replace the $\hat{L}^2$ operator with $\ell(\ell+1)\hbar^2$?

   (b) Assuming the Coulomb potential $V(r) = -e^2/4\pi\varepsilon r$, explain the concept of an effective potential $V_{\text{eff}}(r)$ and what is the expression for this? What is the physical significance of the relative sign in front of the $\hat{L}^2$ and Coulomb terms?

   (c) Sketch the Coulomb potential, $\ell(\ell+1)\hbar^2/2m_e r^2$ piece, and effective $V_{\text{eff}}(r)$ for cases when the electron has (i) no angular momentum $\ell = 0$ and (ii) one unit of angular momentum $\ell = 1$. In the $\ell = 1$ case, explain the physical behaviour near and far from $r = 0$.

   (d) Change variables to $u(r) = rR(r)$ and show the radial equation becomes

$$\frac{\mathrm{d}^2u}{\mathrm{d}r^2} - \frac{\ell(\ell+1)}{r^2}u + \frac{2}{r_B r}u = k^2u. \tag{19.38}$$

   What are the expressions for k and r_B? Explain their physical interpretation.

   (e) Show the derivations for the functional form of the radial wavefunction $R(r)$ in the asymptotic limits: (i) near $r/r_B \to 0$ and far (ii) $r/r_B \to \infty$ from the nucleus. You do not need to normalize the wavefunction.

20 Discovery of spin half

20.1 Stern-Gerlach experiment

In 1922, Otto Stern and Walter Gerlach passed silver-ion beams through an inhomogeneous magnetic field. They observed the beam split into two discrete components in space (figure 33 left). The results of this Stern-Gerlach experiment surprised expectations in two ways:

- classical physics predicts a continuous spatial distribution,
- if angular momentum were integer quantized, there should be three not two parts.

Classical electromagnetism describes a loop of current $I = Qf$ as a charge Q circulating with angular frequency $\omega = 2\pi f$ about an area A with normal vector $\mathbf{n}$. This has a magnetic dipole

$$\boldsymbol{\mu} = IA\mathbf{n} = Q\left(\frac{\omega}{2\pi}\right)(\pi r^2)\mathbf{n} = -\frac{e}{2m_e}\mathbf{L}, \tag{20.1}$$

where the angular momentum is $\mathbf{L} = m_e\mathbf{r}\times\mathbf{v} = (m_e r^2\omega)\mathbf{n}$. The potential energy of $\boldsymbol{\mu}$ in an external magnetic field $\mathbf{B}$ is the scalar product $U = -\boldsymbol{\mu}\cdot\mathbf{B}$. If we align the field along the z axis $\mathbf{B} = (0,0,B_z)$ and it is inhomogeneous $\partial B_z/\partial z \neq 0$, the force is given by the gradient

$$f_z = -\boldsymbol{\nabla}_z U = -\frac{\partial}{\partial z}(\boldsymbol{\mu}\cdot\mathbf{B}) = \frac{e}{2m_e}L_z\frac{\partial B_z}{\partial z}. \tag{20.2}$$

In the classical picture, L_z and therefore f_z can take any continuous value, which disagrees with the Stern-Gerlach experiment.

If we naïvely quantized $L_z \to \hat{L}_z$ following orbital angular momentum formalism, the magnetic quantum number is $\hat{L}_z|\psi\rangle = m_\ell\hbar|\psi\rangle$. Historically in 1922, the prevailing Bohr model hypothesized integer-quantized orbital angular momentum $L \stackrel{?}{=} n\hbar$ to describe atoms. The expectation value of the force $\langle m_\ell|\hat{f}_z|m_\ell\rangle$ acting on the silver ions in state $|m_\ell\rangle$ becomes

$$\langle\hat{f}_z\rangle = \frac{e}{2m_e}(m_\ell\hbar)\frac{\partial B_z}{\partial z} = m_\ell\mu_B\frac{\partial B_z}{\partial z}, \tag{20.3}$$

where $\mu_B = e\hbar/2m_e$ is the Bohr magneton. Orbital angular momentum is integer quantized

$$\ell = 1 \quad\Rightarrow\quad \langle f_z\rangle = m_\ell\mu_B\frac{\partial B_z}{\partial z}, \quad \text{where} \quad m_\ell = -1,0,+1. \tag{20.4}$$

So the force splits the beam into three discrete directions. This disagrees with the two observed in the Stern-Gerlach experiment. This implies there exists half-integer quantized angular momentum "$\ell = 1/2$" giving two values $m_\ell = \pm 1/2$.

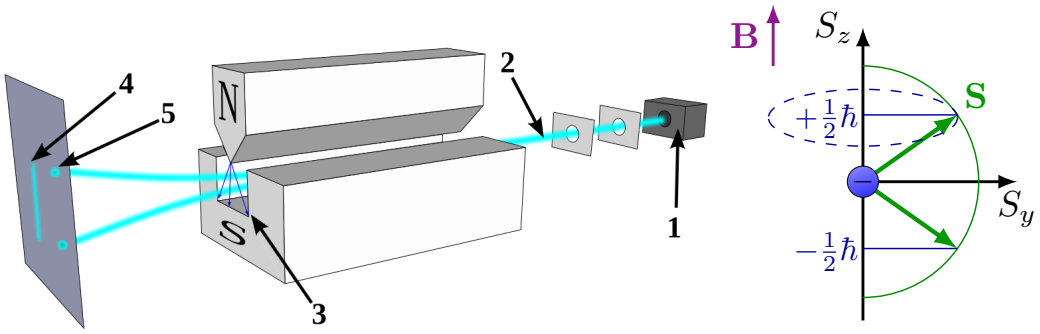


Figure 33: Left: Stern-Gerlach experiment showing charged particles (1,2) entering an inhomogeneous magnetic field (3), giving the continuous classical expectation (4) versus observed quantized result (5). Right: quantized spin for a spin-half particle, where the $\hat{S}_z$ axis is aligned to the external magnetic field $\mathbf{B}$. Images: Tatoute CC-BY-4.0, Izaak Neutelings.

Samuel Goudsmit and George Uhlenbeck realized in 1925 that this effect arose from intrinsic angular momentum called spin $\hat{\mathbf{S}}$ that is *distinct* from the orbital counterpart $\hat{\mathbf{L}}$. The electron carries an intrinsic magnetic moment μ_s given by

$$\hat{\boldsymbol{\mu}}_s = -\frac{g_s e}{2m_e}\hat{\mathbf{S}}. \tag{20.5}$$

where g_s is the electron spin **g-factor**. Historically, this dimensionless parameter was found empirically to be $g_s = 2$. The spin-magnetic Hamiltonian with quantum operators is

$$\hat{H} = -\hat{\boldsymbol{\mu}}_s \cdot \mathbf{B} = \frac{g_s e}{2m_e}\hat{\mathbf{S}} \cdot \mathbf{B}. \tag{20.6}$$

The $\hat{S}_z|s,m_s\rangle = m_s\hbar|s,m_s\rangle$ is analogous to the orbital angular momentum eigenvalue equation. This correctly describes the Stern-Gerlach experiment, where the resulting force (20.3) deflects the ions in one of two directions (figure 33) depending on the spin state

$$f_z^+ = +\mu_B \frac{\partial B_z}{\partial z}, \quad m_s = +\frac{1}{2} \tag{20.7}$$

$$f_z^- = -\mu_B \frac{\partial B_z}{\partial z}, \quad m_s = -\frac{1}{2}. \tag{20.8}$$

The Stern-Gerlach experiment provided direct evidence of quantized spin. But indirect hints of spin were already seen in cadmium line spectra known as the *anomalous Zeeman effect*, discovered by Thomas Preston in 1897 and we revisit in section 25.2. The original setup used silver atoms and it is fortuitous in hindsight that these atoms are indeed spin-half.

The relativistic counterpart of Schrödinger's equation is called the Dirac equation, which fixes the value of $g_s = 2$ theoretically. The actual value deviates from two, which is called the

anomalous magnetic moment. This discovery occurred a short subway ride away (take the 1 line to 116th St) at Columbia University in 1947, where Polykarp Kusch and Henry Foley precisely measured gallium and sodium spectra to determine [37, 38]

$$g_s = 2(1.00119 \pm 0.00005). \tag{20.9}$$

Calculating the deviation requires graduate quantum field theory, which Julian Schwinger derived in 1947 to be $\alpha/2\pi \simeq 0.00116$ [39] and stunningly agrees with measurement.

20.2 Two-state qubit system

Electrons are spin-half particles and can have one of two spin states. We call the $m_s = +1/2$ state spin-up $|\uparrow\rangle$ and $m_s = -1/2$ state spin-down $|\downarrow\rangle$, which form a two-state basis:

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{20.10}$$

This forms a complete orthonormal basis of a two-dimensional Hilbert space. The $|\uparrow\rangle$ and $|\downarrow\rangle$ span the two basis states of a quantum system called a qubit. Some textbooks write spin-up and spin-down using plus-minus symbols $(|+\rangle, |-\rangle)$ to match the eigenvalue sign or $(|0\rangle, |1\rangle)$ notation to evoke quantum bits of information

$$|\uparrow\rangle = |0\rangle = |+\rangle, \tag{20.11}$$

$$|\downarrow\rangle = |1\rangle = |-\rangle. \tag{20.12}$$

It's a matter of taste and textbooks use them interchangeably. A general spin state is called a **spinor** $|\chi\rangle$, which we can write as quantum superposition of spin-up and spin-down

$$|\chi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \tag{20.13}$$

where α and β are complex numbers that satisfy the normalization condition $|\alpha|^2 + |\beta|^2 = 1$. The spin operator that acts on these states is

$$\hat{\mathbf{S}} = \frac{\hbar}{2}\hat{\boldsymbol{\sigma}} = \frac{\hbar}{2}\begin{pmatrix} \hat{\sigma}_x \\ \hat{\sigma}_y \\ \hat{\sigma}_z \end{pmatrix}, \tag{20.14}$$

where $\hat{\boldsymbol{\sigma}}$ is a vector of three **Pauli matrices** that look like

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \hat{\sigma}_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \qquad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{20.15}$$

In this basis, the $\hat{S}_z$ operator acts on the states $|\uparrow\rangle, |\downarrow\rangle$ to give the eigenvalue equations, which we can write in explicit matrix representation as

$$\hat{S}_z|\uparrow\rangle = +\frac{\hbar}{2}|\uparrow\rangle \quad \longleftrightarrow \quad \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = +\frac{\hbar}{2}\begin{pmatrix}1\\0\end{pmatrix} \tag{20.16}$$

The corresponding spin-down eigenvalue equation is

$$\hat{S}_z|\downarrow\rangle = -\frac{\hbar}{2}|\downarrow\rangle \quad \longleftrightarrow \quad \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = -\frac{\hbar}{2}\begin{pmatrix}0\\1\end{pmatrix}. \tag{20.17}$$

The measured eigenvalue $m_s = \pm\hbar/2$ has units of angular momentum. Like m_ℓ for orbital angular momentum, use subscripts m_s assiduously to disambiguate with mass m_e.

20.3 Spin in spherical coordinates

Let us generalize to any direction via a three-dimensional unit vector in spherical coordinates

$$\mathbf{n} = \begin{pmatrix}\sin\theta\cos\phi\\ \sin\theta\sin\phi\\ \cos\theta\end{pmatrix} \tag{20.18}$$

The Pauli operator projected onto direction $\mathbf{n}$ in spherical coordinates $\hat{\sigma}_{\theta,\phi}$ is

$$\hat{\sigma}_{\theta,\phi} = \hat{\boldsymbol{\sigma}}\cdot\mathbf{n} = \hat{\sigma}_x\sin\theta\cos\phi + \hat{\sigma}_y\sin\theta\sin\phi + \hat{\sigma}_z\cos\theta = \begin{pmatrix}\cos\theta & \mathrm{e}^{-\mathrm{i}\phi}\sin\theta\\ \mathrm{e}^{\mathrm{i}\phi}\sin\theta & -\cos\theta.\end{pmatrix} \tag{20.19}$$

We can then write the eigenvalue equation

$$\hat{\boldsymbol{\sigma}}\cdot\mathbf{n}\begin{pmatrix}a\\b\end{pmatrix} = \lambda\begin{pmatrix}a\\b\end{pmatrix}, \tag{20.20}$$

from which we solve for the eigenvalues λ via the characteristic equation

$$\begin{vmatrix}\cos\theta-\lambda & \mathrm{e}^{-\mathrm{i}\phi}\sin\theta\\ \mathrm{e}^{\mathrm{i}\phi}\sin\theta & -(\lambda+\cos\theta)\end{vmatrix} = \lambda^2 - \sin^2\theta - \cos^2\theta = 0. \tag{20.21}$$

The eigenvalues are $\lambda_+ = +1$ for spin-up and $\lambda_- = -1$ for spin-down projected along $\mathbf{n}$. The eigenstates must satisfy the normalization condition $|a|^2+|b|^2 = 1$. For the spin-up state $\lambda_+ = +1$, we find the eigenvector $|\lambda_+,\mathbf{n}\rangle$ via usual eigenvalue methods

$$\frac{b}{a} = \frac{1-\cos\theta}{\mathrm{e}^{-\mathrm{i}\phi}\sin\theta} \quad \Rightarrow \quad |\lambda_+,\mathbf{n}\rangle = \begin{pmatrix}\cos\frac{\theta}{2}\\ \mathrm{e}^{\mathrm{i}\phi}\sin\frac{\theta}{2}\end{pmatrix}, \tag{20.22}$$

where we used half-angle identities $\sin 2\theta = 2\sin\theta\cos\theta$. For simplicity, let us set $\phi = 0$. If we start with a state $|\uparrow, \theta = 0\rangle$ and rotate by 180°, we obtain $|\downarrow\rangle$ as intuition expects:

$$|\uparrow, \theta = 0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |\uparrow, \theta = \pi\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |\downarrow\rangle. \tag{20.23}$$

However, if we rotate by 360°, we do not get back the same state $|\uparrow\rangle$ but instead has a minus sign! We must rotate by 720° to return to the same state

$$|\uparrow, \theta = 2\pi\rangle = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -|\uparrow, \theta = 0\rangle, \quad |\uparrow, \theta = 4\pi\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{20.24}$$

A spinor does not behave like a vector. Rotating a momentum vector $\mathbf{p}$ until it is unchanged requires a 360° rotation. By contrast, rotating a spinor $|\uparrow\rangle$ until it is unchanged requires a 720° rotation. This also happens when one transports an arrow around a Möbius strip.

Now consider an initial Stern-Gerlach filter SG_A aligned in the $\mathbf{z}$ direction, which transmits spin-$\frac{1}{2}$ particles in the $|\uparrow\rangle$ state and block those in the $|\downarrow\rangle$. These filtered particles enter a second Stern-Gerlach filter SG_B with magnetic field oriented in a different direction $\mathbf{n}$, which transmits particles in the $|\lambda_+, \mathbf{n}\rangle$ state:

$$\xrightarrow[\text{spin-half particles}]{} \boxed{\text{SG}_A} \xrightarrow[\text{transmits } |\uparrow\rangle \text{ along } \mathbf{z}]{} \boxed{\text{SG}_B} \xrightarrow[\text{transmits } |\lambda_+\rangle \text{ along } \mathbf{n}]{} \tag{20.25}$$

We ask: given particles enter SG_B in the spin-up $|\uparrow\rangle$ along $\mathbf{z}$ state, what is the probability $P(\lambda_+, \mathbf{n}|\uparrow)$ that they are in spin-up along the $\mathbf{n}$ direction? This probability is given by

$$P(\lambda_+, \mathbf{n}|\uparrow) = |\langle\lambda_+ \mathbf{n}|\uparrow\rangle|^2 = \left| \left(\cos\tfrac{\theta}{2} \; e^{i\phi} \sin\tfrac{\theta}{2} \right) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right|^2 = \cos^2 \frac{\theta}{2}. \tag{20.26}$$

21 Dynamics of spin

21.1 Pauli matrix algebra

The spin operators obey the commutation relation

$$[\hat{S}_x, \hat{S}_y] = \mathrm{i}\hbar\hat{S}_z, \quad [\hat{S}_y, \hat{S}_z] = \mathrm{i}\hbar\hat{S}_x, \quad [\hat{S}_z, \hat{S}_x] = \mathrm{i}\hbar\hat{S}_y. \tag{21.1}$$

We write the spin (Pauli matrix) operator $\hat{S}_i = (\hbar/2)\hat{\sigma}_i$ commutators compactly as

$$[\hat{S}_i, \hat{S}_j] = \mathrm{i}\hbar\varepsilon_{ijk}\hat{S}_k \quad \longleftrightarrow \quad [\hat{\sigma}_i, \hat{\sigma}_j] = 2\mathrm{i}\varepsilon_{ijk}\hat{\sigma}_k. \tag{21.2}$$

Explicit matrix multiplication verifies these commutators, where we do one in full detail:

$$\begin{aligned}
[\hat{S}_x, \hat{S}_y] &= \hat{S}_x\hat{S}_y - \hat{S}_y\hat{S}_x && (21.3)\\
&= \frac{\hbar^2}{4}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} - \frac{\hbar^2}{4}\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} && (21.4)\\
&= \frac{\hbar^2}{4}\begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix} - \frac{\hbar^2}{4}\begin{pmatrix} -\mathrm{i} & 0 \\ 0 & \mathrm{i} \end{pmatrix} = \mathrm{i}\frac{\hbar^2}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \mathrm{i}\hbar\hat{S}_z. && (21.5)
\end{aligned}$$

The Pauli matrices square to the identity $\hat{\sigma}_i^2 = \hat{I}$, so the total spin squared operator $\hat{\mathbf{S}}^2$ is

$$\hat{S}^2 = \hat{\mathbf{S}}\cdot\hat{\mathbf{S}} = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2 = \frac{3}{4}\hbar^2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{21.6}$$

As $\hat{S}^2$ is proportional to the identity, this operator commutes with each spin component

$$[\hat{S}^2, \hat{S}_x] = 0, \quad [\hat{S}^2, \hat{S}_y] = 0, \quad [\hat{S}^2, \hat{S}_z] = 0. \tag{21.7}$$

This commutation relation shows $\hat{S}_z$ and $\hat{S}^2$ are simultaneously observable. You may sense déjà vu: these commutators (21.1) are the same as those of orbital angular momentum operators $\hat{L}_i$ in (16.18). Formally, commutators define an *algebra* and the differential operators or matrices form a *representation* of the algebra.

For a particle with general spin number s and m_s, the total and azimuthal spin operators have the following general eigenvalue equations

$$\hat{S}^2|s, m_s\rangle = \hbar^2 s(s+1)|s, m_s\rangle, \tag{21.8}$$

$$\hat{S}_z|s, m_s\rangle = \hbar m_s|s, m_s\rangle. \tag{21.9}$$

Generally, s and m_s can be half-integer or integer numbers. Moreover, the azimuthal spin quanutm numbers are bounded by the total spin quantum number $|m_s| \leq s$:

$$m_s = -s, -s+1, \ldots, s-1, s. \tag{21.10}$$

In this notation, we can write the spin-$\frac{1}{2}$ particle eigenstates as

$$|\uparrow\rangle = \left|s = \tfrac{1}{2}, m_s = +\tfrac{1}{2}\right\rangle, \quad |\downarrow\rangle = \left|s = \tfrac{1}{2}, m_s = -\tfrac{1}{2}\right\rangle. \tag{21.11}$$

For the $\hat{S}_x$ component, we proceed by solving the eigenvalue equation $\hat{S}_x\chi = \lambda_x\chi$:

$$\frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} = \lambda_x \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix}. \tag{21.12}$$

Applying the characteristic equation of a 2×2 eigenvalue problem gives

$$\lambda_x^2 - \mathrm{tr}(\hat{S}_x)\lambda_x + \det(\hat{S}_x) = 0 \tag{21.13}$$

The Pauli matrices are traceless $\mathrm{tr}(\hat{S}_x) = 0$ and the determinant is $\det(\hat{S}_x) = -(\hbar/2)^2$, so the eigenvalues are $\lambda_x = \pm\hbar/2$. The corresponding normalized eigenvectors are

$$\chi_+ = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ +1 \end{pmatrix} \quad \text{for} \quad \lambda_x = +\frac{\hbar}{2}, \tag{21.14}$$

$$\chi_- = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{for} \quad \lambda_x = -\frac{\hbar}{2}. \tag{21.15}$$

Like the orbital angular momentum ladder operators $\hat{L}_\pm$, there are spin counterparts

$$\hat{S}_\pm = \hat{S}_x \pm \mathrm{i}\hat{S}_y. \tag{21.16}$$

In matrix form, we can write them as

$$\hat{S}_+ = \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \hat{S}_- = \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{21.17}$$

Now we can see what happens when we act these on the spin states

$$\hat{S}_+|\uparrow\rangle = \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{21.18}$$

$$\hat{S}_+|\downarrow\rangle = \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \hbar\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \hbar|\uparrow\rangle. \tag{21.19}$$

We see that the raising spin operator indeed changes the spin-down into a spin-up state. It also makes the spin-up state vanish, as we would expect, because there are no more states above that. Similarly, the lowering operator has similar results

$$\hat{S}_-|\uparrow\rangle = \hbar|\downarrow\rangle, \quad \hat{S}_-|\downarrow\rangle = 0. \tag{21.20}$$

21.2 Spin magnetic precession

Let us now revisit magnet moments but now in a homogeneous magnetic field aligned along the z-coordinate $\mathbf{B} = B\mathbf{z}$. Recall the spin-magnetic Hamiltonian $\hat{H} = -\hat{\boldsymbol{\mu}} \cdot \mathbf{B}$ is

$$\hat{H} = -\gamma B \hat{S}_z, \quad \text{where} \quad \gamma = -\frac{g_s e}{2m}. \tag{21.21}$$

The parameter γ is called the gyromagnetic ratio. The energy eigenvalues E_i corresponding to the spin-up $|\uparrow\rangle$ and spin-down $|\downarrow\rangle$ eigenstates are

$$E_\uparrow = -\frac{\gamma B\hbar}{2}, \quad |\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{21.22}$$

$$E_\downarrow = +\frac{\gamma B\hbar}{2}, \quad |\downarrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{21.23}$$

The time-evolved states have a multiplicative complex phase and have no impact on the observed probabilities

$$|\uparrow, t\rangle = \exp\left(+\frac{\mathrm{i}\gamma B t}{2}\right)|\uparrow, t = 0\rangle, \tag{21.24}$$

$$|\downarrow, t\rangle = \exp\left(-\frac{\mathrm{i}\gamma B t}{2}\right)|\downarrow, t = 0\rangle. \tag{21.25}$$

However, we could consider the spin pointing along the x axis instead

$$|\rightarrow\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}}\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right]. \tag{21.26}$$

Then the time-evolved state is a superposition of the $|\uparrow\rangle, |\downarrow\rangle$

$$|\rightarrow, t\rangle = \frac{1}{\sqrt{2}}\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mathrm{e}^{+\mathrm{i}\gamma B t/2} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mathrm{e}^{-\mathrm{i}\gamma B t/2}\right] = \frac{1}{\sqrt{2}}\begin{pmatrix} \mathrm{e}^{+\mathrm{i}\gamma B t/2} \\ \mathrm{e}^{-\mathrm{i}\gamma B t/2} \end{pmatrix}. \tag{21.27}$$

We write this more insightfully by factorizing out a global phase that has no observable effect

$$|\rightarrow, t\rangle = \mathrm{e}^{\mathrm{i}\gamma B t/2}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \mathrm{e}^{-\mathrm{i}\gamma B t} \end{pmatrix}. \tag{21.28}$$

We can now see that the lower component of the spinor changes to $-\mathrm{i}$ for $t = \pi/\gamma B$ and $\mathrm{e}^{\mathrm{i}\pi} = -1$ for $t = 2\pi/\gamma B$

$$|\rightarrow, t = 2\pi/\gamma B\rangle = -\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = -|\rightarrow, t = 0\rangle. \tag{21.29}$$

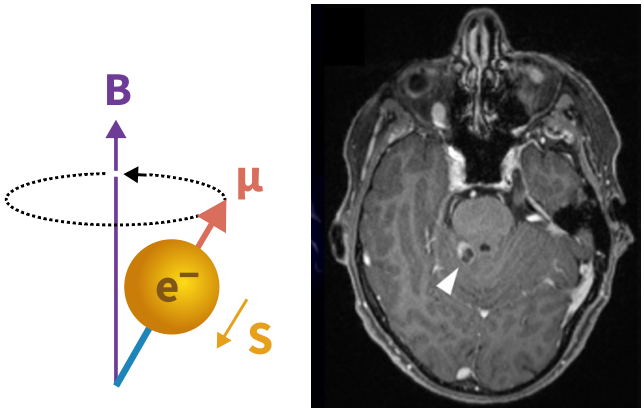


Figure 34: Left: Sketch of electron spin precession in a magnetic field. Right: application of nuclear magnetic precession to magnetic resonance imaging in medical diagnosis [40].

This analysis also follows from Ehrenfest's theorem (13.23), which tells us how $\langle \hat{\mathbf{S}} \rangle$ evolves with time. Because the Hamiltonian is time-independent and only contains the $\hat{S}_z$ operator that commutes $[\hat{S}_z, \hat{H}] = 0$, we immediately find $\langle \hat{S}_z \rangle$ is a constant of motion:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \hat{S}_z \rangle = 0 \quad \Rightarrow \quad \langle \hat{S}_z \rangle = \text{constant}. \tag{21.30}$$

However, $\hat{S}_x$ and $\hat{S}_y$ do not commute with $\hat{H}$, giving

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\langle \hat{S}_x \rangle &= \gamma B \langle \hat{S}_y \rangle \\ \frac{\mathrm{d}}{\mathrm{d}t}\langle \hat{S}_y \rangle &= -\gamma B \langle \hat{S}_x \rangle \end{aligned} \tag{21.31}$$

We can write these equations in a form that corresponds to classical precession of $\langle \hat{\mathbf{S}} \rangle$

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \hat{\mathbf{S}} \rangle = \gamma \langle \hat{\mathbf{S}} \rangle \times \mathbf{B}. \tag{21.32}$$

This equation governs **magnetic precession** about the **B**-field axis with angular frequency

$$\omega = \frac{g_s e B}{2m}. \tag{21.33}$$

This is the **Larmor frequency**, named after Joseph Larmor. Precession frequency provides a helpful way to remember the value of the electron Bohr magneton is $\mu_B = h \times 14.0\ \mathrm{GHz\ T^{-1}}$. The time-evolved state is the original state rotated about the z axis by amount ωt. The general time evolution of a spin state $|\psi(t)\rangle$ in the spin-up and spin-down basis is given by

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}} \mathrm{e}^{-\mathrm{i}\omega t/2} |\uparrow\rangle + \frac{1}{\sqrt{2}} \mathrm{e}^{\mathrm{i}\omega t/2} |\downarrow\rangle. \tag{21.34}$$

From these, we can obtain explicit forms of $\langle \hat{S}_x \rangle$ and $\langle \hat{S}_y \rangle$

$$\langle \hat{S}_x \rangle = \langle \uparrow, x | \hat{S}_x | \psi(t) \rangle = \frac{\hbar}{2} \cos \omega t, \tag{21.35}$$

$$\langle \hat{S}_y \rangle = \langle \uparrow, y | \hat{S}_y | \psi(t) \rangle = \frac{\hbar}{2} \sin \omega t. \tag{21.36}$$

Spin precession finds amazing applications with local history. Atomic nuclei precess in strong magnetic fields and can absorb radio waves for magnetic resonance imaging (MRI), which has proven transformative for medical imaging (figure 34). Paul Lauterbur pioneered MRI in 1973 at Stony Brook University [41], an hour train ride away on the Long Island Rail Road. Nuclei have a magnetic moment given by

$$\hat{\boldsymbol{\mu}}_N = g_N \left(\frac{e\hbar}{2m_p} \right) \hat{\mathbf{I}} = g_N \mu_N \hat{\mathbf{I}}, \tag{21.37}$$

defined with the nuclear magneton $\mu_N = e\hbar/2m_p$, proton mass m_p and nuclear spin $\hat{\mathbf{I}}$. Isidor Rabi, Jerome Kellogg and Jerrold Zacharias measured the proton magnetic moment to be very large $\mu_p = (3.2 \pm 0.3)\mu_N$ at Columbia, published in 1934 [42]. For several decades, the anomalously large proton and neutron g-factors were a longstanding mystery

$$g_{\text{proton}} \simeq +5.6, \quad g_{\text{neutron}} \simeq -3.8. \tag{21.38}$$

Naïvely, neutral point-like particles should have zero magnetic moment $g = 0$, contradicting the observed neutron g-factor. Moreover, while Leona Marshall Woods was NYU faculty in 1962, she measured a non-zero magnetic moment for an "exotic neutron" (called a Λ^0 baryon) using the Brookhaven Cosmotron on Long Island [43]. Lindsay Schachinger of Rutgers University later precisely measured this to be $\mu_{\Lambda^0} = (-0.6138 \pm 0.0047)\mu_N$ for her thesis in 1978 [44]. With hindsight, these surprises indicated these particles were composite long before the direct discovery of quarks after 1968 and helped decipher the strong force.

22 Identical particles

So far we have only studied one particle at a time, or at most two distinguishable particles of hydrogen comprising an electron and proton. We now study a fundamental concept in quantum mechanics: two particles with the same quantum numbers are indistinguishable.

22.1 Bosons and fermions

Let us denote a state with two identical particles as

$$|a,b\rangle = |a\rangle_1 |b\rangle_2, \tag{22.1}$$

which represents particle 1 in state a and particle 2 in state b. Let us define the exchange operator $\hat{P}_{12}$ that exchanges labels 1 and 2

$$\hat{P}_{12}|a,b\rangle = |b,a\rangle. \tag{22.2}$$

Because quantum particles are **identical**, we cannot tell if we have exchanged the particles' labels for any physical state. Their probabilities must be equal:

$$|\langle a,b|b,a\rangle|^2 = |\langle b,a|a,b\rangle|^2 \tag{22.3}$$

This is the **exchange principle** of identical particles. In the position representation, we can equally write this as an exchange of location $x_a \leftrightarrow x_b$

$$|\psi(x_b,x_a)|^2 = |\psi(x_a,x_b)|^2 \tag{22.4}$$

For the general case where state $a \neq b$, we can represent the $\hat{P}_{12}$ operator as a matrix in the $(|a,b\rangle, |b,a\rangle)$ basis

$$\hat{P}_{12} = \begin{pmatrix} \langle a,b|\hat{P}_{12}|a,b\rangle & \langle a,b|\hat{P}_{12}|b,a\rangle \\ \langle b,a|\hat{P}_{12}|a,b\rangle & \langle b,a|\hat{P}_{12}|b,a\rangle \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{22.5}$$

Its eigenstates are the general symmetric $|\psi_S\rangle$ and antisymmetric $|\psi_A\rangle$ two-particle states

$$|\psi_S\rangle = \frac{1}{\sqrt{2}}(|a,b\rangle + |b,a\rangle) \quad \text{bosons, integer spin,} \tag{22.6}$$

$$|\psi_A\rangle = \frac{1}{\sqrt{2}}(|a,b\rangle - |b,a\rangle) \quad \text{fermions, half-integer spin.} \tag{22.7}$$

We identify the symmetric and antisymmetric wavefunctions with the spin-statistics theorem:

- Particles with symmetric wavefunctions are **bosons** and have **integer spin**. They obey Bose-Einstein statistics.
- Particles with antisymmetric wavefunctions are **fermions** and have **half-integer spin**. They obey Fermi-Dirac statistics.

Thermal & Statistical Physics (PHYS-UA 140) explores the deep physics of this theorem.

22.2 Pauli exclusion principle

Let's try putting both particles in the same state, say $|a,a\rangle$. For bosons, the state becomes

$$|\psi_S\rangle = \frac{2}{\sqrt{2}}|a,a\rangle. \tag{22.8}$$

There is a non-zero probability of this happening and all is well. If we repeat this with fermions, we obtain

$$|\psi_A\rangle = \frac{1}{\sqrt{2}}\left(|a,a\rangle - |a,a\rangle\right) = 0. \tag{22.9}$$

This means there is zero probability this can occur. We deduce an important property of identical fermions called the **Pauli exclusion principle**:

> Two fermions cannot occupy the same state.

22.3 Adding two spin-half particles

We illustrate adding angular momentum in quantum mechanics with two spin-half particles. This finds an important application for the two indistinguishable electrons of a helium atom (e^-, e^-) in section 27. There are four natural ways to write two spin-half particles:

$$|\uparrow\rangle_1|\uparrow\rangle_2, \quad |\downarrow\rangle_1|\downarrow\rangle_2, \quad |\uparrow\rangle_1|\downarrow\rangle_2, \quad |\downarrow\rangle_1|\uparrow\rangle_2. \tag{22.10}$$

The left label denotes particle 1 and the right is particle 2. The summed spin angular momentum operator and its projection along the z axis are

$$\hat{\mathbf{S}}_{12} = \hat{\mathbf{S}}_1 + \hat{\mathbf{S}}_2, \quad \hat{S}_{z,12} = \hat{S}_{z,1} + \hat{S}_{z,2}. \tag{22.11}$$

Here, $\hat{\mathbf{S}}_1$ only acts on particle 1 and $\hat{\mathbf{S}}_2$ only acts on particle 2. We can write the summed spin eigenvalue equations as

$$\hat{S}_{12}^2|\chi_{12}\rangle = S(S+1)\hbar^2|\chi_{12}\rangle, \tag{22.12}$$

$$\hat{S}_{z,12}|\chi_{12}\rangle = M_S\hbar|\chi_{12}\rangle. \tag{22.13}$$

Acting on the combined spin state $|\chi_{12}\rangle = |\chi_1\rangle|\chi_2\rangle$ gives

$$\hat{S}_{z,12}|\chi_{12}\rangle = \left(\hat{S}_{z,1} + \hat{S}_{z,2}\right)|\chi_1\rangle|\chi_2\rangle = \left(\hat{S}_{z,1}|\chi_1\rangle\right)|\chi_2\rangle + \left(\hat{S}_{z,2}|\chi_2\rangle\right)|\chi_1\rangle \tag{22.14}$$

The operator labelled 1 only acts on the state vector with label 1 $\hat{S}_{z,1}|\chi_1\rangle = m_1\hbar|\chi_1\rangle$ and similarly for label 2 $\hat{S}_{z,2}|\chi_2\rangle = m_2\hbar|\chi_2\rangle$. We find

$$\hat{S}_{z,12}|\chi_{12}\rangle = (m_1 + m_2)\hbar|\chi_{12}\rangle. \tag{22.15}$$

We denote the combined spinor state $|\chi_{12}\rangle$ by the eigenvalues of the total magnitude $\hat{S}^2_{12}$ and azimuthal $\hat{S}_{z,12}$ spin

$$|S, M_S\rangle = |\chi_{12}\rangle. \tag{22.16}$$

Given $m_1 = \{+\frac{1}{2}, -\frac{1}{2}\}$ and $m_2 = \{+\frac{1}{2}, -\frac{1}{2}\}$, there are four combinations of total M_S

$$M_S = +1, 0, 0, -1. \tag{22.17}$$

With the eigenvalue constraint $-S \leq M_S \leq +S$, we can immediately identify the $M_S = \pm 1$ states to $S = 1$. These correspond to the following two states from equation (22.10)

$$|1, +1\rangle = |\uparrow\rangle_1|\uparrow\rangle_2, \tag{22.18}$$

$$|1, -1\rangle = |\downarrow\rangle_1|\downarrow\rangle_2. \tag{22.19}$$

This is consistent with our intuition of adding arrowheads: the two spin-half arrows adding to make one unit of spin $1\hbar = \frac{1}{2}\hbar + \frac{1}{2}\hbar$:

$$\xrightarrow{+1\hbar} = \xrightarrow{\frac{\hbar}{2}} + \xrightarrow{\frac{\hbar}{2}}. \tag{22.20}$$

Intuitively, $M_S = 0$ states correspond to spins being anti-aligned so they sum to zero

$$0 = \xrightarrow{\frac{\hbar}{2}} + \xleftarrow{\frac{\hbar}{2}}. \tag{22.21}$$

But there are two distinct states in equation (22.10): $|\uparrow\rangle_1|\downarrow\rangle_2$ and $|\uparrow\rangle_2|\downarrow\rangle_1$. We proceed by applying the combined lowering operator

$$\hat{S}_{-,12} = \hat{S}_{-,1} + \hat{S}_{-,2}. \tag{22.22}$$

Applying this onto the $|1, +1\rangle$ state gives

$$\hat{S}_{-,12}|1, +1\rangle = \hat{S}_{-,1}|\uparrow\rangle_1|\uparrow\rangle_2 + \hat{S}_{-,2}|\uparrow\rangle_1|\uparrow\rangle_2 \tag{22.23}$$

Recast the angular momentum ladder operators from (17.20) into spin notation

$$\hat{S}_{\pm}|s,m_s\rangle = \sqrt{s(s+1)-m_s(m_s\pm 1)}|s,m_s\pm 1\rangle. \tag{22.24}$$

Applying this to both sides of equation (22.23) gives

$$|1,0\rangle = \frac{1}{\sqrt{2}}\left(|\downarrow\rangle_1|\uparrow\rangle_2 + |\uparrow\rangle_1|\downarrow\rangle_2\right). \tag{22.25}$$

There is an orthogonal state that has a negative sign corresponding to the $S=0$ eigenvalue

$$|0,0\rangle = \frac{1}{\sqrt{2}}\left(|\downarrow\rangle_1|\uparrow\rangle_2 - |\uparrow\rangle_1|\downarrow\rangle_2\right). \tag{22.26}$$

We find both $M_S = 0$ states are 50:50 superpositions of $|\downarrow\rangle_1|\uparrow\rangle_2$ and $|\uparrow\rangle_1|\downarrow\rangle_2$, differing by a negative sign. To summarize, two spin-half particles add to form one spin-0 states (*singlet*) state and three spin-1 states (*triplet*) that is sometimes written as "$\frac{1}{2}\oplus\frac{1}{2} = 0\otimes 1$":

$$\text{“singlet spin-0”} = |0,0\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle\right), \tag{22.27}$$

$$\text{“triplet spin-1”} = \begin{cases} |1,+1\rangle & = |\uparrow\rangle|\uparrow\rangle, \\ |1,0\rangle & = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle|\downarrow\rangle + |\downarrow\rangle|\uparrow\rangle\right), \\ |1,-1\rangle & = |\downarrow\rangle|\downarrow\rangle. \end{cases} \tag{22.28}$$

The singlet spin-0 state is antisymmetric upon swapping the labels of particle 1 and 2. The triplet spin-1 state is symmetric upon swapping the labels of particle 1 and 2.

22.4 Separable and entangled states

To introduce a little more terminology, a two-electron system is **separable** if we can write it in a product form

$$|\chi_{12}\rangle = |\chi\rangle_1|\chi\rangle_2. \tag{22.29}$$

This property applies to the following two triplet states:

$$|1,+1\rangle = |\uparrow\rangle_1|\uparrow\rangle_2, \quad |1,-1\rangle = |\downarrow\rangle_1|\downarrow\rangle_2. \tag{22.30}$$

For $|\uparrow\rangle_1|\uparrow\rangle_2$, separable means that if we measure electron 1, it is definitely spin-up and if we measure electron 2, it is definitely spin-up. Similarly, the $|\downarrow\rangle_1|\downarrow\rangle_2$ state is also separable.

By contrast, the other two states $|0,0\rangle$ and $|1,0\rangle$ are **entangled**. Look at the singlet state:

$$|0,0\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle_1|\downarrow\rangle_2 - |\downarrow\rangle_1|\uparrow\rangle_2\right). \tag{22.31}$$

Entangled means we cannot write this in the separable form (22.29). Now if we measure electron 1, there is a 50% probability it is spin-up $|\uparrow\rangle$ or spin-down $|\downarrow\rangle$. Suppose the wavefunction collapses to the spin-up state

$$|\text{before}\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle_1|\downarrow\rangle_2 - |\downarrow\rangle_1|\uparrow\rangle_2\right) \xrightarrow[\text{measurement}]{} |\text{after}\rangle = |\uparrow\rangle_1|\downarrow\rangle_2 \tag{22.32}$$

Upon measuring electron 1, if we get spin-up $|\uparrow\rangle$, we immediately know electron 2 must be spin-down $|\downarrow\rangle$ *even if* we do not measure electron 2. This feature of entanglement is quantum mechanical with no classical analogue.

If "immediately know" sounds unsettling, you're not alone. Quantum information literature [45] entertains Alice and Bob in thought experiments, who carry Stern-Gerlach devices far across the universe, awaiting entangled electron pairs ejected in opposite directions:

$$\boxed{\text{Alice}} \xleftarrow[e_1^- \text{ ejected south}]{} \boxed{\frac{1}{\sqrt{2}}\left(|\uparrow\rangle_1|\downarrow\rangle_2 - |\downarrow\rangle_1|\uparrow\rangle_2\right)} \xrightarrow[e_2^- \text{ ejected north}]{} \boxed{\text{Bob}} \tag{22.33}$$

If Alice measures $|\uparrow\rangle$, the wavefunction collapses and Bob can only measure $|\downarrow\rangle$. Alice immediately knows Bob's measurement result with certainty. This is the Einstein-Podolsky-Rosen (EPR) thought experiment as framed by David Bohm. Section 23 provides some optional reading on this topic in the context of Bell inequalities.

23 Bell inequalities (optional reading)

This class is when mid-term exam 2 takes place.

> These notes instead present optional reading of Bell's inequality that may be delivered at the end of the course if time permits. *The material in this section 23 is optional and not part of the course.*

Wavefunction collapse in the EPR thought experiment (22.33) appears to propagate and influence results across distances faster than light, in conflict with relativity. Einstein critiqued this *spukhafte Fernwirkung* "spooky action at a distance". It also raises metaphysical questions (section 15). Heuristically, does the wavefunction collapse *really physically* (ontology) propagate instantly to Bob? How does Alice *know* Bob's measurement if she cannot see it (epistemology)?

23.1 CHSH inequality

A concrete way forward is to consider whether classical determinism can describe the EPR thought experiment. Suppose we have a pair of old-fashioned classical gloves and seal each one into separate boxes. After shuffling, we randomly pick one box to send to Alice and the other to Bob. When Alice opens the box and finds say a left-handed glove, she immediately knows Bob must have the right-handed glove without any communication or causal influence. This is because the results are *classically correlated*. The glove for Alice was pre-determined to be left-handed from start to finish, but the state was simply hidden. This classical story motivates counterparts for quantum mechanics called *hidden-variables theories*.

Consider a source emitting a pair of electrons as in (22.33):

$$\hat{A}_1, \hat{A}_2 \quad \boxed{\text{Alice}} \xleftarrow[e^-]{} \boxed{\text{Source}} \xrightarrow[e^-]{} \boxed{\text{Bob}} \quad \hat{B}_1, \hat{B}_2$$
$$\searrow \qquad\qquad \swarrow$$
$$\text{Counter} \tag{23.1}$$

Upon receiving the electron, Alice uses Stern-Gerlach device $\hat{A}_1$ or $\hat{A}_2$, which measures binary values $a_1 = \pm 1$ and $a_2 = \pm 1$, respectively. Bob also has two Stern-Gerlach devices $\hat{B}_1$ or $\hat{B}_2$, which measure binary values $b_1 = \pm 1$ and $b_2 = \pm 1$:

$$\hat{A}_1 \to a_1, \quad \hat{A}_2 \to a_2, \quad \hat{B}_1 \to b_1, \quad \hat{B}_2 \to b_2. \tag{23.2}$$

Alice and Bob send their results to a counter for multiplication. To derive a correlation amenable for experimental test, we define the linear combination

$$\hat{S} = (\hat{A}_1 + \hat{A}_2)\hat{B}_1 + (\hat{A}_1 - \hat{A}_2)\hat{B}_2. \tag{23.3}$$

We now make the **locality assumption** that measurements by Alice cannot causally influence those by Bob and vice versa, so their measurements are statistically independent. We next make the **realism assumption** that a hidden variable (λ) pre-determines the measured values. Like the glove story with pre-determined handedness from start to finish, if Alice uses Stern-Gerlach $\hat{A}_1$ and measures $+1$, the electron had the $a_1 = +1$ property since its emission at source. The values are well-defined such that

$$S_\lambda = (a_1 + a_2)b_1 + (a_1 - a_2)b_1. \tag{23.4}$$

The linear combination is designed such that one of the parentheses must vanish. This arises because either $a_1 = \pm 1$ or $a_2 = \pm 1$, so there are two possibilities:

$$a_1 = +a_2 \quad \Rightarrow \quad S_\lambda = \pm 2b_1, \tag{23.5}$$

$$a_1 = -a_2 \quad \Rightarrow \quad S_\lambda = \pm 2b_2. \tag{23.6}$$

Either way, we find $S_\lambda = \pm 2b_i$, where b_i can be either ± 1. Repeating the experiment many times, the average values of b_i, denoted $\langle b_i \rangle$, must satisfy $-1 \leq \langle b_i \rangle \leq +1$. This implies

$$-2 \leq \langle S_\lambda \rangle \leq +2. \tag{23.7}$$

This is the Clauser-Horne-Shimony-Holt (CHSH) inequality, presented by John Clauser, Michael Horne, Abner Shimony, and Richard Holt in 1969 [46]. This adapts the inequality by John Bell from 1964 [35] to be more amenable for experimental tests.

23.2 Quantum correlations

We now calculate $\langle S \rangle$ for quantum mechanics. Consider the entangled singlet state

$$|\chi_{0,0}\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle_A|\downarrow\rangle_B - |\downarrow\rangle_A|\uparrow\rangle_B). \tag{23.8}$$

The spin-up $|\uparrow\rangle$ and spin-down $|\downarrow\rangle$ states are eigenstates of the $\hat{\sigma}_z$ Pauli operator. We use the Pauli matrices (23.9) with $\phi = 0$, rotating via the polar coordinate $\hat{\sigma}_\theta$ in the $y-z$ plane

$$\hat{\sigma}_\theta = \hat{\boldsymbol{\sigma}} \cdot \boldsymbol{\theta} = \hat{\sigma}_x \sin\theta + \hat{\sigma}_z \cos\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}. \tag{23.9}$$

Alice aligns one Stern-Gerlach $\hat{A}_1$ along the z-axis, $\hat{A}_1 = \hat{\sigma}_z$ ($\theta = 0$) and the other $\hat{A}_2 = \hat{\sigma}_x$ ($\theta = \pi/2$). Bob aligns his Stern-Gerlach devices mutually perpendicular but rotated 45° relative to Alice so $\hat{B}_1 = \hat{\sigma}_{\theta=\pi/4}$ and $\hat{B}_1 = \hat{\sigma}_{\theta=3\pi/4}$. This arrangement maximizes the CHSH quantity. Evaluating all correlations using the singlet state $\langle \hat{A}_i \hat{B}_j \rangle = \langle \chi_{0,0} | \hat{A}_i \hat{B}_j | \chi_{0,0} \rangle$ gives

$$\langle \hat{A}_1 \hat{B}_1 \rangle = \langle \hat{A}_1 \hat{B}_2 \rangle = \langle \hat{A}_2 \hat{B}_1 \rangle = -\frac{1}{\sqrt{2}}, \quad \langle \hat{A}_2 \hat{B}_2 \rangle = +\frac{1}{\sqrt{2}}. \tag{23.10}$$

The quantum correlation result can therefore exceed the CHSH inequality (23.7)

$$\langle \hat{S}_{\text{quantum}} \rangle = -2\sqrt{2} < -2. \tag{23.11}$$

To test this prediction in practice, photons are preferred for experimental control. In 1950, Chien-Shiung Wu and her student Irving Shaknov experimentally demonstrated entangled photon pairs arising from electron-positron annihilation [47]. At Orsay, France in 1982, Alain Aspect, Philippe Grangier, and Gérard Roger used polarized photons from a calcium source to empirically measure [48]

$$\langle S_{\text{experiment}} \rangle = 2.70 \pm 0.05 \tag{23.12}$$

This disagrees with the CHSH inequality and favours the quantum prediction. Quantum mechanics questions the assumptions of (i) locality or (ii) hidden variables.

V Atomic spectra

24 Hydrogen gross structure

Let us return to hydrogen, whose importance cannot be understated given Cecilia Payne [13] discovered it is the dominant element in stars and thus the universe. Three observables describe hydrogen: the Hamiltonian $\hat{H}$, total orbital angular momentum $\hat{L}^2$ and its azimuthal projection $\hat{L}_z$. Three quantum numbers $|n,\ell,m_\ell\rangle$ define the electron state, which specifies the *gross structure* of hydrogen

$$\begin{aligned} n: &\quad \text{principal quantum number for energy } \hat{H},\\ \ell: &\quad \text{orbital angular momentum for } \hat{L}^2,\\ m_\ell: &\quad \text{azimuthal angular momentum for } \hat{L}_z. \end{aligned}$$

The eigenvalues of $\hat{H}$ are the energy levels E_n from equation (19.28)

$$E_n = -\left(\frac{1}{2}\alpha^2 m_e c^2\right)\frac{1}{n^2} = -\frac{E_R}{n^2} \simeq -\frac{13.6\text{ eV}}{n^2}. \tag{24.1}$$

Table 3 summarizes useful physical quantities that set the scales and dynamics of hydrogen.

Symbol	Expression	Value (3 s.f.)	Quantity
α	$\dfrac{e^2}{4\pi\varepsilon_0\hbar c}$	$1/137$	Fine-structure constant
E_R	$\frac{1}{2}\alpha m_e c^2$	13.6 eV	Rydberg energy
r_B	$\dfrac{\hbar}{\alpha m_e c}$	0.529 Å	Bohr radius
μ_B	$\dfrac{e\hbar}{2m_e}$	$h\times 14.0$ GHz T^{-1}	Bohr magneton
g_s	$2\left(1+\frac{\alpha}{\pi}+\ldots\right)$	2×1.00116	Electron spin g-factor

Table 3: Useful physical quantities relevant for atomic physics to three significant figures.

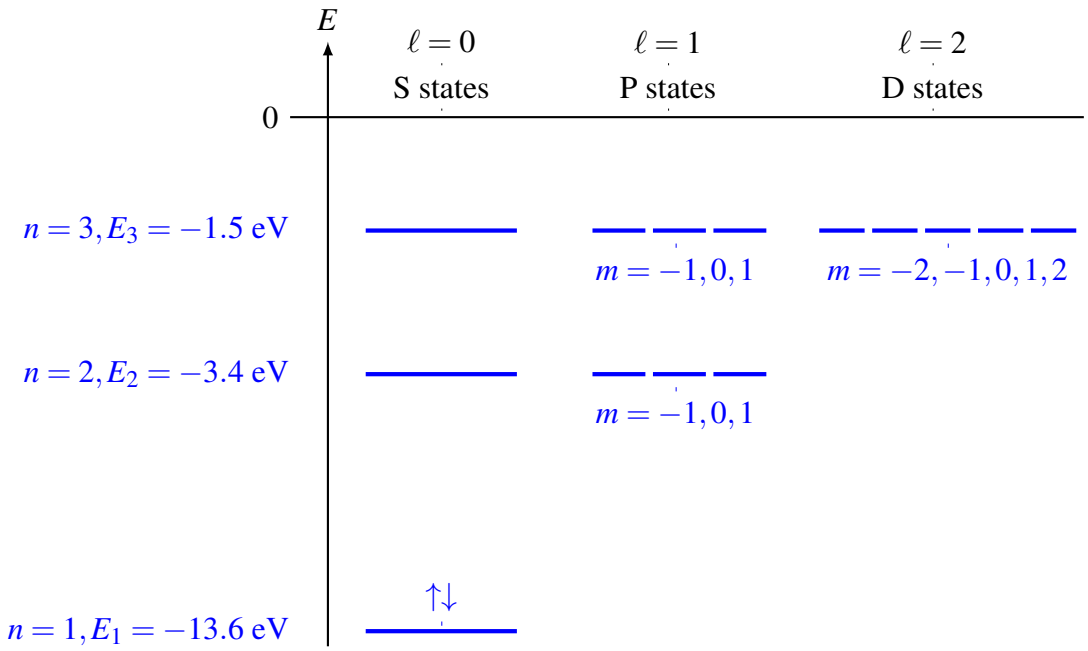


Figure 35: Gross structure of hydrogen atom $E_n = -13.6\text{ eV}/n^2$ for $n = 1,2,3$. The degenerate states are shown for $\ell = 0,1,2$ along with the m quantum number states. Electrons can occupy each state in one the spin-up $\uparrow$ and spin-down $\downarrow$ states.

24.1 Degeneracy and chemistry

We introduce the concept of *degeneracy*, which counts the linearly independent eigenstates sharing the same energy. The Bohr formula for hydrogen energies (19.28) only depends on n while the angular momentum quantum numbers (ℓ, m_ℓ) may take multiple values

$$\ell = 0,1,2,\ldots,n-1 \tag{24.2}$$

$$m_\ell = -\ell, -\ell+1, \ldots, 0, \ldots, \ell-1, \ell. \tag{24.3}$$

Given a value of ℓ, we have $2\ell+1$ states specified by m_ℓ. The degeneracy of each energy level scales quadratically with n:

$$\sum_{\ell=0}^{n-1}(2\ell+1) = 2\frac{n(n-1)}{2} + n = n^2. \tag{24.4}$$

Figure 35 shows how the degenerate states $|n,\ell,m_\ell\rangle$ correspond to the angular momentum (ℓ, m_ℓ) quantum numbers for each energy quantum number n.

- For $n = 1$, there is one distinct state with one value of $\ell = 0$ and $m_\ell = 0$: $|1,0,0\rangle$

Shell n	Subshell	ℓ	Electron states $(\uparrow\downarrow)\times(\ell, m_\ell)$
1	1s	0	$2 = 2\times 1$
2	2s2p	0,1	$8 = 2\times(1+3)$
3	3s3p3d	0,1,2	$18 = 2\times(1+3+5)$
4	4s4p4d4f	0,1,2,3	$32 = 2\times(1+3+5+7)$

Table 4: Electronic shell degeneracy of states.

- For $n = 2$, ℓ takes two values: $\ell = 0, 1$. The m_ℓ quantum number takes three values $m_\ell = -1, 0, 1$ for $\ell = 1$, and one value $m_\ell = 0$ for $\ell = 0$. So the four states are

$$n = 2: \quad |2,0,0\rangle, \quad |2,1,0\rangle, \quad |2,1,1\rangle, \quad |2,1,-1\rangle. \tag{24.5}$$

From the Pauli exclusion principle, electrons can only occupy distinct states. Each energy level is $2\times n^2$-fold degenerate due to two spin states: $|\uparrow\rangle, |\downarrow\rangle$. Electrons occupy each energy level called *shells*. Table 4 shows this filling of electron shells. Spectroscopic notation L label the ℓ states called *subshells* (or suborbitals)

$$\text{L} = \text{s, p, d, f} \quad \longleftrightarrow \quad \ell = 0, 1, 2, 3. \tag{24.6}$$

These historically stand for sharp, principal, diffuse and fundamental, e.g. nL = 2p denotes $n = 2, \ell = 1$. The filling of electron subshells then follows the periodic table of chemistry:

- 1s shell: hydrogen and helium.
- 2s subshell: lithium, beryllium.
- 2p subshell: boron, carbon, nitrogen, oxygen, flourine and neon.

An electron shell is "full" when all its spin states are occupied, which correspond to the noble gases: helium, neon, argon. In chemistry, *Hund's rules* systematize the shell filling of electrons. For more chemistry perspectives and spectroscopy, see CHEM-UA 651. It is remarkable that quantum physics explains the structure of the periodic table from first principles. Quantum mechanics plus electromagnetism is the basis of chemistry and life! In 1949, Maria Goeppert Mayer [49] and Hans Jensen [50] extended the shell model to explain the relative stability of atomic nuclei.

There are further refinements we can consider:

- **Reduced mass**. The hydrogen gross structure analysis effectively treats the proton as a stationary point in the limit of infinite mass. Like orbital mechanics, we could the

electron mass with the reduced mass of the hydrogen system $m_e \to \mu$ given by

$$\mu = \frac{m_e m_p}{m_e + m_p} = \frac{m_e}{1 + m_e/m_p}. \tag{24.7}$$

Here, m_e is the electron mass and m_p is the proton mass. Given $m_e/m_p \simeq 1/1836$, this is a small but non-negligible correction.

- **Hydrogenic atoms**. For atoms with proton number Z, we can ionize all but one electron: helium He^+, lithium Li^{2+}, and so on. These hydrogenic atoms are two-body problems and enable us to reuse the exact solutions for hydrogen. We replace the proton charge e with Ze in the Coulomb potential $V_Z(r) = -Ze^2/4\pi\varepsilon_0 r$, so the fine-structure constant is rescaled $\alpha \to Z\alpha$. The energy is rescaled by a factor of Z^2

$$E_n \to E_{n,Z} = -\frac{1}{2}(Z\alpha)^2 m_e c^2 \simeq -\frac{Z^2 13.6\text{ eV}}{n^2}. \tag{24.8}$$

The Bohr radius r_B is rescaled by $1/Z$ as the electron is pulled closer to the nucleus

$$r_B \to \frac{r_B}{Z} = \frac{\hbar}{Z\alpha m_e c}. \tag{24.9}$$

24.2 Hydrogen wavefunctions

Recalling equation (19.13) from section 19, the hydrogen radial wavefunctions $R_{n,\ell}(r)$ are

$$R_{n,\ell}(r) = [\text{normalization}] \times [\text{near origin}] \times [\text{intermediate}] \times [\text{far from origin}] \tag{24.10}$$

$$= A_{n,\ell} \times r^\ell \times y_n^\ell(r) \times \mathrm{e}^{-r/nr_B}, \tag{24.11}$$

where $y_n^\ell(r)$ is a polynomial of degree $j_{\max} = n - \ell - 1$

$$y_n^\ell(r) = \sum_{b=0}^{n-\ell-1} c_b r^b. \tag{24.12}$$

The recurrence relation (19.21) fully defines these generalized Laguerre polynomials, which are tabulated in textbooks and are often written as

$$y_n^\ell(r) = L_{n-\ell-1}^{2\ell+1}\left(\frac{2r}{nr_B}\right). \tag{24.13}$$

The constant $A_{n,\ell}$ is determined by the normalization condition

$$\int_0^\infty |R_{n,\ell}(r)|^2 r^2 \mathrm{d}r = 1. \tag{24.14}$$

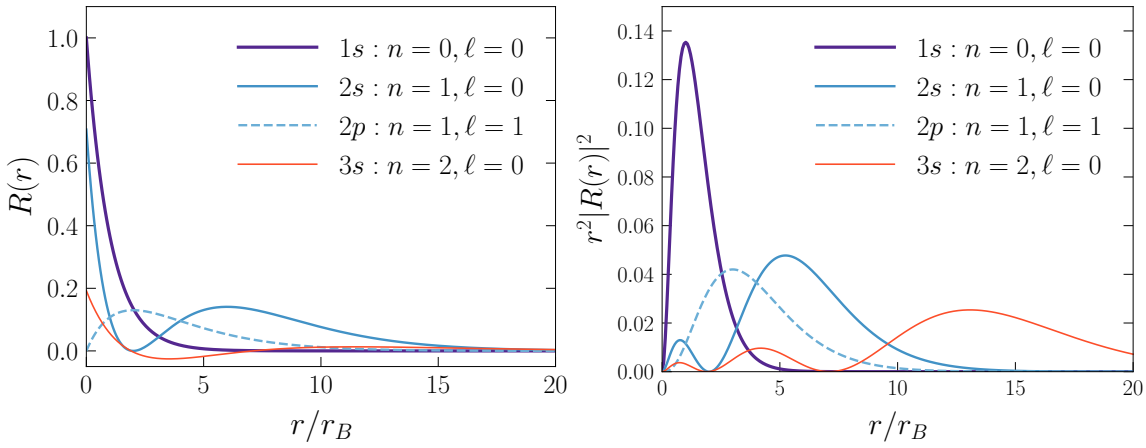


Figure 36: Hydrogen radial wavefunctions $\psi \propto R(r)$ and probability densities $|\psi|^2 \propto r^2|R(r)|^2$ plotted with `matplotlib`.

The full hydrogen wavefunction is the product of $R_{n,\ell}(r)$ the spherical harmonics $Y_\ell^m(\theta,\phi)$ detailed in section 18.1

$$\psi_{n,\ell,m}(r,\theta,\phi) = R_{n,\ell}(r)Y_\ell^m(\theta,\phi). \tag{24.15}$$

In the ground state $n = 1, \ell = 0$, the normalized radial eigenfunction is

$$R_{1,0} = 2\left(\frac{1}{r_B}\right)^{3/2} \mathrm{e}^{-r/r_B}. \tag{24.16}$$

The $\ell = 0, m_\ell = 0$ spherical harmonic is $Y_0^0(\theta,\phi) = 1/\sqrt{4\pi}$, so the full wavefunction is

$$\psi_{100}(r,\theta,\phi) = \frac{1}{\sqrt{\pi}}\left(\frac{1}{r_B}\right)^{3/2} \mathrm{e}^{-r/r_B}. \tag{24.17}$$

We see that the Bohr radius r_B indeed sets the characteristic size of the hydrogen ground state. There is zero orbital angular momentum in the ground state. The e^{-r/r_B} dependence is reminiscent of the bound state in the delta function potential we studied way back in one-dimensional wave mechanics of section 5.2! The probability density for an electron to be in the spherical shells r and $r + \mathrm{d}r$ is the modulus square

$$\rho_{100}(r)\mathrm{d}r = |R_{1,0}(r)|r^2\mathrm{d}r = \frac{1}{\pi}\left(\frac{1}{r_B}\right)^3 r^2\mathrm{e}^{-2r/r_B}. \tag{24.18}$$

The maximum of this probability density corresponds to the radius where we are most likely to find an electron

$$\frac{\mathrm{d}\rho_{100}(r)}{\mathrm{d}r} = 0 = \frac{1}{\pi}\left(\frac{1}{r_B}\right)^3 2r\left(1 - \frac{r}{r_B}\right)\mathrm{e}^{-2r/r_B}. \tag{24.19}$$

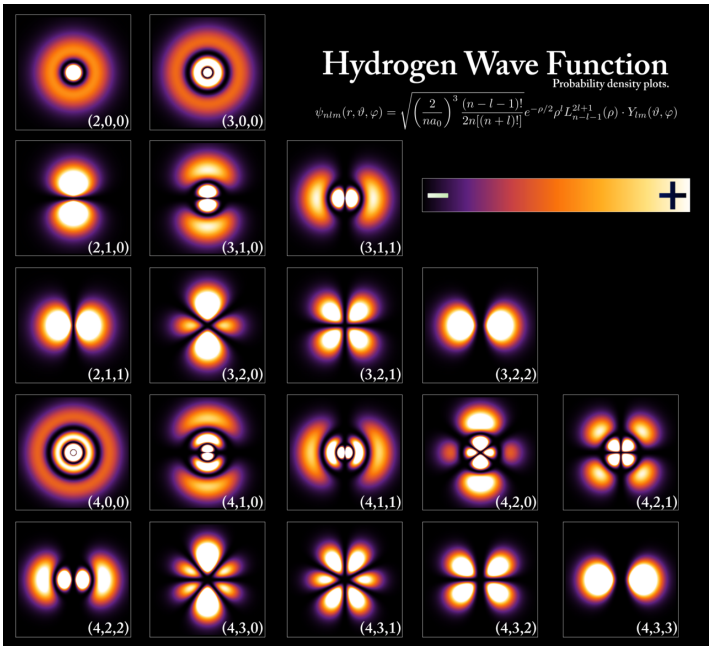


Figure 37: Hydrogen probability density plots in a two-dimensional plane. Image: PoorLeno

This demonstrates that there is a maximum at the Bohr radius $r_B = \hbar/m_e c\alpha \simeq 0.529$ Å.

The first excited $n = 2, \ell = 0$ and $n = 2, \ell = 1$ states are

$$\psi_{200} = \frac{1}{\sqrt{8\pi}} \left(\frac{1}{r_B}\right)^{3/2} \left(1 - \frac{r}{2r_B}\right) \mathrm{e}^{-r/2r_B}, \tag{24.20}$$

$$\psi_{210} = \frac{1}{\sqrt{32\pi}} \left(\frac{1}{r_B}\right)^{3/2} \cos\theta \frac{r}{r_B} \mathrm{e}^{-r/2r_B}. \tag{24.21}$$

Figure 36 plots the radial wavefunctions and probability densities of the electron in hydrogen. Note how $n = 2, \ell = 0$ has a node with zero probability at $r = 2r_B$ and the probability peaks further out than $n = 1$. Figure 37 includes the angular dependence in θ. They are beautiful and strikingly different from the planetary orbit cartoons of the Bohr model!

HW10: Spin half and hydrogen (set Nov 24, due Dec 3)

1. The spin operator is $\hat{\mathbf{S}} = \hat{\boldsymbol{\sigma}}/2$. Write down the three 2×2 Pauli matrices $\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z$.

   (a) Describe, in terms of operators and eigenvalues, the physical and mathematical similarities and differences between electron spin and orbital angular momentum.

   (b) By explicitly matrix multiplication, show that: (i) commutator $[\hat{S}_y, \hat{S}_z] = \mathrm{i}\hbar\hat{S}_x$, and (ii) anticommutators $\{\hat{\sigma}_x, \hat{\sigma}_y\} = 0$ and $\{\hat{\sigma}_x, \hat{\sigma}_x\} = 2\hat{I}$, where $\hat{I}$ is the identity.

2. A system of two Stern-Gerlach (SG) machines sees the first aligned along the z-axis then a second SG device aligned along $\mathbf{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)^{\mathrm{T}}$:

$$\xrightarrow[\text{spin-half particles}]{} \boxed{\mathrm{SG}_A} \xrightarrow[\text{transmits } |\uparrow\rangle \text{ along } \mathbf{z}]{} \boxed{\mathrm{SG}_B} \xrightarrow[\text{transmits } |\lambda_+\rangle \text{ along } \mathbf{n}]{} \tag{24.22}$$

   (a) Electrons in state $|\chi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$ enter SG_A. Show that the expectation value of the force on the electrons is $\langle\hat{f}_z\rangle = (|\alpha|^2 - |\beta|^2)\mu_B\frac{\partial B_z}{\partial z}$, where $\mu_B = e\hbar/2m_e$.

   (b) An electron enters SG_B in a spin-up state $|\uparrow\rangle$. By solving the eigenvalue problem $\hat{\boldsymbol{\sigma}}\cdot\mathbf{n}|\lambda_-, \mathbf{n}\rangle = \lambda_-|\lambda_-, \mathbf{n}\rangle$, show that the probability of measuring the spin-down state $|\lambda_-, \mathbf{n}\rangle$ after exiting SG_B is $P(\lambda_-, \mathbf{n}|\uparrow) = \sin^2\left(\frac{\theta}{2}\right)$.

3. Consider a system of two identical electrons.

   (a) Explain the concept of identical particles and the Pauli exclusion principle.

   (b) By applying the combined spin raising operator $\hat{S}_{+,12} = \hat{S}_{+,1} + \hat{S}_{+,2}$ onto the $|1,-1\rangle = |\downarrow\rangle_1|\downarrow\rangle_2$ state, show that this gives $|1,0\rangle = \frac{1}{\sqrt{2}}(|\downarrow\rangle_1|\uparrow\rangle_2 + |\uparrow\rangle_1|\downarrow\rangle_2)$. Explain why the $|1,0\rangle$ state is considered entangled whereas $|1,-1\rangle$ is separable?

4. Describe the three quantum numbers and their physical interpretation that characterize the states of hydrogen gross structure. How are their allowed values?

   (a) Express the energies E_n in terms of the fine-structure constant α and electron mass m_e. What is the value of E_1 in eV and Joules? How does this change if we use the reduced mass $\mu = m_e m_p/(m_e + m_p)$ instead of a stationary proton?

   (b) What states can electrons occupy in the ground, first and second excited energies? What does *degeneracy* mean in this context? Explain your answers by counting states with the hydrogen quantum numbers and concept of electron (sub)shells.

   (c) What is meant by a hydrogenic atom? Write the expressions for the energy formula, ground and first excited state wavefunctions in terms of atomic number Z. How does the most probable radial electron position scale with Z?

25 Time-independent perturbation theory

So far, we have studied special cases that enjoy exact solutions: step potentials, harmonic oscillator, and hydrogen. Exact solutions are rare in nature. We take a first look at a powerful approximation method called time-independent perturbation theory applied to atoms.

25.1 Non-degenerate theory

The idea of perturbation theory is to add a small change to the Hamiltonian. Colloquially, we gently prod an atom. We wish to find solutions to the Hamiltonian

$$\hat{H} = \hat{H}^{(0)} + \lambda \hat{H}^{(1)}. \tag{25.1}$$

Here, we write the unperturbed $\hat{H}^{(0)}$ and perturbation $\hat{H}^{(1)}$ Hamiltonian, with $\lambda \ll 1$ being a small dimensionless perturbation parameter. The thesis of perturbation theory is: we know the exact solutions for $\hat{H}^{(0)}$ and we want the correction terms for our wavefunctions and eigenvalues for the perturbation $\hat{H}^{(1)}$. We expand the eigenstates and eigenvalues as a perturbation series

$$\begin{aligned} |n\rangle &= |n^{(0)}\rangle + \lambda |n^{(1)}\rangle + \lambda^2 |n^{(2)}\rangle + \dots \\ E_n &= E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \end{aligned} \tag{25.2}$$

This is a similar spirit to a binomial expansion in a small x:

$$(1+x)^n = 1 + nx + \frac{1}{2} n(n-1)x^2 + \dots, \quad x \ll 1. \tag{25.3}$$

We adopt the following notation for perturbation theory:

- superscripts in brackets show the perturbation order, e.g. $E^{(0)}$ = zeroth order energy;
- subscripts such as E_n denote the n-th energy level;
- $|n\rangle$ denotes the n-th eigenstate.

Substitute the perturbation series (25.2) into Schrödinger's equation $\hat{H}|n\rangle = E_n|n\rangle$ and keep terms to first order

$$\hat{H}^{(1)}|n^{(1)}\rangle + \hat{H}^{(1)}|n^{(0)}\rangle = E_n^{(0)}|n^{(1)}\rangle + E_n^{(1)}|n^{(0)}\rangle. \tag{25.4}$$

Multiply from the left with $\langle n^{(0)}|$ lets us isolate the energy correction $E_n^{(1)}$

$$E_n^{(1)} = \langle n^{(0)}|\hat{H}^{(1)}|n^{(0)}\rangle. \tag{25.5}$$

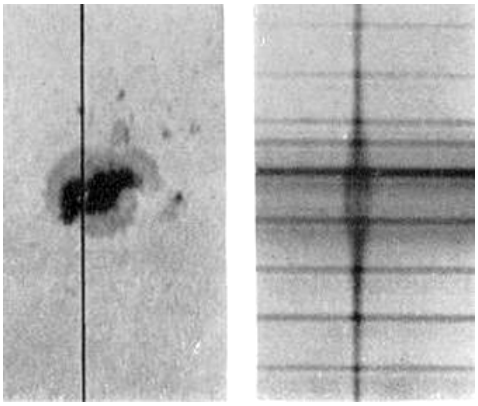


Figure 38: Image of Lorentz-Zeeman splitting of the iron 617.3 nm line in the magnetic field of the Sun's sunspot taken in 1916 [51].

This is the expectation value of the perturbation $H^{(1)}$ in the unperturbed eigenbasis $|n^{(0)}\rangle$. We can expand the first order eigenstate correction $|n^{(1)}\rangle$ in the unperturbed basis $|m^{(0)}\rangle$

$$|n^{(1)}\rangle = \sum_{m\neq n} |m^{(0)}\rangle\langle m^{(0)}|n^{(1)}\rangle \tag{25.6}$$

Multiply through (25.4) by $\langle m^{(0)}|$ to evaluate the expansion coefficients

$$(E_n^{(0)} - E_m^{(0)})\langle m^{(0)}|n^{(1)}\rangle = \langle m^{(0)}|\hat{H}^{(1)}|n^{(0)}\rangle \tag{25.7}$$

This gives the first-order correction to the eigenstates as

$$|n^{(1)}\rangle = \sum_{m\neq n} \frac{\langle m^{(0)}|\hat{H}^{(1)}|n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}|m^{(0)}\rangle. \tag{25.8}$$

For second-order perturbation theory, the energy corrections are given by

$$E_n^{(2)} = \langle n^{(0)}|\hat{H}^{(1)}|n^{(1)}\rangle. \tag{25.9}$$

Upon using (25.8), we recast this as

$$E_n^{(2)} = \sum_{m\neq n} \frac{|\langle m^{(0)}|\hat{H}^{(1)}|n^{(0)}\rangle|^2}{E_n^{(0)} - E_m^{(0)}}. \tag{25.10}$$

25.2 Zeeman effect

Equipped with more powerful gratings in 1896, Pieter Zeeman reattempted Michael Faraday's unsuccessful attempts to see if magnetism affects light. Zeeman found emission lines

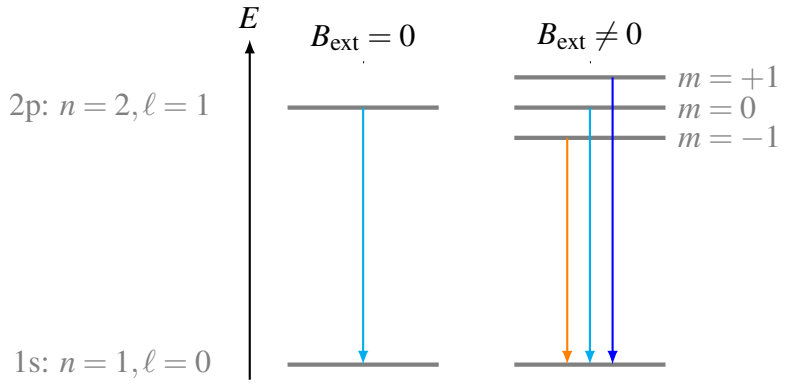


Figure 39: Schematic energy diagram of normal Zeeman splitting for B_{ext} on and off.

from sodium broadening and cadmium splitting when switching on a strong $\mathcal{O}(1)$ T magnetic field. Before quantum mechanics, these observations were understood using Hendrik Lorentz's electron force theory.

We model this splitting using perturbation theory of quantum mechanics. Applying an external magnetic field $\mathbf{B}_{\text{ext}}$ aligned along the z-axis, the atomic magnetic moment $\hat{\boldsymbol{\mu}}_{\text{atom}}$ precesses according a Hamiltonian

$$\Delta\hat{H}_{\text{mag}} = -\hat{\boldsymbol{\mu}}_{\text{atom}} \cdot \mathbf{B}_{\text{ext}}. \tag{25.11}$$

We define the strength of the fields by comparing the magnetic energies with the spin-orbit energy $\hat{H}_{\text{magnetic}} \ll \hat{H}_{\text{spin-orbit}}$ (discussed in the next section). The atomic magnetic moment $\hat{\boldsymbol{\mu}}_L = -g_L\mu_B\hat{\mathbf{L}}$ performs Larmor precession around the z-axis with energy

$$\begin{aligned}\Delta E_{\text{Zeeman}} &= \langle -\hat{\mathbf{L}} \cdot \mathbf{B}_{\text{ext}} \rangle = g_L\mu_B \langle \hat{\mathbf{L}} \cdot \mathbf{B}_{\text{ext}} \rangle \\ &= g_L\mu_B B_{\text{ext}} \langle \hat{L}_z \rangle.\end{aligned} \tag{25.12}$$

With a weak magnetic-field perturbation, the atom is in the ordinary eigenstates $|n, \ell, m\rangle$, where we only consider ℓ and m_ℓ with $g_L = 1$. Thus the energy shifts evaluate to

$$\Delta E_{\text{Zeeman}} = \mu_B B_{\text{ext}} m_\ell. \tag{25.13}$$

As m_ℓ satisfies $-\ell \leq m_\ell \leq +\ell$, there are $(2\ell+1)$ sub-levels of energy for each ℓ. The following changes in m_ℓ are allowed, called selection rules:

$$\Delta m_\ell = 0, \pm 1. \tag{25.14}$$

This gives the *Lorentz-Zeeman triplet*: three lines with frequencies ω_0 and $\omega_0 \pm \Delta\omega_Z$. This simplified treatment illustrates the main idea and is called the *normal Zeeman effect*. The

full treatment is complicated by the electron possessing its own magnetic moment from its intrinsic spin $\hat{\mathbf{S}}$. We would need to consider the total angular momentum $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$ coupling to the external magnetic field $\hat{\mathbf{J}} \cdot \mathbf{B}_{\text{ext}}$. This gives rise to the *anomalous Zeeman effect*, which has this historical name as it was not understood before electron spin was discovered. This effect is detailed in the Quantum Mechanics II (PHYS-UA 124) course.

25.3 Degenerate perturbation theory

We next introduce the idea of degenerate perturbation theory. Suppose we have a *two-fold degeneracy*: there are two states $|n_1\rangle, |n_2\rangle$ corresponding to one energy E_n. We write the perturbation $H^{(1)}$ as a 2×2 matrix in the basis of the degenerate states $|n_1\rangle, |n_2\rangle$

$$\begin{pmatrix} \langle n_1|H^{(1)}|n_1\rangle & \langle n_1|H^{(1)}|n_2\rangle \\ \langle n_2|H^{(1)}|n_1\rangle & \langle n_2|H^{(1)}|n_2\rangle \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = E_n^{(1)} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}. \tag{25.15}$$

We solve this eigenvalue problem to obtain two values for the energy corrections $E_n^{(1)} = E_\pm$. The resulting eigenstates are linear combinations of the original Hamiltonian and exact eigenstates of $H^{(1)}$. This process "lifts" the degeneracy: the perturbation $H^{(1)}$ turns two states corresponding to one energy into two states corresponding to two energies. This process diagonalizes the $H^{(1)}$ matrix in the subspace of degenerate states.

Let us illustrate this concept with an example called the *Stark-Lo Surdo effect*. We apply a weak external electric field aligned with the z-axis to hydrogen atom. Writing the unperturbed eigenfunctions in the form $|n,\ell,m\rangle$, the first excited states are

$$|2,0,0\rangle, \quad |2,1,-1\rangle \quad |2,1,0\rangle \quad |2,1,-1\rangle. \tag{25.16}$$

To see why only matrix elements with the same m contribute, consider only the ϕ integral of such matrix elements

$$\begin{aligned} \langle m'|m\rangle &= \frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{-\mathrm{i}m'\phi}\mathrm{e}^{\mathrm{i}m\phi}\mathrm{d}\phi = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{\mathrm{i}(m-m')\phi}\mathrm{d}\phi \\ &= \begin{cases} \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} \mathrm{d}\phi = 1 & \text{when } m = m' \\ \dfrac{\overbrace{\mathrm{e}^{\mathrm{i}(m-m')2\pi}}^{=1} - 1}{2\pi\mathrm{i}(m-m')} = 0 & \text{when } m \neq m' \end{cases} \end{aligned} \tag{25.17}$$

This shows that the matrix elements are non-zero only if $m = m'$ contribute. We now apply degenerate perturbation theory (25.15). With the perturbation $\Delta\hat{H}_{\text{electric}} = e\mathscr{E}\hat{z}$, the perturbation matrix is

$$\Delta\hat{H}_{\text{electric}} = e\mathscr{E}\begin{pmatrix} \langle 2,0,0|\hat{z}|2,0,0\rangle & \langle 2,0,0|\hat{z}|2,1,0\rangle \\ \langle 2,1,0|\hat{z}|2,0,0\rangle & \langle 2,1,0|\hat{z}|2,1,0\rangle \end{pmatrix}. \tag{25.18}$$

The diagonal elements vanish because of parity. The eigenfunctions with $\ell = 0$ have even parity while $\ell = 1$ have odd parity while z is odd parity so the integral over all space for the diagonal elements vanish. Evaluating the off-diagonal matrix elements gives

$$\langle 2,0,0|z|2,1,0\rangle = -3r_B \tag{25.19}$$

so the perturbation matrix becomes

$$\Delta\hat{H}_{\text{electric}} = -3r_B e\mathscr{E}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{25.20}$$

The electric field perturbation therefore gives energy shifts of

$$\Delta E_{\pm} = \pm 3r_B e\mathscr{E}. \tag{25.21}$$

This lifts the degeneracy into two non-degenerate higher and lower energy levels. The corresponding eigenstates are linear combinations of the unperturbed Hamiltonian's eigenstates

$$\Delta E_{+} = +3r_B e\mathscr{E}: \quad |\psi_{+}\rangle = \frac{1}{\sqrt{2}}(|2,0,0\rangle + |2,1,0\rangle) \tag{25.22}$$

$$\Delta E_{-} = -3r_B e\mathscr{E}: \quad |\psi_{-}\rangle = \frac{1}{\sqrt{2}}(|2,0,0\rangle - |2,1,0\rangle). \tag{25.23}$$

In 1926 at McGill University, John Stuart Foster and Mary Laura Chalk performed detailed measurements of hydrogen in an electric field [52, 53]. Jane Dewey of Copenhagen published analogous measurements of electric fields applied to helium [54]. These overlooked results are among the earliest and important tests of the then nascent predictions of Schrödinger's equation. They established key experimental evidence supporting the quantum mechanics we learn today.

26 Fine structure

26.1 Relativistic kinematic correction

We now apply time-independent perturbation theory to refine the gross structure picture of hydrogen. These small corrections are called fine structure. Let us treat relativistic kinematic motion as a perturbation. Recall the energy-momentum relation from special relativity $E^2 + (pc)^2 = (mc^2)^2$. We can binomial expand the relativistic kinetic energy

$$E_{\text{kin.}} = \sqrt{(m_e c^2)^2 + (pc)^2} - m_e c^2 = \frac{p^2}{2m_e}\left[1 - \frac{1}{4}\left(\frac{p}{m_e c}\right)^2 + \dots\right]. \tag{26.1}$$

The first term is the non-relativistic kinetic energy. The second term is the first-order relativistic correction, which we can model as a time-independent perturbation Hamiltonian

$$\hat{H}_{\text{rel. kin.}} = -\frac{1}{8}\frac{\hat{p}^4}{m_e^3 c^2}. \tag{26.2}$$

This perturbation commutes with the angular momentum operators $[\hat{H}_1, \hat{L}^2] = 0$ and $[\hat{H}_1, \hat{L}_z] = 0$. This implies the off diagonal terms of degenerate perturbation theory are zero

$$\langle n, \ell, m | \hat{H}_1 | n, \ell', m' \rangle = 0 \quad \text{if } \ell \neq \ell', m \neq m'. \tag{26.3}$$

The perturbation is smaller than the non-relativistic kinetic energy by

$$\frac{p^2}{m_e^2 c^2} = \frac{v^2}{c^2} \simeq \alpha^2. \tag{26.4}$$

We can thus use non-degenerate perturbation theory. The unperturbed Schrödinger equation is $\hat{p}^2 |n\ell m\rangle = 2m_e(E_n - V(r))|n\ell m\rangle$. So we can write the perturbation Hamiltonian as

$$\hat{H}_{\text{rel. kin.}} = -\frac{1}{8m_e^3 c^2}\hat{p}^4 = -\frac{1}{2m_e c^2}\left(\hat{H}_0 - V(r)\right)^2 \tag{26.5}$$

The evaluation of the energy becomes

$$\Delta E_{\text{rel. kin.}} = \langle \hat{H}_1 \rangle = \langle n, \ell, m | \hat{H}_1 | n, \ell, m \rangle = -\frac{1}{2m_e c^2}\left(E_n^2 - 2E_n \langle V(r) \rangle + \langle V^2(r) \rangle\right)^2. \tag{26.6}$$

The expectation values that we quote identities from textbooks

$$\left\langle \frac{1}{r} \right\rangle_{n\ell} = \frac{1}{r_B n^2}, \quad \left\langle \frac{1}{r^2} \right\rangle_{n\ell} = \frac{1}{r_B^2 n^3 (\ell + 1/2)}. \tag{26.7}$$

We obtain the relativistic energy shift $\Delta E_{\text{rel. kin.}}$ depends on ℓ, lifting the degeneracy of states

$$\Delta E_{\text{rel. kin.}} = \langle \hat{H}_1 \rangle = -\frac{m_e c^2}{2}\left(\frac{\alpha}{n}\right)^4 \left(\frac{n}{\ell + 1/2} - \frac{3}{4}\right). \tag{26.8}$$

26.2 Spin-orbit coupling

We take a first look at spin-orbit coupling $\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$. In classical electrostatics, the proton has an electric field $\mathbf{E} = -\boldsymbol{\nabla}\phi(r)$ given by the gradient of the Coulomb potential $\phi(r) = Ze/4\pi\varepsilon_0 r$. From the frame of the electron, the proton is moving with velocity $\mathbf{v}$. A moving charge is a non-zero current that produces a magnetic field $\mathbf{B}$ by the Biot-Savart law

$$\mathbf{B} = \frac{\mu_0}{4\pi}\frac{q\mathbf{v}\times\mathbf{r}}{r^2} = \frac{1}{c^2}\mathbf{v}\times\mathbf{E}, \tag{26.9}$$

The electron experiences a magnetic field related to the orbital angular momentum $\mathbf{L}$

$$\mathbf{B} = -\frac{1}{m_e c^2}(\mathbf{v}\times\boldsymbol{\nabla}\phi) = -\frac{1}{m_e c^2}\frac{\partial\phi}{\partial r}\left(\mathbf{v}\times\frac{\mathbf{r}}{r}\right) = \frac{1}{(m_e c)^2}\frac{1}{r}\frac{\partial\phi}{\partial r}\mathbf{L}, \tag{26.10}$$

where we used the orbital angular momentum is $\mathbf{L} = \mathbf{r}\times\mathbf{p} = -\mathbf{p}\times\mathbf{r} = \mathbf{r}\times(m_e\mathbf{v}_p)$. A magnetic dipole $\boldsymbol{\mu}$ in a magnetic field has potential energy $-\boldsymbol{\mu}\cdot\mathbf{B}$. In 1926, Llewellyn Thomas correctly derived a relativistic effect called Thomas precession introduces an extra factor of $1/2$ that we quote from textbooks. The resulting Hamiltonian is

$$H = -\frac{1}{2}\boldsymbol{\mu}\cdot\mathbf{B}. \tag{26.11}$$

Applying the electron magnetic moment $\hat{\boldsymbol{\mu}} = -g_s e/2m_e\hat{\mathbf{S}}$ and putting hats to denote operators, the Hamiltonian for spin-orbit coupling is

$$\Delta\hat{H}_{\text{spin-orbit}} = -\frac{e^2}{4\pi\varepsilon_0}\frac{1}{2(m_e c)^2}\frac{1}{r^3}\hat{\mathbf{L}}\cdot\hat{\mathbf{S}}. \tag{26.12}$$

To evaluate $\hat{\mathbf{L}}\cdot\hat{\mathbf{S}}$, the trick is to introduce the total angular momentum $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$ and square

$$\hat{J}^2 = \hat{L}^2 + \hat{S}^2 + 2\hat{\mathbf{L}}\cdot\hat{\mathbf{S}} \quad\Rightarrow\quad \hat{\mathbf{L}}\cdot\hat{\mathbf{S}} = \frac{1}{2}\left(\hat{J}^2 - \hat{L}^2 - \hat{S}^2\right). \tag{26.13}$$

The $\hat{J}^2$ operator commutes with its azimuthal component $\hat{J}_z$ and also the $\hat{L}^2$, $\hat{S}^2$ operators

$$[\hat{J}_z, \hat{J}^2] = [\hat{L}^2, \hat{J}^2] = [\hat{S}^2, \hat{J}^2] = 0. \tag{26.14}$$

However, the components of the original angular momenta do not commute with $\hat{J}^2$

$$[\hat{L}_z, \hat{J}^2] \neq 0, \quad [\hat{S}_z, \hat{J}^2] \neq 0. \tag{26.15}$$

So we can only simultaneously measure the following quantum numbers (but not m_ℓ, m_s)

$$|n, j, m_j, \ell, s\rangle. \tag{26.16}$$

The m_j quantum numbers are constrained as usual $m_j = -j, -j+1, \ldots, j-1, j$. The minimum and maximum values of j are:

$$j_{\min} = |\ell - s|, \quad j_{\max} = \ell + s. \tag{26.17}$$

We are only concerned with an electron spin $s = 1/2$, which keeps things simpler[7]:

$$j_+ = \ell + \frac{1}{2}, \quad j_- = \ell - \frac{1}{2} \text{ for } \ell \neq 0. \tag{26.18}$$

Applying equation (26.13), the spin-orbit coupling $\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$ operator has eigenvalue equation

$$\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} |n, j, m_j, \ell, s\rangle = \frac{\hbar^2}{2} \left[j(j+1) - \ell(\ell+1) - s(s+1) \right] |n, j, m_j, \ell, s\rangle = \frac{\hbar^2}{2} \lambda |n, j, m_j, \ell, s\rangle. \tag{26.19}$$

Fixing $s = 1/2$, we then evaluate the dimensionless eigenvalue $\lambda_\pm$ to be

$$\lambda_\pm = \begin{cases} -(\ell+1), & j = \ell + \frac{1}{2}, \\ \ell, & j = \ell - \frac{1}{2} \quad (\ell \neq 0). \end{cases} \tag{26.20}$$

The energy shift for degenerate perturbation theory is

$$\Delta E_{LS}^{n,j,\ell} = \langle \Delta \hat{H}_{\text{spin-orbit}} \rangle_{n,j,\ell} = \frac{e^2 \hbar^2}{4\pi\varepsilon_0 2 (m_e c)^2} \lambda_\pm \left\langle \frac{1}{r^3} \right\rangle_{n,j,\ell}. \tag{26.21}$$

For brevity, we quote the expectation value from textbooks without proof to be

$$\left\langle \frac{1}{r^3} \right\rangle_{n,j,\ell} = \frac{1}{(r_B)^3} \frac{1}{n^3} \frac{1}{\ell \left(\ell + \frac{1}{2}\right) (\ell + 1)}. \tag{26.22}$$

Therefore, the energy shift from spin-orbit coupling is

$$\Delta E_{LS}^{n,j,\ell} = \langle \Delta \hat{H}_{\text{spin-orbit}} \rangle_{n,j,\ell} = \frac{\alpha^4 m_e c^2}{4} \frac{1}{n^3} \frac{1}{\ell \left(\ell + \frac{1}{2}\right) (\ell + 1)} \lambda_\pm. \tag{26.23}$$

We can rewrite this with $j = \ell \pm \frac{1}{2}$ as

$$\Delta E_{LS}^{n,j,\ell} = \frac{1}{4} \left(\frac{\alpha}{n} \right)^2 m_e c^2 \frac{n}{j + 1/2} \times \begin{cases} \frac{1}{j}, \\ -\frac{1}{j+1}. \end{cases} \tag{26.24}$$

[7] In general, there are $2N+1$ states, where $N = s$ if $s < \ell$ or $N = \ell$ if $\ell < s$. The spectrum of j values increments by unity according to ladder operators $j = |\ell - s|, |\ell - s| + 1, \ldots, (\ell + s) - 1, (\ell + s)$.

There is one last relativistic correction called the Darwin term related to the electron rest mass. Full details require the Dirac equation, where the physical cartoon is that the electron undergoes *Zitterbewegung*, or "trembling motion", and sees a smeared potential. We simply quote the result that is also $\sim \alpha^4$

$$\langle \hat{H}_{\text{Darwin}} \rangle = \frac{1}{2} m_e c^2 \left(\frac{\alpha}{n}\right)^4 n \delta_{\ell,0}. \tag{26.25}$$

We combine all three perturbations into $\hat{H}_{\text{fine-structure}} = \hat{H}_{\text{rel. kin.}} + \hat{H}_{\text{spin-orbit}} + \hat{H}_{\text{Darwin}}$ and after all the dust settles we find:

$$\Delta E_{\text{fine-structure}} = \langle \hat{H}_{\text{fine-structure}} \rangle = \frac{1}{2} m_e c^2 \left(\frac{\alpha}{n}\right)^4 \left(\frac{3}{4} - \frac{n}{j+1/2}\right). \tag{26.26}$$

The impact of all the terms in fine structure is to lift the degeneracy of states for orbital angular momentum.

26.3 Discussion of higher-order corrections

Fine structure is not the end of the story. We can keep going in perturbation theory with higher-order corrections, sometimes known as *hyperfine structure*

$$\hat{H} = \hat{H}_{\text{Coulomb}} + \underbrace{\hat{H}_{\text{rel. kin.}} + \hat{H}_{\text{spin-orbit}} + \hat{H}_{\text{Darwin}}}_{\hat{H}_{\text{fine-structure}}} + \hat{H}_{\text{nuclear}} + \hat{H}_{\text{Lamb}}. \tag{26.27}$$

Let us qualitatively discuss these different terms:

- $\hat{H}_{\text{Coulomb}}$, $E \sim \alpha^2$: This governs the gross structure Hamiltonian with the static Coulomb potential. It results in the Bohr energy levels that depend quadratically on the fine-structure constant $E_n = -\frac{1}{2}\alpha^2 m_e c^2/n^2$.

- $\hat{H}_{\text{fine-structure}}$, $\Delta E \sim \alpha^4$. This contains all the relativistic kinetic motion, spin-orbit coupling, and Darwin rest-mass terms. These are accounted for using the Dirac equation, the relativistic counterpart of the Schrödinger equation with a Coulomb potential. The net effect is to partly lift the angular momentum degeneracy. It is proportional to the fourth power in fine-structure constant: $\frac{1}{2} m_e c^2 \left(\frac{\alpha}{n}\right)^4 \left(\frac{3}{4} - \frac{n}{j+1/2}\right)$.

- $\hat{H}_{\text{Lamb}}$, $\Delta E \sim \alpha^5$. In 1947 at Columbia University, an even smaller correction was measured by Lamb and Retherford [55]. This contribution splits the $2P_{1/2}$ and $2S_{1/2}$ states that degenerate in fine structure (spectroscopic notation nL_J) by 4.3×10^{-6} eV. The Lamb shift and the electron anomalous magnetic moment (section 21) was the

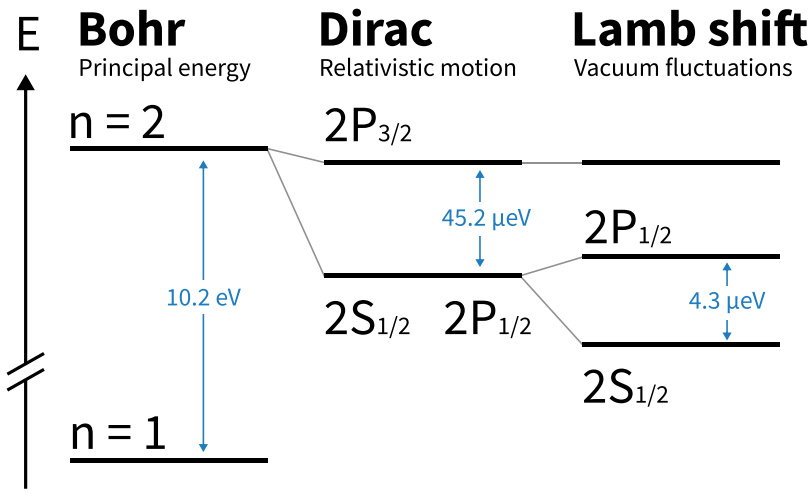


Figure 40: Sketch of Lamb shift.

topic of the 1947 Shelter Island conference off the eastern shores of Long Island. The Lamb shift is the groundbreaking discovery that the vacuum is not static but rather has quantum fluctuations. The result paved the way to quantum electrodynamics, which is covered in graduate courses on quantum field theory.

- $\hat{H}_{\text{nuclear}}$. There are also small corrections from more accurate modelling of the atomic nucleus. First, nuclei can have sizeable magnetic moments $\hat{\boldsymbol{\mu}}_N = g_N(e/2m_N)\hat{\mathbf{I}}$ from section 21.2, which contributes to electron-nucleus spin-spin interactions $\hat{\mathbf{I}}\cdot\hat{\mathbf{S}}$. Second, the nucleus is not a point charge. The so-called isotope shift by accounting for finite-volume effects of the atomic nucleus was derived by NYU faculty Gregory Breit and graduate student Jenny Rosenthal[8] in 1932 [56]. Working in the University Heights campus that is now Bronx Community College, this work provided an early way to determine the finite size of the nucleus via precision spectroscopy.

[8]NYU alumna profile: https://physics.nyu.edu/alumni/Doctoral/bramley.jenny.html

27 Helium

At first glance, helium is "just" hydrogen with an extra electron. We can add a hydrogen Hamiltonian with charge Ze for each electron and an additional term to model the repulsive Coulomb force of the two electrons

$$\hat{H} = \frac{\hat{\mathbf{p}}_1^2 + \hat{\mathbf{p}}_2^2}{2m_e} - \frac{Ze^2}{4\pi\varepsilon_0}\left(\frac{1}{|\hat{\mathbf{r}}_1|} + \frac{1}{|\hat{\mathbf{r}}_2|}\right) + \frac{e^2}{4\pi\varepsilon_0|\hat{\mathbf{r}}_{12}|}. \tag{27.1}$$

We describe the following terms:

- The first group of operators represent the kinetic energies of the electrons.
- The second group of operators is the electron-nucleus (attractive) Coulomb interaction due to nuclear charge $Z = 2$ for helium.
- The final operator is the electron-electron (repulsive) Coulomb interaction, where $\hat{\mathbf{r}}_{12}$ is the relative co-ordinate of the two electrons $\hat{\mathbf{r}}_{12} = \hat{\mathbf{r}}_1 - \hat{\mathbf{r}}_2$.

Unfortunately, this is the three-body problem of quantum mechanics. It is intractable to solve exactly. Nonetheless, we can qualitatively extract important physics unseen in hydrogen because we have a system of two indistinguishable electrons from section 22.

27.1 Ground state

We write the ground $|G\rangle$ and first excited $|F\rangle$ spatial states of an electron respectively as

$$|1,0,0\rangle = |G\rangle, \tag{27.2}$$
$$|2,\ell,m\rangle = |F\rangle. \tag{27.3}$$

For the helium ground state $|\psi_{100,100}\rangle$, both electrons are in the $|G\rangle$ state

$$|\psi_{100,100}\rangle = |G\rangle_1|G\rangle_2 \tag{27.4}$$

where the subscript labels $|\rangle_1$, $|\rangle_2$ represent particle 1 and particle 2, respectively. Let the eigenstate of helium be the product of the total spatial state and the total spin state

$$|\psi\rangle = |\psi_{\text{spatial}}\rangle|\psi_{\text{spin}}\rangle \tag{27.5}$$

Electrons are fermions so the combined state $|\psi\rangle$ must be antisymmetric. We can only do this via the singlet state $|A\rangle = |0,0\rangle$ from (22.27), giving the helium ground state

$$|\psi_{100,100}\rangle|A\rangle = |G\rangle_1|G\rangle_2|0,0\rangle. \tag{27.6}$$

To find the energy eigenvalues, it is tempting to treat the helium Hamiltonian (27.1) using perturbation theory. The unperturbed Hamiltonian is the sum of two hydrogenic Hamiltonians $\hat{H}_0(\hat{\mathbf{r}}_1)+\hat{H}_0(\hat{\mathbf{r}}_2)$, for which we know the solutions to, plus a perturbation term $\hat{H}_1(\hat{\mathbf{r}}_{12})$ due to the electron-electron repulsion:

$$\hat{H}_1(\hat{\mathbf{r}}_{12}) = \frac{e^2}{4\pi\varepsilon|\hat{\mathbf{r}}_{12}|}. \tag{27.7}$$

Unfortunately, the perturbation is not small compared to the size of the electron-nucleus Coulomb potential, so perturbation theory will not give very accurate results.

27.2 First excited states

To consider the first excited states of helium $|\psi_{100,2\ell m}\rangle$, we must first make the total spatial states symmetric or antisymmetric by taking linear combinations

$$|\psi^{(S)}_{100,2\ell m}\rangle = \frac{1}{\sqrt{2}}\left(|G\rangle_1|F\rangle_2+|F\rangle_1|G\rangle_2\right), \tag{27.8}$$

$$|\psi^{(A)}_{100,2\ell m}\rangle = \frac{1}{\sqrt{2}}\left(|G\rangle_1|F\rangle_2-|F\rangle_1|G\rangle_2\right). \tag{27.9}$$

We now use perturbation theory to evaluate the energy shifts corresponding to the first excited states by evaluating the expectation of $\hat{H}_1$. First use the symmetric eigenbasis to obtain

$$\begin{aligned} E^{(1)}_{100,2\ell m} &= \langle\psi^{(S)}_{100,2\ell m}|\hat{H}_1|\psi^{(S)}_{100,2\ell m}\rangle \\ &= \frac{1}{2}\int \mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2(F_1^*G_2^*+F_2^*G_1^*)H_1(F_1G_2+F_2G_1) \end{aligned} \tag{27.10}$$

When we expand out the integrand, the non-cross term give the *direct integral* J

$$J = \int \mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2(F_1^*G_2^*H_1F_1G_2) = \int \mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2|F_1|^2|G_2|^2H_1 \tag{27.11}$$

Applying exchange symmetry (by swapping particle labels) to the other non-cross term, we see it is the same as the direct integral (27.11)

$$\int \mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2(F_2^*G_1^*H_1F_2G_1) = \int \mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2(F_1^*G_2^*H_1F_1G_2) = J \tag{27.12}$$

For the cross terms in (27.10), we similarly apply exchange symmetry to obtain the *exchange integral* K

$$\int \mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2(F_1^*G_2^*H_1F_2G_1) = \int \mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2(F_2^*G_1^*H_1F_1G_2) = K \tag{27.13}$$

Together, we obtain the first order energy correction due to the repulsive electron-electron interaction in the symmetric eigenbasis

$$\langle \psi^{(S)}_{100,2\ell m} | \hat{H}_1 | \psi^{(S)}_{100,2\ell m} \rangle = J + K \tag{27.14}$$

So the symmetric spatial eigenstate $|\psi^{(S)}_{100,2\ell m}\rangle$ has energy eigenvalue

$$E^{(S)}_{100,200} = E_{n_1,n_2} + J + K. \tag{27.15}$$

where E_{n_1,n_2} is the energy eigenvalue of the Hamiltonian without the electron-electron interaction. The corresponding eigenstate is the *singlet state*, found by combining $|\psi^{(S)}_{100,200}\rangle$ with the only antisymmetric spin state (22.27)

$$|\psi^{(S)}_{100,200}\rangle |A\rangle = \frac{1}{\sqrt{2}} \left(|G\rangle_1 |F\rangle_2 + |F\rangle_1 |G\rangle_2 \right) |0,0\rangle. \tag{27.16}$$

We can similarly find the expectation value of $\hat{H}_1$ using the antisymmetric eigenstates

$$\begin{aligned} E^{(1)}_{100,2\ell m} &= \langle \psi^{(A)}_{100,2\ell m} | \hat{H}_1 | \psi^{(A)}_{100,2\ell m} \rangle = \frac{1}{2} \int \mathrm{d}^3\mathbf{r}_1 \mathrm{d}^3\mathbf{r}_2 (F_1^* G_2^* - F_2^* G_1^*) H_1 (F_1 G_2 - F_2 G_1) \\ &= J - K. \end{aligned} \tag{27.17}$$

So the energy eigenvalue for the antisymmetric spatial eigenstate $|\psi^{(A)}_{100,2\ell m}\rangle$ is different:

$$E^{(A)}_{100,21m} = E_{n_1,n_2} + J - K. \tag{27.18}$$

The resulting triplet states are found by combining $|\psi^{(A)}_{100,21m}\rangle$ with the symmetric spin states (22.28)

$$|\psi^{(A)}_{100,21m}\rangle |S\rangle = \frac{1}{\sqrt{2}} \left(|G\rangle_1 |F\rangle_2 - |F\rangle_1 |G\rangle_2 \right) \times \begin{cases} |1,1\rangle \\ |1,0\rangle \\ |1,-1\rangle \end{cases}. \tag{27.19}$$

27.3 Sudden approximation

Strictly speaking, the sudden approximation is not time-independent perturbation theory. But the perturbations are so fast that they effectively happen near-instantly such that we can neglect the time dependence. There is an application to a simple calculation of hydrogen

turning into helium nucleus as an overlap integral, which can model radioactive decay. The time scale τ for a system to evolve is from the time-dependent Schrödinger equation

$$|\psi\rangle = \sum_n |\psi(0)\rangle \exp\left(-\frac{\mathrm{i}E_n t}{\hbar}\right) \quad \Rightarrow \quad \tau \sim \hbar/E_n. \tag{27.20}$$

The sudden approximation is applicable when the time scale for a perturbation that changes much faster than τ. The general procedure:

- Before perturbation: The system is in eigenstate $|k^{(i)}\rangle$ of the Hamiltonian $\hat{H}^{(i)}$.
- A sudden perturbation occurs at $t = 0$.
- After perturbation: $|n^{(f)}\rangle$ are the eigenstates of the Hamiltonian $\hat{H}^{(f)}$.

The sudden approximation predicts that the states before and after the perturbation are equal at $t = 0$

$$|\psi^{(i)}(0)\rangle = |\psi^{(f)}(0)\rangle. \tag{27.21}$$

The state $|\psi^{(f)}\rangle$ immediately after the perturbation is a superposition of $\hat{H}^{(f)}$ eigenstates

$$|\psi^{(f)}\rangle = \sum_n |n^{(f)}\rangle\langle n^{(f)}|k^{(i)}\rangle. \tag{27.22}$$

The probability the system has transferred from the k^{th} to the n^{th} state is thus given by the modulus square of the expansion coefficients

$$P(n|k) = |\langle n^{(f)}|k^{(i)}\rangle|^2 \tag{27.23}$$

Example: tritium beta decay Tritium is an isotope of hydrogen with one proton and two neutrons ^{3}H. The ground state $|\psi_{100}^{Z=1}\rangle$ spontaneously decays to $^3\mathrm{He}^+$ via β decay

$${}^3_1\mathrm{H} \to {}^3_2\mathrm{He}^+ + e^- + \bar{\nu}_e. \tag{27.24}$$

Microscopically, a neutron n turns into a proton p while emitting an electron e^- and anti-neutrino $\bar{\nu}_e$. The probability P that $^3\mathrm{He}^+$ is also in the ground state $|\psi_{100}^{Z=2}\rangle$ is given by (27.23):

$$\begin{aligned} P &= |\langle \psi_{100}^{Z=2}|\psi_{100}^{Z=1}\rangle|^2 \\ &= 4\pi \int_0^\infty \sqrt{\frac{2^3}{\pi a_0^3}} \mathrm{e}^{-2r/a_0} \sqrt{\frac{1}{\pi a_0^3}} \mathrm{e}^{-r/a_0} r^2 \mathrm{d}r \\ &= 2^3 \left(\frac{2}{3}\right)^6 \approx 70\%. \end{aligned} \tag{27.25}$$

28 Review of course

The last lecture provides a comprehensive review of the course.

28.1 Foundations and formalism

Experimental foundations There are several important experiments that demonstrate quantum behaviour

- Wave-particle duality. Davisson-Germer experiment showed electrons behaved as waves. Compton scattering experiment showed photons behaved as particles.
- Atomic radiation. Frank-Hertz experiment showed discrete atomic energy states. Radioactivity and the gold-foil alpha scattering experiments showed the existence of an atomic nucleus. Together, these results motivate quantization and the Bohr model that successfully accounts for the Rydberg formula for line spectra.
- Radioactive decay and energy transitions are probabilistic, motivating a reformulation of mechanics using the wavefunction and probability amplitudes.

Operator formalism We can restate what we have learnt as the key postulates of quantum mechanics:

1. **Wavefunction postulate**. Quantum states are described by a wavefunction state vector $|\psi\rangle$ that is square integrable for normalizability.
2. **Operator postulate**. Observables are eigenvalues q_n of Hermitian operators $\hat{Q}$ according to the eigenvalue equation $\hat{Q}|\phi_n\rangle = q_n|\phi_n\rangle$ for eigenstate $|\phi_n\rangle$.
3. **Measurement postulate**. Given an initial state $|\psi\rangle$ can be expanded in terms of an eigenbasis $\sum_n q_n|\phi_n\rangle$, the probability amplitude that the state is in $|\phi_n\rangle$ upon measurement is given by the Born rule
$$P(\phi_n|\psi) = |\langle\phi_n|\psi\rangle|^2. \tag{28.1}$$
4. **Time evolution postulate**: the wavefunction evolves according to Schrödinger's equation.
$$\mathrm{i}\hbar\frac{\partial|\psi\rangle}{\partial t} = \hat{H}|\psi\rangle. \tag{28.2}$$

The standard deviations for measurements of two operators $\hat{A}$ and $\hat{B}$ satisfy this generalized uncertainty relation

$$\sigma_A \sigma_B \geq \left| \frac{\mathrm{i}}{2} \langle [\hat{A}, \hat{B}] \rangle \right|. \tag{28.3}$$

Taking the canonical commutation relation $[\hat{x}, \hat{p}] = \mathrm{i}\hbar$, this becomes the Heisenberg uncertainty principle

$$\sigma_x \sigma_p \geq \frac{\hbar}{2}. \tag{28.4}$$

28.2 Wave mechanics

General prescription The time-independent Schrödinger equation with potential $V(x)$ is

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + k^2\psi = 0, \quad k^2 = \frac{2m(E-V)}{\hbar^2}. \tag{28.5}$$

There are two classes of solutions:

- **Scattering states**. Particles with energy $E > 0$ have plane-wave solutions for real k

$$\psi(x) = A\mathrm{e}^{\mathrm{i}kx} + B\mathrm{e}^{-\mathrm{i}kx}, \qquad E = \frac{\hbar^2 k^2}{2m}. \tag{28.6}$$

- **Bound states**. Particles with energy $E < 0$ have with real κ with exponential solutions

$$\psi(x) = A\mathrm{e}^{\kappa x} + B\mathrm{e}^{-\kappa x}, \qquad E = -\frac{\hbar^2 \kappa^2}{2m} < 0. \tag{28.7}$$

Boundary conditions fix the integration constants A and B. We obtain boundary conditions by integrating the Schrödinger equation over a small interval $[a-\varepsilon, a+\varepsilon]$ at $x = a$ across a change in potential

$$\int_{a-\varepsilon}^{a+\varepsilon} \frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}\mathrm{d}x = -\frac{2m}{\hbar^2}\int_{a-\varepsilon}^{a+\varepsilon} [E - V(x)]\psi(x)\mathrm{d}x. \tag{28.8}$$

The wavefunction is continuous across a boundary at $x = a$ between two potentials

$$\psi_1(a) = \psi_2(a), \tag{28.9}$$

for ψ_1 in $x < a$ and ψ_2 in $x > a$. The wavefunction's first derivative at $x = a$ is continuous

$$\left.\frac{\mathrm{d}\psi_1}{\mathrm{d}x}\right|_a = \left.\frac{\mathrm{d}\psi_2}{\mathrm{d}x}\right|_a. \tag{28.10}$$

Harmonic oscillator The Hamiltonian for the harmonic oscillator is

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2. \tag{28.11}$$

We factorize the Hamiltonian into a pair of ladder operators

$$\hat{A} = \frac{1}{\sqrt{2m\hbar\omega}}(m\omega x + \mathrm{i}\hat{p}), \quad \hat{A}^\dagger = \frac{1}{\sqrt{2m\hbar\omega}}(m\omega x - \mathrm{i}\hat{p}). \tag{28.12}$$

They satisfy the commutation relation $[\hat{A},\hat{A}^\dagger] = 1$. We find $\hat{A}^\dagger$ raises and $\hat{A}$ lowers energy eigenstates by $\hbar\omega$

$$\hat{H}|\psi_\pm\rangle = (E_n \pm \hbar\omega)|\psi_\pm\rangle, \quad |\psi_+\rangle = \hat{A}^\dagger|E_n\rangle, \quad |\psi_-\rangle = \hat{A}|E_n\rangle. \tag{28.13}$$

28.3 Angular momentum

The angular momentum operator is defined $\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$. The total angular momentum operator $\hat{L}^2$ and azimuthal projection $\hat{L}_z$. This applies for orbital $\hat{L}$, labelled by ℓ, m_ℓ:

$$\hat{L}^2|\ell, m_\ell\rangle = \hbar^2\ell(\ell+1)|\ell, m_\ell\rangle, \tag{28.14}$$

$$\hat{L}_z|\ell, m_\ell\rangle = \hbar m_\ell|\ell, m_\ell\rangle. \tag{28.15}$$

Angular momentum operators satisfy the following commutation relations

$$[\hat{L}_i, \hat{L}_j] = \mathrm{i}\hbar\varepsilon_{ijk}\hat{L}_k, \quad [\hat{L}^2, \hat{L}_i] = 0. \tag{28.16}$$

We can replace all instances of $\hat{L}$ with the spin $\hat{S}$ operators and the eigenvalues are labelled s, m_s. The spin operator is given by

$$\hat{\mathbf{S}} = \frac{\hbar}{2}\hat{\boldsymbol{\sigma}}, \tag{28.17}$$

where $\hat{\boldsymbol{\sigma}}$ is a vector of three Pauli matrices

$$\hat{\boldsymbol{\sigma}} = \begin{pmatrix} \hat{\sigma}_x \\ \hat{\sigma}_y \\ \hat{\sigma}_z \end{pmatrix}, \quad \text{where} \quad \hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{28.18}$$

These obey the commutation relation

$$[\hat{\sigma}_i, \hat{\sigma}_j] = 2\mathrm{i}\varepsilon_{ijk}\hat{\sigma}_k. \tag{28.19}$$

28.4 Atoms

Hydrogen gross structure is the solution of the time-independent Schrödinger equation in spherical coordinates with a Coulomb potential

$$\left[-\frac{\hbar^2}{2m_e}\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}}{\mathrm{d}r}\right)+\frac{\ell(\ell+1)\hbar^2}{2m_e r^2}-\frac{e^2}{4\pi\varepsilon_0 r}\right]\psi(r,\theta,\phi)=E\,\psi(r,\theta,\phi). \tag{28.20}$$

A state is denoted by three $|n,\ell,m\rangle$ quantum numbers:

- Principal quantum number n in Bohr energy formula

$$E_n=-\frac{1}{2}\alpha^2 m_e c^2\frac{1}{n^2}=-\frac{E_R}{n^2}\simeq-\frac{13.6\text{ eV}}{n^2},\quad n=1,2,3,\ldots. \tag{28.21}$$

- Angular momentum quantum number ℓ for $\hat{L}^2$ operator: $\ell=0,1,2,\ldots,n-1$.
- The m_ℓ magnetic quantum number for $\hat{L}_z$ operator $m_\ell=-\ell,-\ell+1,\ldots,0,\ldots,\ell-1,\ell$. Given a value of ℓ we have $2\ell+1$ states specified by m_ℓ.

References

[1] C. Davisson and L. H. Germer, *The Scattering of Electrons by a Single Crystal of Nickel*, *Nature* **119** (1927) 890.

[2] C. Davisson and L. H. Germer, *Diffraction of Electrons by a Crystal of Nickel*, *Phys. Rev.* **30** (1927) 705.

[3] L. de Broglie, *Recherches sur la théorie des Quanta*, *Annals Phys.* **2** (1925) 22.

[4] C. T. R. Wilson, *On an expansion apparatus for making visible the tracks of ionising particles in gases and some results obtained by its use*, *Proc. R. Soc. Lond. A* **87** (1912) 277.

[5] R. Brown, U. Camerini, P. H. Fowler, H. Muirhead, C. F. Powell and D. Ritson, *Observations With Electron Sensitive Plates Exposed to Cosmic Radiation*, *Nature* **163** (1949) 47.

[6] G. Thomson and A. Reid, *Diffraction of Cathode Rays by a Thin Film*, *Nature* **119** (1927) 890.

[7] J. J. Thomson, *Lviii. on the masses of the ions in gases at low pressures*, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **48** (1899) 547.

[8] C. M. G. Lattes, H. Muirhead, G. P. S. Occhialini and C. F. Powell, *Processes Involving Charged Mesons*, *Nature* **159** (1947) 694.

[9] T. Young, *II. The Bakerian Lecture. On the theory of light and colours*, *Philosophical Transactions of the Royal Society of London* (1802) 12.

[10] A. Einstein, *Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt*, *Annalen der Physik* **322** (1905) 132.

[11] A. H. Compton, *A Quantum Theory of the Scattering of X-rays by Light Elements*, *Phys. Rev.* **21** (1923) 483.

[12] A. H. Compton, *The Spectrum of Scattered X-Rays*, *Phys. Rev.* **22** (1923) 409.

[13] C. H. Payne, *Stellar Atmospheres; a Contribution to the Observational Study of High Temperature in the Reversing Layers of Stars*, *Harvard Observatory Monographs* (1925) .

[14] H. Geiger and E. Marsden, *Lxi. the laws of deflexion of a particles through large angles*, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **25** (1913) 604.

[15] E. Rutherford, *Lxxix. the scattering of α and β particles by matter and the structure of the atom*, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **21** (1911) 669.

[16] J. Franck and G. Hertz, *Über zusammenstöße zwischen elektronen und den molekülen des quecksilberdampfes und die ionisierungsspannung desselben*, *Physikalische Blätter* **23** (1967) 294.

[17] S. B. McKagan, K. K. Perkins and C. E. Wieman, *Why we should teach the Bohr model and how to teach it effectively*, *Phys. Rev. ST Phys. Educ. Res.* **4** (2008) 010103.

[18] D. J. Fixsen, E. S. Cheng, J. M. Gales, J. C. Mather, R. A. Shafer and E. L. Wright, *The Cosmic Microwave Background spectrum from the full COBE FIRAS data set*, *Astrophys. J.* **473** (1996) 576 [astro-ph/9605054].

[19] M. Born, *Zur Quantenmechanik der Stoßvorgänge*, *Z. Phys.* **37** (1926) 863.

[20] C. Jönsson, *Elektroneninterferenzen an mehreren künstlich hergestellten Feinspalten*, *Zeitschrift für Physik* **161** (1961) 454.

[21] A. Tonomura, J. Endo, T. Matsuda, T. Kawasaki and H. Ezawa, *Demonstration of single-electron buildup of an interference pattern*, *Am. J. Phys.* **57** (1989) 117.

[22] R. Bach, D. Pope, S. H. Liou and H. Batelaan, *Controlled double-slit electron diffraction*, *New J. Phys.* **15** (2013) 033018 [1210.6243].

[23] S. Pedalino, B. E. Ramírez-Galindo, R. Ferstl, K. Hornberger, M. Arndt and S. Gerlich, *Probing quantum mechanics with nanoparticle matter-wave interferometry*, *Nature* **649** (2026) 033018 [2507.21211].

[24] P. Grangier, G. Roger and A. Aspect, *Experimental Evidence for a Photon Anticorrelation Effect on a Beam Splitter: A New Light on Single-Photon Interferences*, *Europhysics Letters* **1** (1986) 173.

[25] K. Harada, T. Akashi, K. Niitsu, K. Shimada, Y. A. Ono, D. Shindo et al., *Interference experiment with asymmetric double slit by using 1.2-MV field emission transmission electron microscope*, *Sci Rep.* **8** (2018) 1008.

[26] E. Schrödinger, *Über das Verhältnis der Heisenberg-Born-Jordanschen Quantenmechanik zu der meinen*, *Annalen der Physik* **384** (1926) 734.

[27] E. Schrödinger, *An Undulatory Theory of the Mechanics of Atoms and Molecules*, *Phys. Rev.* **28** (1926) 1049.

[28] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes and R. P. Hudson, *Experimental Test of Parity Conservation in β Decay*, *Phys. Rev.* **105** (1957) 1413.

[29] R. L. Garwin, L. M. Lederman and M. Weinrich, *Observations of the Failure of Conservation of Parity and Charge Conjugation in Meson Decays: the Magnetic Moment of the Free Muon*, *Phys. Rev.* **105** (1957) 1415.

[30] G. Binnig and H. Rohrer, *Scanning tunneling microscopy—from birth to adolescence*, *Rev. Mod. Phys.* **59** (1987) 615.

[31] P. A. M. Dirac, *A new notation for quantum mechanics*, *Mathematical Proceedings of the Cambridge Philosophical Society* **35** (1939) 416–418.

[32] P. Dirac, *The Principles of Quantum Mechanics*, International Series of Monographs on Physics. Clarendon Press, 1981 (First Edition 1930).

[33] W. Heisenberg, *Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik*, *Zeitschrift für Physik* **43** (1927) 172.

[34] A. Einstein, B. Podolsky and N. Rosen, *Can quantum mechanical description of physical reality be considered complete?*, *Phys. Rev.* **47** (1935) 777.

[35] J. S. Bell, *On the Einstein-Podolsky-Rosen paradox*, *Physics* **1** (1964) 195.

[36] M. Schlosshauer, J. Kofler and A. Zeilinger, *A Snapshot of Foundational Attitudes Toward Quantum Mechanics*, *Stud. Hist. Phil. Sci. B* **44** (2013) 222 [1301.1069].

[37] P. Kusch and H. M. Foley, *Precision Measurement of the Ratio of the Atomic 'g Values' in the* $^2P_{\frac{3}{2}}$ *and* $^2P_{\frac{1}{2}}$ *States of Gallium*, *Phys. Rev.* **72** (1947) 1256.

[38] P. Kusch and H. M. Foley, *The Magnetic Moment of the Electron*, *Phys. Rev.* **74** (1948) 250.

[39] J. S. Schwinger, *On Quantum-Electrodynamics and the Magnetic Moment of the Electron*, *Phys. Rev.* **73** (1948) 416.

[40] A. Lasocki and R. J. Hicks, *How we read: the combined use of MRI and novel PET tracers for the characterisation and treatment planning of masses in neuro-oncology*, *Cancer Imaging* **19** (2019) 57.

[41] P. C. Lauterbur, *Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance*, *Nature* **242** (1973) 190.

[42] I. I. Rabi, J. M. B. Kellogg and J. R. Zacharias, *The magnetic moment of the proton*, *Phys. Rev.* **46** (1934) 157.

[43] R. L. Cool, E. W. Jenkins, T. F. Kycia, D. A. Hill, L. Marshall and R. A. Schluter, *Measurement of the Magnetic Moment of the Lambda0 Hyperon*, *Phys. Rev.* **127** (1962) 2223.

[44] L. Schachinger et al., *A Precise Measurement of the* Λ^0 *Magnetic Moment*, *Phys. Rev. Lett.* **41** (1978) 1348.

[45] A. Ekert, T. Hosgood, A. Kay and C. Macchiavello, "Introduction to Quantum Information Science." https://qubit.guide, 2025.

[46] J. F. Clauser, M. A. Horne, A. Shimony and R. A. Holt, *Proposed Experiment to Test Local Hidden-Variable Theories*, *Phys. Rev. Lett.* **23** (1969) 880.

[47] C. S. Wu and I. Shaknov, *The Angular Correlation of Scattered Annihilation Radiation*, *Phys. Rev.* **77** (1950) 136.

[48] A. Aspect, P. Grangier and G. Roger, *Experimental Realization of*

Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities, *Phys. Rev. Lett.* **49** (1982) 91.

[49] M. G. Mayer, *On closed shells in nuclei. 2*, *Phys. Rev.* **75** (1949) 1969.

[50] O. Haxel, J. H. D. Jensen and H. E. Suess, *On the "Magic Numbers" in Nuclear Structure*, *Phys. Rev.* **75** (1949) 1766.

[51] G. E. Hale, F. Ellerman, S. B. Nicholson and A. H. Joy, *The Magnetic Polarity of Sun-Spots*, *The Astrophysical Journal* **49** (1919) 153.

[52] J. S. Foster and M. L. Chalk, *Observed Relative Intensities of Stark Components in Hydrogen*, *Nature* **118** (1926) 592.

[53] J. S. Foster and M. L. Chalk, *Observed Relative Intensities of Stark Components of Hα*, *Nature* **121** (1928) 830–831.

[54] J. M. Dewey, *Intensities in the Stark Effect of Helium*, *Phys. Rev.* **28** (1926) 1108.

[55] W. E. Lamb and R. C. Retherford, *Fine Structure of the Hydrogen Atom by a Microwave Method*, *Phys. Rev.* **72** (1947) 241.

[56] J. E. Rosenthal and G. Breit, *The Isotope Shift in Hyperfine Structure*, *Phys. Rev.* **41** (1932) 459.