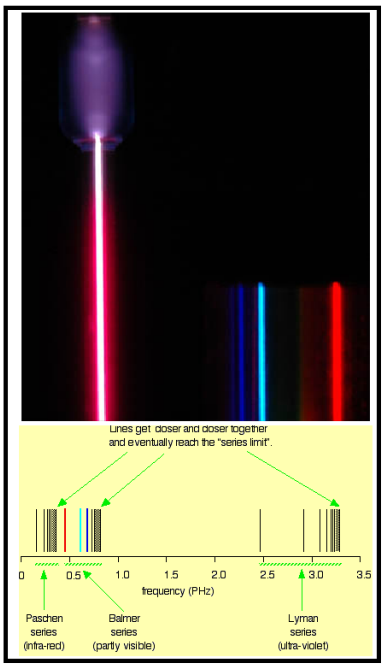
University College Dublin
An Coláiste Ollscoile, Baile Átha Cliath

School of Mathematical Sciences
Scoil na nEolaíochtaí Matamaitice
Foundations of Quantum Mechanics (ACM30210)



Dr Lennon Ó Náraigh

Lecture notes in Quantum Mechanics, January 2015

Foundations of Quantum Mechanics (ACM30210)

- Subject: Applied and Computational Maths
- School: Mathematical Sciences
- Module coordinator: Dr Lennon Ó Náraigh
- Credits: 5
- Level: 3
- Semester: Second

This module introduces Quantum Mechanics in its modern mathematical setting. Several canonical, exactly-solvable models are studied, including one-dimensional piecewise constant potentials, Dirac potentials, the harmonic oscillator, and the Hydrogen atom. Three calculational techniques are introduced: time-independent perturbation theory, variational methods, and numerical (spectral) methods.

The postulates of Quantum Mechanics, **Mathematical background** Complex vector spaces and scalar products, linear forms and duality, the natural scalar product derived from linear forms, Hilbert spaces, linear operators, commutation relations, expectation values, uncertainty, **Time evolution and the Schrödinger equation** Derivation of the Schrodinger equation for time-independent Hamiltonians, the position and momentum representations, the probability current, the free particle **Piecewise constant one-dimensional potentials** Bound and unbound states, wells and barriers, scattering, transmission coefficients, tunneling, **The harmonic oscillator** Solution by power series, Hermite polynomials, creation and annihilation operators, coherent states, **The Hydrogen atom** Solution by separation of variables, quantization of energy and angular momentum, general treatment of central potentials in terms of spherical harmonics, **Angular momentum** Motivation: angular momentum in the hydrogen atom, as derived from spherical harmonics, angular momentum in the abstract setting, intrinsic angular momentum, addition of angular momenta, Clebsch-Gordan coefficients, **Approximation methods** Time-independent perturbation theory: the non-degenerate case, variational methods for estimating the ground-state energy Further topics may include: Spin coherent states, how to build a microwave laser, the Dyson series for time-evolution for time-dependent Hamiltonians, one-dimensional Dirac potentials, time-independent perturbation theory for degenerate eigenstates, the fine structure of Hydrogen, numerical (spectral) methods for solving the Schrödinger equation

What will I learn?

On completion of this module students should be able to

1. Perform standard linear-algebra calculations as they relate to the mathematical foundations of Quantum Mechanics;
2. Solve standard problems for systems with finite-dimensional Hilbert spaces, e.g. the two-level system
3. Solve standard one-dimensional models including piecewise constant potential wells and barriers, Dirac potentials, and the Harmonic oscillator;
4. Perform calculations based on Hermite polynomials, including the characterization of coherent states;
5. Compute expectation values for appropriate observables for the Hydrogen atom;
6. Explain the quantum theory of angular momentum and compute expectation values for appropriate observables. These computations will involve both the matrix representation of intrinsic angular momentum, and the spherical-harmonic representation of orbital angular momentum;
7. Add independent angular momenta in the quantum-mechanical fashion;
8. Perform time-independent non-degenerate perturbation theory up to and including the second order

First edition: January 2011
Second edition: January 2012
Third edition: January 2013
Fourth edition: January 2014
This edition: January 2015

Contents

Chapter 1

Introduction

1.1 Overview

Here is the executive summary of the course:

> After this course, you will understand the vital need for quantum mechanics in reconciling experiments with theories of particle behaviour. You will be able to apply the mathematical machinery of quantum mechanics to characterize several physical systems of great practical importance: the square well, the harmonic oscillator, and the hydrogen atom. In doing so, you will develop intuition about quantum tunnelling, angular momentum, uncertainty, and the theory of operators.

In more detail, we will follow the following programme of work:

1. We review the evidence that points to the failure of classical mechanics and introduce an alternative treatment;
2. We study the theory of linear operators on Hilbert spaces;
3. We formulate the postulates of quantum mechanics;
4. We examine the Schrödinger equation;
5. We apply the Schrödinger equation to several standard systems;
6. We introduce perturbation theory to solve non-standard problems where a certain parameter is small;
7. We introduce spectral methods to study problems that do not have an analytical solution;
8. We introduce variational methods for the same reason.

1.2 Learning and Assessment

Learning:

- Thirty six classes, three per week.
- In some classes, we will solve problems together or look at supplementary topics.
- To develop an ability to *solve problems autonomously*, you will be given homework exercises, and it is recommended that you do *independent study*.

Assessment:

- Three homework assignments, for a total of 20%;
- Three in-class tests, for a total of 20%;
- One end-of-semester exam, 60%

Policy on late submission of homework:

The official UCD policy explained in the Science handbook will be strictly adhered to: coursework that is late by up to one week after the due date will have the grade awarded reduced by two grade points (e.g. from B- to C); coursework submitted up to two weeks after the due date will have the grade reduced by four grade points (e.g. B- to D+). Coursework received more than two weeks after the due date may not be accepted.

Textbooks

- Lecture notes will be put on the web. These are self-contained. They will be available *before* class. It is anticipated that you will print them and bring them with you to class. You can then annotate them and follow the proofs and calculations done on the board. Thus, you are still expected to attend class, and I will occasionally deviate from the content of the notes, give hints about solving the homework problems, or give a revision tips for the final exam.
- To a certain extent, I have based my notes on the book by Mandl:
  - *Quantum Mechanics*, F. Mandl, Wiley (Four copies in UCD library, 530.12).
- I have also used material from the following sources:
  - *University Physics*, H. D. Young and R. A. Freedman, Addison–Wesley (10^{th} edition, 2000);

  - *The Feynman Lectures on Physics*, R. P. Feynman, Addison–Wesley–Longman (1st edition, 1970);
  - *Quantum Mechanics Non-Relativistic Theory*, L. D. Landau and L. M. Lifshitz, Butterworth–Heinemann (3rd edition, 1981).

- The lecture notes by Prof. David Simms (Course 211) will be helpful in understanding the mathematical formulation of Quantum Mechanics.

1.3 On the failures of classical mechanics

Reading material for this chapter: Young and Freedman, Chapters 40–41

In other classes (e.g. ACM/MAPH 10030) you will have learned that the equation

$$m\frac{d^2\boldsymbol{x}}{dt^2} = -\nabla\mathcal{U},$$

is sufficient to describe the trajectory $\boldsymbol{x}(t)$ of a particle of mass m, for all time. In other classes (e.g. ACM 40010) you will have learned that the equations (SI units)

$$\nabla\cdot\boldsymbol{E} = \frac{1}{\epsilon_0}\rho, \tag{1.1}$$

$$\nabla\cdot\boldsymbol{B} = 0, \tag{1.2}$$

$$\nabla\times\boldsymbol{E} = -\frac{\partial\boldsymbol{B}}{\partial t}, \tag{1.3}$$

$$\nabla\times\boldsymbol{B} = \mu_0\boldsymbol{J} + \mu_0\epsilon_0\frac{\partial\boldsymbol{E}}{\partial t}, \tag{1.4}$$

suffice to describe electromagnetic phenomena. We now examine what happens when these two sets of equations are combined.

1.3.1 Blackbody radiation

Rayleigh, c. 1890

A *blackbody* is a perfect emitter (and absorber) of electromagnetic radiation, and is in thermal equilibrium. Such a body can be modelled as a box (or cavity) containing normal modes of electromagnetic radiation. To see what such normal modes look like, we take the curl of Eq. (1.3) and combine it with Eq. (1.4). There are no sources and sinks of radiation in the box, hence $\boldsymbol{J} = \rho = 0$, and

$$\nabla\times(\nabla\times\boldsymbol{E}) = -\frac{\partial}{\partial t}(\nabla\times\boldsymbol{B}) = -\mu_0\epsilon_0\frac{\partial^2\boldsymbol{E}}{\partial t^2}.$$

We use the vector-calculus identity (ACM 20150)

$$\nabla \times (\nabla \times \boldsymbol{E}) = \nabla (\nabla \cdot \boldsymbol{E}) - \nabla^2 \boldsymbol{E}, \qquad \nabla \cdot \boldsymbol{E} = 0,$$

hence

$$\frac{1}{c^2}\frac{\partial^2 \boldsymbol{E}}{\partial t^2} = \nabla^2 \boldsymbol{E}, \qquad c = \frac{1}{\sqrt{\mu_0 \epsilon_0}}. \tag{1.5}$$

The solution to this **wave equation** is

$$\boldsymbol{E} = \boldsymbol{E}_0 \mathrm{e}^{\mathrm{i}\omega t} \sin(k_x x) \sin(k_y y) \sin(k_z z),$$

with **dispersion relation**

$$\frac{\omega^2}{c^2} = \boldsymbol{k}^2, \qquad \boldsymbol{k} = (k_x, k_y, k_z).$$

For reasons that will become clear in what follows, we label the solution by the wavenumber $\boldsymbol{k}$:

$$\boldsymbol{E}_{\boldsymbol{k}} = \boldsymbol{E}_{0\boldsymbol{k}} \mathrm{e}^{\mathrm{i}\omega t} \sin(k_x x) \sin(k_y y) \sin(k_z z). \tag{1.6}$$

Now the domain of the problem is a box, $\boldsymbol{x} \in [0, L]^3$, and the **boundary conditions** on the box wall specify that no vibrations can occur there (ACM 30220), hence $\boldsymbol{E}(x = 0) = \boldsymbol{E}(x = L) = 0$ etc., or

$$k_x = \frac{n_x \pi}{L}, \qquad k_y = \frac{n_y \pi}{L}, \qquad k_z = \frac{n_z \pi}{L}, \qquad n_x, n_y, n_z \in \mathbb{N}$$

(that is why no cosines appear in the solution; they cannot satisfy the boundary conditions). Going back to the dispersion relation, we have

$$\begin{aligned} \frac{\omega^2}{c^2} &= \boldsymbol{k}^2, \\ &= \frac{\pi^2}{L^2}\left(n_x^2 + n_y^2 + n_z^2\right), \\ &= \left(\frac{2\pi}{\lambda}\right)^2, \end{aligned}$$

hence

$$\lambda = \frac{2L}{\sqrt{n_x^2 + n_y^2 + n_z^2}}, \qquad (n_x, n_y, n_z) \in \mathbb{N}^3.$$

An allowed $\boldsymbol{k}$-value is called a **normal mode**. To each normal mode, there corresponds a wavelength λ. Note, however, that different integer triples can produce the same wavelength. We wish to compute the total energy in the cavity. To do so, we will resort to **density-of-modes** calculations.

First, note that

$$\begin{aligned} u &= \frac{\text{Total energy}}{\text{Unit volume}}, \\ &= \int_0^\infty \frac{\text{Total energy in a wavelength interval from } \lambda \text{ to } \mathrm{d}\lambda}{\text{Unit volume, unit wavelength}} \mathrm{d}\lambda, \\ &= \int_0^\infty u_\lambda(\lambda) \mathrm{d}\lambda. \end{aligned}$$

The function $u_\lambda(\lambda)$ is the **spectral density**, which we compute now.

The number of normal modes with wavelength λ is obtained by counting points in $\boldsymbol{k}$-space. This is a three-dimensional discrete space where (k_x, k_y, k_z) form the axes, and where each allowed point (k_x, k_y, k_z) is given by $(n_x, n_y, n_z)\pi/L$, where $(n_x, n_y, n_z) \in \mathbb{N}^3$. There is one such point in a box of volume $(\pi/L)^3$ in this space[1]:

$$\text{Number of points per unit volume in } \boldsymbol{k}\text{-space} = \frac{1}{\text{Box volume}} = \frac{L^3}{\pi^3}.$$

The number of normal modes of **magnitude** $k = \sqrt{k_x^2 + k_y^2 + k_z^2}$, in the range $[k, k + \mathrm{d}k]$, is given by

$$\begin{aligned} &\text{Number of normal modes in the range } [k, k + \mathrm{d}k] = \\ &\qquad [\text{Number of points per unit volume in } \boldsymbol{k}\text{-space}] \times \\ &\qquad\qquad [\text{Volume occupied by normal modes in the range } [k, k + \mathrm{d}k]] \end{aligned}$$

The volume element in $\boldsymbol{k}$-space is

$$\mathrm{d}k_x \, \mathrm{d}k_y \, \mathrm{d}k_z = k^2 \mathrm{d}k \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi,$$

where $k = \sqrt{k_x^2 + k_y^2 + k_z^2}$ is one of the three spherical-polar coordinates (k, θ, φ). We are concerned only with the magnitude of the wavenumbers, and not with their directions, hence

$$\begin{aligned} [\text{Volume occupied by normal modes in the range } [k, k + \mathrm{d}k]] &= \int_{\text{Positive octant}} \mathrm{d}k_x \, \mathrm{d}k_y \, \mathrm{d}k_z, \\ &= \int_{\text{Positive octant}} k^2 \mathrm{d}k \, \sin\theta \mathrm{d}\theta \, \mathrm{d}\varphi, \\ &= \frac{4\pi}{8} k^2 \mathrm{d}k, \\ &= \tfrac{1}{2}\pi k^2 \mathrm{d}k. \end{aligned}$$

[1] The dimensions of volume in $\boldsymbol{k}$-space are $1/$ [dimensions of volume on ordinary space]

Putting these last two results together, we have

$$\text{Number of normal modes in the range } [k, k+\mathrm{d}k] = \frac{L^3}{\pi^3} \times \tfrac{1}{2}\pi k^2 \mathrm{d}k$$

The solution (1.6) to the wave equation satisfies the equation of a harmonic oscillator:

$$\frac{\partial^2 \boldsymbol{E_k}}{\partial t^2} + c^2 \boldsymbol{k}^2 \boldsymbol{E_k} = 0.$$

We therefore recall some facts about statistical ensembles of classical harmonic oscillators. Given such a collection of oscillators, in thermal equilibrium, the average energy of each oscillator is $k_B T$, where k_B is Boltzmann's constant and T is temperature. The energy per unit wavenumber is therefore

$$\begin{aligned}\text{Energy in a wavenumber interval } [k, k+\mathrm{d}k] = \\ \text{Energy of a normal mode} \times \text{Number of normal modes in the range } [k, k+\mathrm{d}k],\\ = (k_B T) \times \left(\frac{L^3}{2\pi^2} k^2 \mathrm{d}k\right) \times 2,\end{aligned}$$

where the factor of 2 is introduced because each normal mode of vibration contains two polarisation states of light. Finally, we pass over to the wavelength variable, $\lambda = 2\pi/k$:

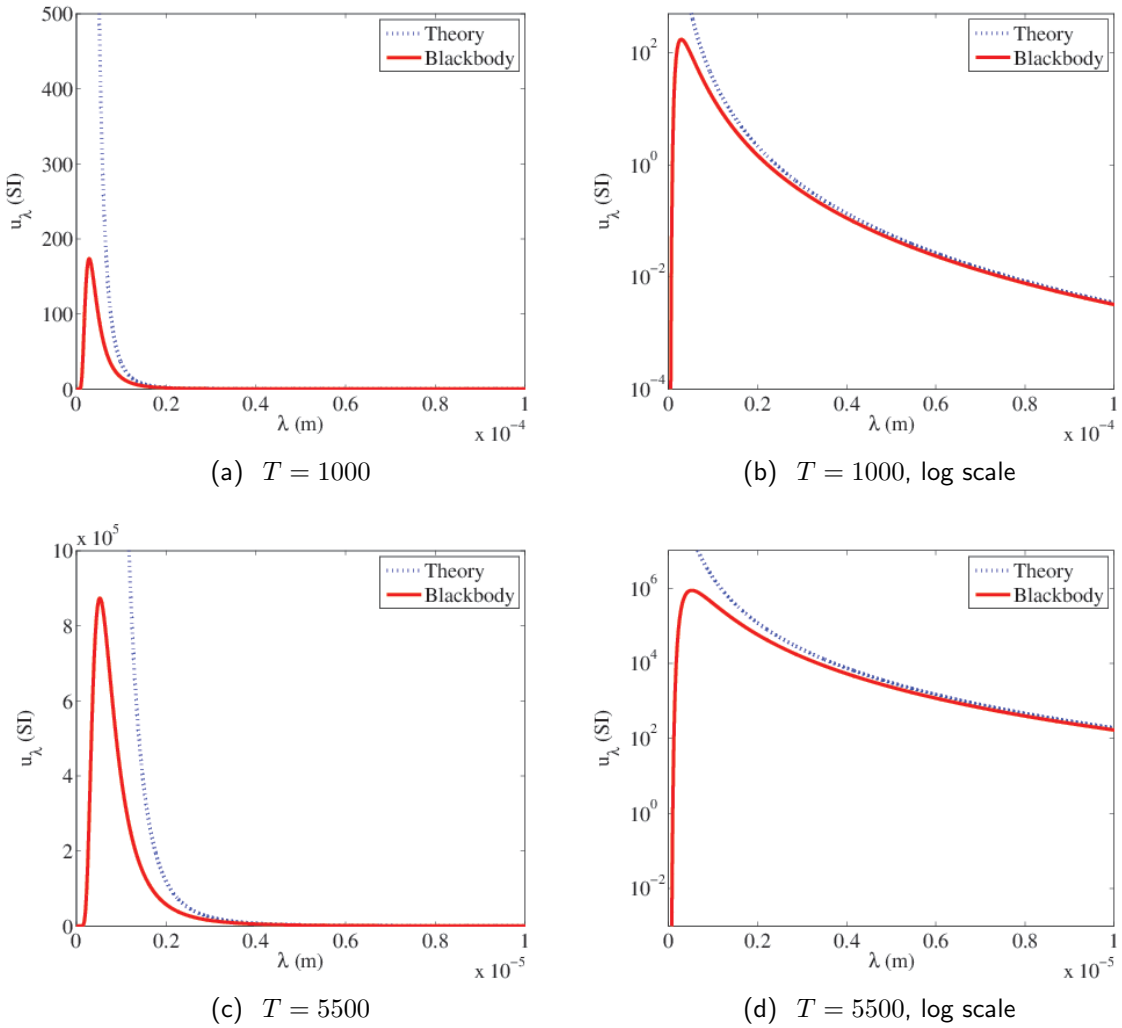
$$\begin{aligned}\text{Energy in a wavelength interval } [\lambda, \lambda+\mathrm{d}\lambda] &= k_B T \frac{L^3}{\pi^2}\left(\frac{2\pi}{\lambda}\right)^2 \left|\frac{dk}{d\lambda}\right| \mathrm{d}\lambda, \qquad k = \frac{2\pi}{\lambda},\\ &= \frac{8\pi k_B T L^3}{\lambda^4}\mathrm{d}\lambda,\\ \frac{\text{Energy in a wavelength interval } [\lambda, \lambda+\mathrm{d}\lambda]}{L^3} &= \frac{8\pi k_B T}{\lambda^4}\mathrm{d}\lambda,\end{aligned}$$

hence

$$u_\lambda(\lambda) = \frac{8\pi k_B T}{\lambda^4}.$$

We have computed the spectral density $u_\lambda(\lambda)$. This enables us to compute the total energy density of the blackbody:

$$\begin{aligned} u &= \int_0^\infty u_\lambda(\lambda)\mathrm{d}\lambda,\\ &= 8\pi k_B T \int_0^\infty \lambda^{-4}\mathrm{d}\lambda,\\ &= \tfrac{8}{3}\pi k_B T \lim_{\delta\to 0} \delta^{-3},\\ &= \infty. \end{aligned}$$



(a) $T = 1000$ (b) $T = 1000$, log scale

(c) $T = 5500$ (d) $T = 5500$, log scale

Figure 1.1: Spectral density of blackbody radiation, as a function of temperature and wavelength

But what has gone wrong ??? The best place to start a failure analysis for a theory is with experiments. It is a simple experiment to measure the intensity of light coming from (an approximate) blackbody, and hence to find the spectral density. Our failed theory is compared with the true (experimentally correct) curves in Fig. 1.1. The theory does appear to be correct in the long-wavelength limit. Only in the short-wavelength limit does the theory fail. Thus, the classical theory we have just derived is sometimes given the rather florid title of **the ultraviolet catastrophe**. Later on, we shall find out that our assumption that each normal mode of radiation behaves like a classical simple-harmonic oscillator is totally wrong. The classical oscillator can possess any amount of energy; if instead, we assume that the normal modes behave like quantum-mechanical oscillators, with discrete energy levels given by a quantum-mechanical calculation, then we shall recover the experimentally-correct curve. This will be the subject of future chapters.

1.3.2 Photoelectric effect

Hertz, 1887; Einstein, 1905

The **photoelectric effect** is the emission of electrons when light strikes a surface. The liberated electrons absorb energy from the incident radiation and are thus able to overcome the attractive forces that bind them to the surface. Hertz first observed the effect in 1887, and experiments (W. Hallachs and P. Lenard (1886-1900)) on the phenomenon defied classical explanation:

1. For incident light below a certain frequency, NO electrons are emitted. This is called the **threshold frequncy**.

2. Increasing the intensity of the light, while maintaining the frequency below threshold, does NOT cause electrons to be emitted;

3. Indeed, the energy of emitted electrons is independent of the intensity of the incident light. Since the intensity is a measure of the energy carried by the incident light, one would expect a higher intensity to lead to more energetic emitted electrons.

Einstein proposed that the incident light must be quantised. In other words, the incident light has a particle nature. The particles of light are massless and are called 'photons'; a photon carrying light of frequency ν has energy

$$E = h\nu,$$

where h is **Planck's constant**, and is a fundamental unit of angular momentum. Now the bound electrons have an energy $-\phi$, where ϕ is the 'work function', or the potential energy binding the electrons to the surface. Thus, the initial energy of the system (photon+electron) is

$$h\nu - \phi,$$

while the final energy is simply the kinetic energy of the liberated electron, $m_{\mathrm{e}}v^2/2$. Since energy is conserved,

$$\tfrac{1}{2}m_{\mathrm{e}}v^2 = h\nu - \phi.$$

Thus, points 1-2 are explained: The threshold frequency is $h\nu = \phi$, since photons below this frequency would cause the liberated electron to have a negative kinetic energy – impossible.

The intensity of the incident light is a measure of its energy content, per unit time, per unit area. If we divide the intensity of a monochromatic source by $h\nu$, we obtain a measure of the number of photons incident on the surface, per unit time, per unit area. Thus, the intensity controls the number of photons, **but not the photon energy**. Increasing the intensity of the incident light

will increase the number of photons, and hence, the number of emitted electrons, **but it will not increase the energy of individual emitted electrons**.

Einstein's explanation is satisfactory, but it does not fit into any overall theoretical framework. In particular, there is no description of the dynamics of the light- and electron-particles, and no description of their interaction.

1.3.3 The emission spectrum of the hydrogen atom

Bohr, 1913

Hydrogen is the simplest atom, and consists of one electron of negative charge $(-e)$ that is bound to a much more massive proton, of positive charge $(+e)$. Classically, the electron can be thought of as 'in orbit' around a fixed force centre, with potential energy

$$\mathcal{U}(r) = -\frac{1}{4\pi\epsilon_0}\frac{e^2}{r}$$

The energy of the atom is therefore

$$E = \tfrac{1}{2}m_{\rm e}\boldsymbol{v}^2 - \frac{1}{4\pi\epsilon_0}\frac{e^2}{r}.$$

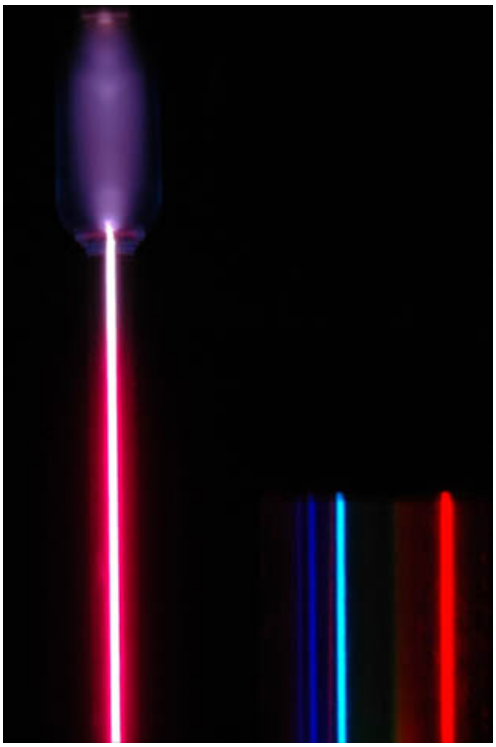
The electron binds to the proton provided $E \leq 0$. One can imagine an electromagnetic interaction where an excited electron $E \lesssim 0$ de-excites to a more stable state (a more negative E-value) by the emission of electromagnetic radiation (a photon of light). In this scenario, the excited electron can have any negative energy E, and therefore, a continuous spectrum of emitted light must be possible. However, this is not the case. It is an experimental fact that the spectrum of hydrogen atom is a sequence of lines (Fig. 1.2)

Before the advent of the Schrödinger equation, Bohr (1913) proposed an *ad-hoc* model to describe the spectrum of hydrogen. He first noticed that circular orbits solve the orbit problem

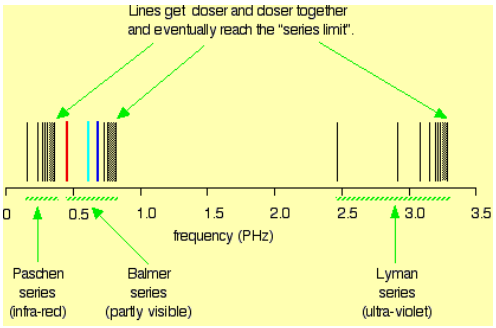
$$m_{\rm e}\frac{d^2\boldsymbol{x}}{dt^2} = -\frac{e^2}{4\pi\epsilon_0}\frac{\boldsymbol{x}}{|\boldsymbol{x}|^3},$$

or

$$\begin{aligned} m_{\rm e}\frac{d^2r}{dt^2} &= -\frac{e^2}{4\pi\epsilon_0}\frac{1}{r^2} + \frac{J^2}{m_{\rm e}r^3}, \\ \frac{d}{dt}\left(m_{\rm e}r^2\frac{d\theta}{dt}\right) &= 0, \qquad m_{\rm e}r^2\frac{d\theta}{dt} = J, \end{aligned}$$



(a)



(b)

Figure 1.2: (a) The photograph comes the HyperPhysics website (Rod Nave, GSU). It shows part of a hydrogen discharge tube on the left, and the three most easily seen lines in the visible part of the spectrum on the right. Ignore the blurring – particularly to the left of the red line. This is caused by flaws in the way the photograph was taken; (b) A schematic interpretation of (a), showing other emission lines not in the visible range (chemguide.co.uk).

in polar coordinates. Indeed, circular orbits are an equilibrium solution, $d^2r/dt^2 = 0$, provided

$$\frac{e^2}{4\pi\epsilon_0}\frac{1}{r^2} = \frac{J^2}{m_e r^3},$$

or

$$\frac{e^2}{4\pi\epsilon_0}\frac{1}{r} = \frac{J^2}{m_e r^2} = m_e v^2. \tag{1.7}$$

Here J is the angular momentum. Bohr hypothesised that the angular momentum should be quantised:

$$J = J_n = n\hbar,$$

where $\hbar = h/2\pi$ is a fundamental unit of angular momentum and n is a positive integer. This in turn implies that the velocity v and radius r of the circular orbits can take only discrete values:

$$J_n = m_e v_n r_n = n\hbar. \tag{1.8}$$

Substituting the quantisation rule (1.8) into the circular-orbit condition (1.7), we have

$$\frac{1}{4\pi\epsilon_0}\frac{e^2}{r_n} = m_e v_n^2.$$

We solve the equations

$$\begin{aligned} m_e v_n r_n &= n\hbar, \\ \frac{1}{4\pi\epsilon_0}\frac{e^2}{r_n} &= m_e v_n^2, \end{aligned}$$

for r_n and obtain

$$r_n = \frac{n^2\hbar^2}{m_e}\frac{4\pi\epsilon_0}{e^2}.$$

Thus, the radii of the electron orbits are not random, but are rather square-integer multiples of the basic radius

$$a_0 = \frac{\hbar^2}{m_e}\frac{4\pi\epsilon_0}{e^2}.$$

Using the circular-orbit expression

$$E = \tfrac{1}{2}m_e v^2 - \frac{e^2}{4\pi\epsilon_0}\frac{1}{r} = -\frac{J^2}{2m_e r^2},$$

n_1	n_2	ΔE/eV	λ/μm	Name
1	2	10.2	0.121	Lyman-alpha (Ultra-violet)
2	3	1.88	0.656	Balmer-alpha (red)
2	4	2.55	0.486	Balmer-beta (blue-green)
2	5	2.85	0.434	Balmer-gamma (Violet)
2	6	3.02	0.410	Balmer-delta (Violet)
2	7	3.12	0.397	Balmer-epsilon (Ultra-violet)

Table 1.1: Some spectroscopic lines of hydrogen (emission spectrum), computed from the Bohr model. The visible lines form part of the Balmer series. These lines correspond exactly with experimental observations of the light emitted from hydrogen.

we obtain the **quantisation of energy**,

$$\begin{aligned} E_n &= -\frac{J_n^2}{2m_{\mathrm{e}}r_n^2}, \\ &= -\frac{n^2\hbar^2}{2m_{\mathrm{e}}}\left(\frac{m_{\mathrm{e}}}{n^2\hbar^2}\frac{e^2}{4\pi\epsilon_0}\right)^2, \\ &= -\tfrac{1}{2}\frac{m_{\mathrm{e}}e^4}{(4\pi\epsilon_0\hbar)^2}\frac{1}{n^2}, \\ &:= -\frac{E_0}{n^2}, \end{aligned} \tag{1.9}$$

The fundamental unit of angular momentum is **Planck's constant**, $h = 6.626\times10^{-34}\,\mathrm{kg\,m^2\,s^{-1}}$, and $\hbar = h/2\pi$. Using this information, the energy E_0 is computed to be

$$E_0 = \tfrac{1}{2}\frac{m_{\mathrm{e}}e^4}{(4\pi\epsilon_0\hbar)^2} = 2.1798\times10^{-18}\,\mathrm{kg\,m^2\,s^{-2}} = 13.60\,\mathrm{eV},$$

where $1\,\mathrm{eV} = 1.602\times10^{-19}\,\mathrm{kg\,m^2\,s^{-2}}$. Sometimes E_0 as written as $E_0 = 1\mathrm{Ryd}$, the **Rydberg**. Remember this value for the whole module!!

From the discussion on the photoelectric effect, we know that light is made up of discrete photons. Let us assume that a single photon is produced as an electron de-excites from a high energy level E_{n_2} down to a less energetic state E_{n_1}, where $n_2 > n_1$. We tabulate some of these energies (Tab. 1.1) and the corresponding photon wavelengths, using

$$\Delta E_{n_1,n_2} = E_0\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right) = h\nu_{n_1,n_2} = \frac{hc}{\lambda_{n_1,n_2}}, \qquad n_2 > n_1.$$

The visible lines of the Balmer series ($n_1 = 2$) correspond exactly with the experimental pictures (Fig. 1.2). Bohr's theory works! The transitions are shown schematically in Fig. 1.3. Unfortunately, there is no reason to assume that the angular momentum is quantised – we need a more complete

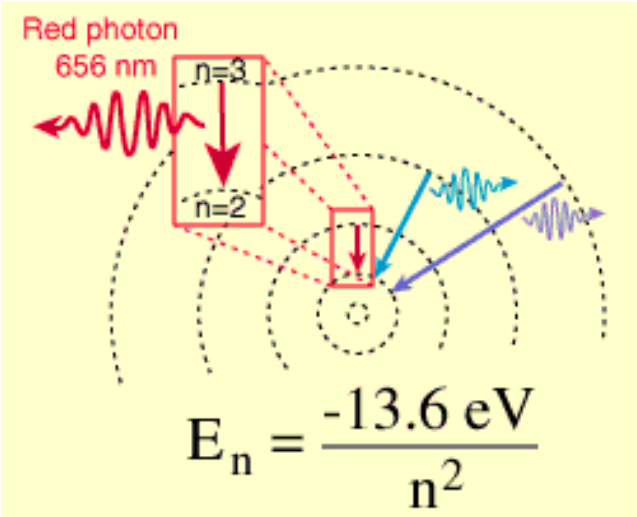


Figure 1.3: Schematic description of the transition of the electron to lower energies, corresponding to the Balmer (visible) series.

theory to justify this.

1.3.4 Diffraction patterns formed by electron beams

de Broglie, 1924; Davisson and Germer, 1927

The examples of blackbody radiation and the photoelectric effect suggest that light comprises particles that obey 'strange dynamics', and are not described by classical mechanics. In this section, we show that particles – electrons – can exhibit wave-like behaviour. Only by invoking a complete theory of quantum mechanics can the apparent contradiction of this **wave-particle duality** be overcome.

As an alternative to Bohr's resolution of the hydrogen problem, de Broglie (1924) proposed that particles have a wave-like behaviour; the particle wavenumber k is related to the particle momentum p through a Planck-type equation,

$$p = \hbar k = \frac{h}{\lambda}. \tag{1.10}$$

Since the electron that is bound to the hydrogen atom is confined in some sense, it must correspond to a 'standing wave', in which the wavelength 'fits' into the confining domain. In other words, the standing wave must be related to the radius of the circular orbit in a rational way:

$$n\lambda_n = 2\pi r_n.$$

But $\lambda_n = h/p = h/mv_n$, hence

$$\frac{nh}{mv_n} = 2\pi r_n,$$

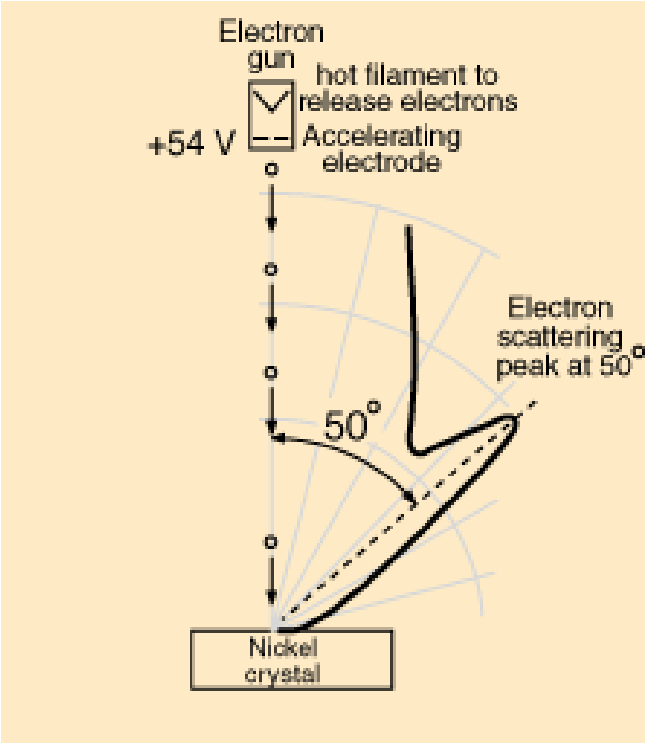


Figure 1.4: Schematic description of the experiment of Davisson and Germer (1927).

or

$$J_n = mv_n r_n = \frac{nh}{2\pi} = n\hbar.$$

This is equivalent to Bohr's quantisation of angular momentum! This consistency is reassuring, and provides support for de Broglie's hypothesis.

However, the only true test of such speculative theories is experiment. In 1927, Davisson and Germer fired a beam of electrons at a crystal sample (Fig. 1.4). The intensity of the scattered beam was very strong at certain scattering angles, and weak at others. They drew a plot of the intensity of the scattered beam as a function of the angle θ, and found a functional form that could only be described by making the assumption that the scattered beam was in fact a wave. Referring to Fig. 1.5, two neighbouring waves emerging from the crystal are in phase (constructive interference) provided

$$m\lambda = d\sin\theta, \qquad m = 1, 2, 3, \cdots, \tag{1.11}$$

where d is the crystal spacing and λ is the de Broglie wavelength.

Example: In a particular electron-diffraction experiment using an accelerating voltage of $54\,\mathrm{V}$, an intensity maximum occurs when the scattering angle is $\theta = 50^o$. The initial kinetic energy of the electrons is negligible. The rows of atoms are known to have a separation $d = 2.15 \times 10^{-10}$m. Find the electron wavelength (a) from the diffraction formula; (b) from the de Broglie hypothesis.

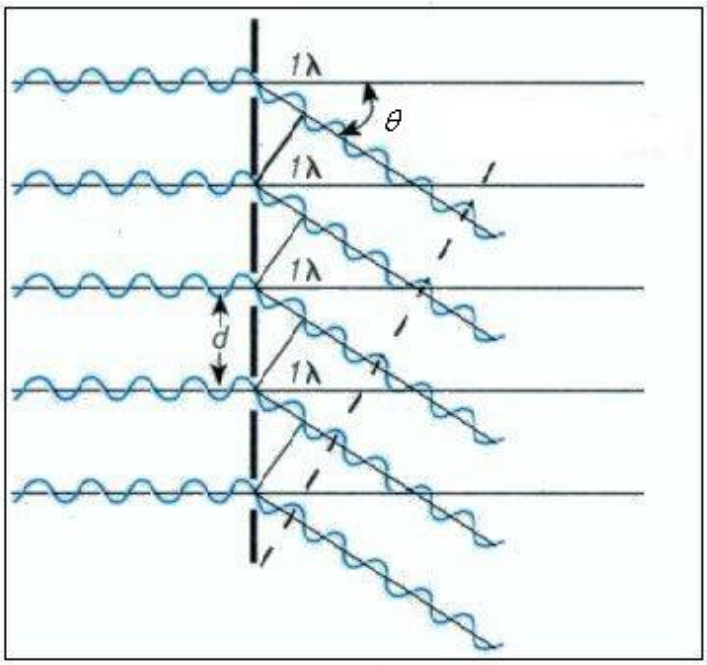


Figure 1.5: Constructive interference: The reflected waves are in phase provided that the difference between the path length of neighbouring waves is an integer number of wavelengths.

Compare the results.

From Eq. (1.11), with $m = 1$,

$$\begin{aligned} \lambda &= d \sin\theta, \\ &= \left(2.15 \times 10^{-10}\mathsf{m}\right) \sin 50^o, \\ &= 1.65 \times 10^{-10}\mathsf{m}. \end{aligned}$$

Using the work-energy theorem, the work done on the electron ($= eV$) is equal to the kinetic energy gained:

$$eV = \frac{p^2}{2m_\mathrm{e}},$$

hence

$$p = \sqrt{2m_\mathrm{e}eV},$$

and

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2m_\mathrm{e}eV}}.$$

Putting in the numbers,

$$\begin{aligned} \lambda &= \frac{6.626 \times 10^{-34}\mathsf{J}\cdot\mathsf{s}}{\sqrt{2\,(9.109 \times 10^{-31}\mathsf{kg})\,(1.602 \times 10^{-19}\mathsf{Coulomb})\,(54\mathsf{V})}}, \\ &= 1.67 \times 10^{-10}\mathsf{m}, \end{aligned}$$

and the two numbers agree to within the accuracy of the experimental results.

Having described several experiments where classical mechanics demonstrably fails, and having advanced several *ad-hoc* theories to describe these phenomena, we turn to the rigorous formulation of the axioms of quantum mechanics.

Chapter 2

The mathematical foundation of quantum mechanics

Reading material for this chapter: Feynman (Vol. 3, Chapters 1 and 3); Simms 211

In this chapter, we develop a framework to describe the phenomena described in Ch. 1 using generic rules. There are two approaches here: the first, very intuitive approach, comes from Feynman. The second is more abstract and mathematical, and can be regarded as a neat distillation of the first approach into a few laws whose form we will investigate in later chapters.

2.1 The two-slit experiment

Consider the double-slit experiment involving waves (water waves, light, sound), shown in Fig. 2.1. Waves emanate from a source and hit a wall containing two slits. We know from Huygens' Principle that the slits act as 'new' wave sources. Therefore, to study the pattern formed by waves moving between the slits and the detector, it suffices to imagine the interaction between two individual wave patterns, sourced at slit 1, and slit 2.

Suppose that a vector $\boldsymbol{E}_i$ describes the wave pattern emanating from source i. This might be the electro-magnetic field of light, or the velocity field of a water wave. The dynamics of such waves is linear. Therefore, the total vector for the combined pattern coming from both secondary sources is

$$\boldsymbol{E} = \boldsymbol{E}_1 + \boldsymbol{E}_2.$$

Typically, this is a complex-valued vector, with phases like $\mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}$. Finally, the intensity pattern observed at the detector is related to energy; the energy of these waves is related to the square of the total wave vector:

$$I_{12} = |\boldsymbol{E}|^2 = |\boldsymbol{E}_1|^2 + |\boldsymbol{E}_2|^2 + 2\Re\left(\boldsymbol{E}_1 \cdot \boldsymbol{E}_2\right).$$

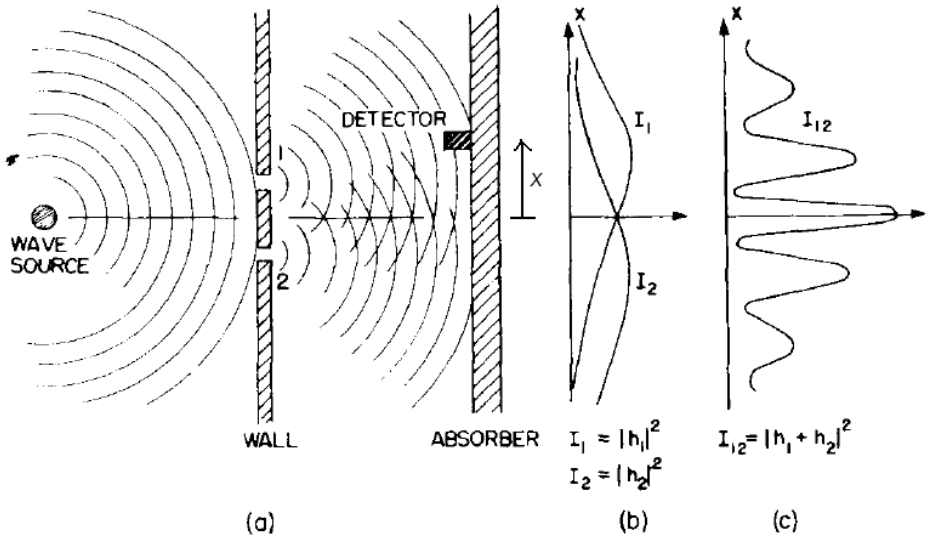


Figure 2.1: Interference pattern from a wave source

That the intensity of the combined waves is not the sum of the individual intensities is called **interference**.

Next, we conduct a thought experiment where we replace the wave source with an electron gun. We also reduce the size of the slits to a width that compares with the spacing of a crystal lattice (as in the experiment of Davisson and Germer). Technically, we should replace the **absorber** with a **backstop**, and replace the detector with one capable of observing electrons. We compute the probability that electrons, starting at the source, pass through slit 1 OR slit 2, and arrive at the detector, located at position x. This is called P_{12}. Practically, this can be computed as

$$P_{12}(x) = \frac{\text{Average number of clicks made by electron detector at location } x\text{, per unit time (both slits open)}}{\text{Number of electrons emitted by gun, per unit time}}$$

Next, we close off slit 2 and compute the probability that electrons, starting at the source, pass through slit 1, and arrive at x:

$$P_1(x) = \frac{\text{Average number of clicks made by electron detector at location } x\text{, per unit time (slit 2 closed)}}{\text{Number of electrons emitted by gun, per unit time}}$$

Similarly, we compute P_2. We find,

$$P_{12} \neq P_1 + P_2,$$

which is the same result as the case of waves. Thus, it appears as though the two sources of electrons interfere.

We are therefore motivated to ascribe a **complex-valued probability amplitude** to events:

ϕ_1 = Probability amplitude for electron to leave the source, pass through slit 1, and end up at x,

ϕ_2 = Probability amplitude for electron to leave the source, pass through slit 2, and end up at x

such that

$$P_1 = |\phi_1|^2, \qquad P_2 = |\phi_2|^2.$$

We know that the probabilities do not add, so we propose instead that the **probability amplitudes add**:

$$\phi_{12} = \phi_1 + \phi_2,$$

In other words,

Probability amplitude for electron to leave source, pass through slit 1 OR slit 2, and arrive at x
= Probability amplitude for electron to leave source, pass through slit 1, and arrive at x
+ Probability amplitude for electron to leave source, pass through slit 2, and arrive at x.

But we have interpreted the modulus-squared of an amplitude as a probability, hence

$$P_{12} = |\phi_1 + \phi_2|^2 = |\phi_1|^2 + |\phi_2|^2 + 2\Re(\phi_1\phi_2) = P_1 + P_2 + 2\Re(\phi_1\phi_2).$$

Thus, the two events interfere, just like ordinary waves.

For the second part of the thought experiment, we imagine **observing** the electrons as they pass through one of the two slits. To do this, we place a light source near the slits, between the slits and the backstop (Fig. 2.2). We know that light (photons) scatter off electrons. Thus, whenever we see a flash of light near A, we know that the electron has gone through slit 2; if we see a flash of light nearer to slit 1, then we conclude that the electron has gone through slit 1. In this way, we build up new probabilities:

P_1' = Probability that the electron leaves the source, passes through slit 1,, scatters light, and ends up at x,

P_2' = Probability that the electron leaves the source, passes through slit 2,, scatters light, and ends up at x,

P_{12}' = Probability that the electron leaves the source, passes through slit 1 OR slit 2, scatters light, and ends up at x.

Remarkably, we find that

$$P_{12}' = P_1' + P_2'.$$

Thus, we no longer get the interference pattern (Probability as a function of x) associated with the original experiment. Indeed, the pattern of probabilities observed is as though we were firing bullets through the slits (Fig. 2.2). Continuing with the thought experiment, **the original interference**

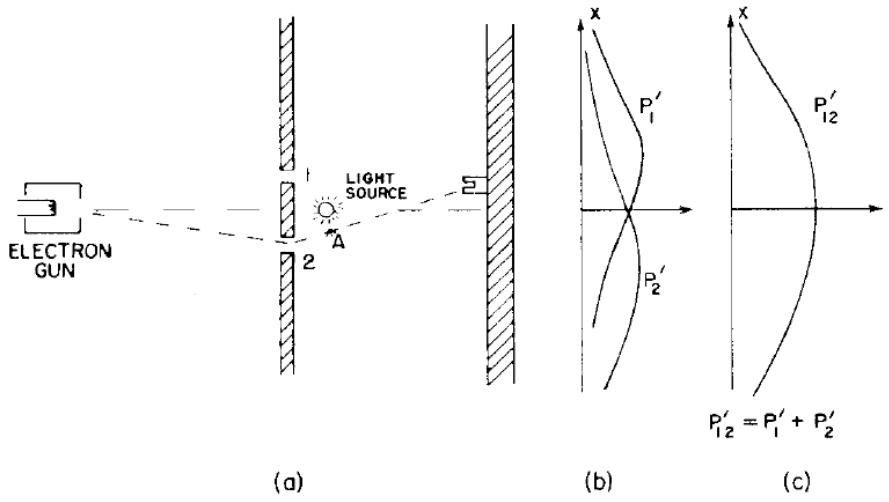


Figure 2.2: Electron source, slits illuminated (interference pattern collapses)

pattern is observed when the light source is switched back off.

This thought experiment (which is supported by real experiments) leads to the following rules for combining probabilities:

1. The probability of an event is given by the square of the absolute value of a complex number ϕ which is called the probability amplitude:

$$\begin{aligned} P &= \text{Probability;} \\ \phi &= \text{probability ampliutde;} \\ P &= |\phi|^2 \end{aligned}$$

2. When an event can occur in several alternative ways, the probability amplitude for the event is the sum of the probability amplitudes for each way considered separately. There is interference:

$$\begin{aligned} \phi &= \phi_1 + \phi_2, \\ P &= |\phi_1 + \phi_2|^2. \end{aligned}$$

   If an experiment is performed which is capable of determining whether one or another alternative is actually taken, the probability of the event is the sum of the probabilities of each alternative. The interference is lost:

$$P = P_1 + P_2.$$

2.2 Manipulating probability amplitudes

Consider again the interference pattern generated by the two-slit experiment with electrons, when there is no way of knowing which slit the electrons have passed through (Fig. 2.1). We are going to use some new notation for the probability amplitudes. For the detector at position x, we have[1]

Amplitude that electron leaves source s and arrives at x

$$= \langle \text{Particle arrives at } x | \text{Particle leaves } s \rangle = \langle x|s\rangle.$$

We know that probability amplitudes add, thus the amplitude for the particle to arrive at x is given by a sum over all possible routes of getting there:

$$\langle x|s\rangle_{\text{both slits open}} = \langle x|s\rangle_{\text{through 1}} + \langle x|s\rangle_{\text{through 2}}.$$

To these results we may add a further general principle (rule 3):

> When a particle goes by some particular route, the amplitude for that route can be written as the product of the amplitude to go part of the way, with the amplitude to go the rest of the way.

Thus,

$$\langle x|s\rangle_{\text{through 1}} = \langle x|1\rangle\langle 1|s\rangle.$$

But

$\langle x|s\rangle_{\text{through 1}}$

$=$ Probability amplitude for electron to leave source, pass through slit 1, and arrive at x

$$= \phi_1;$$

similarly

$$\phi_2 = \langle x|s\rangle_{\text{through 2}} = \langle x|2\rangle\langle 2|s\rangle.$$

Combining these, we have the total amplitude for the electron to reach the detector:

$$\langle x|s\rangle_{\text{both slits open}} = \langle x|1\rangle\langle 1|s\rangle + \langle x|2\rangle\langle 2|s\rangle.$$

Referring back to the law

$$P_{12} = |\phi_1 + \phi_2|^2,$$

[1] The order is rather strange here: final state $\langle f|$ on the left, initial state $|i\rangle$ on the right, leading to an amplitude $\langle f|i\rangle$. According to a friend from undergraduate days, the order of these terms is precisely the same as the order of the first letters of the two-word expression 'Feck it' – a rather curious mnemonic.

we have

$$P(x; \text{both slits open}) = |\langle x|s\rangle_{\text{both slits open}}|^2 = |\langle x|1\rangle\langle 1|s\rangle + \langle x|2\rangle\langle 2|s\rangle|^2 .$$

2.3 Distinguishable alternatives

We return to the problem of measuring which slit the electrons pass through. This is done by placing two detectors between the slits and the backstop (Fig. 2.3). The amplitude for a particle to start

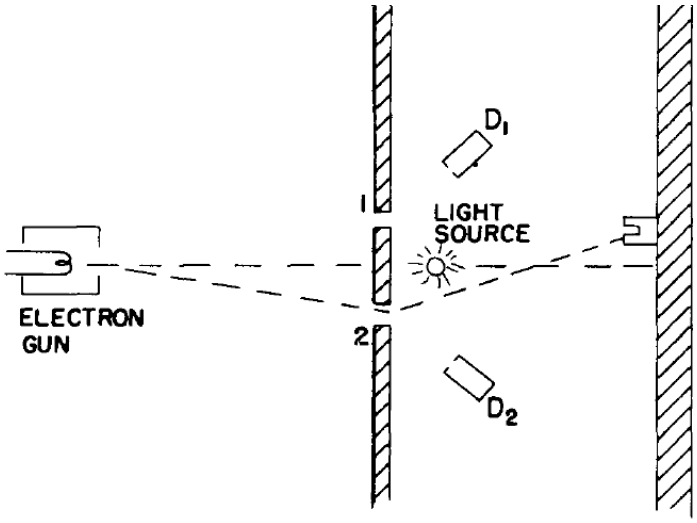


Figure 2.3:

at the source s, go through slit 1, scatter off a photon that goes into detector D_1, and proceeds to location x is

$$\langle x|1\rangle a\langle 1|s\rangle = a\phi_1,$$

where a is the probability amplitude that the electron at slit 1 scatters a photon that goes into detector D_1. Now we also have to allow for the possibility that an electron going through slit 2 scatters a photon into detector D_1, although this would be a poorly-designed experiment, since we wish for detector D_1 to mark out electrons going into slit 1. Nevertheless, we have

$$\langle x|2\rangle b\langle 2|s\rangle = b\phi_2$$

where b is the probability amplitude that the electron at slit 2 scatters a photon that goes into detector D_1. Thus,

$$\langle \text{electron at } x,\ \text{photon at } D_1|\text{electron from } s,\ \text{photon from light source}\rangle = a\phi_1 + b\phi_2, \qquad (2.1)$$

where we sum over the two indistinguishable alternatives. Now we assume that the system is totally symmetric between slits and detectors, so that

$$\langle \text{electron at } x,\ \text{photon at } D_2 | \text{electron from } s,\ \text{photon from light source} \rangle = a\phi_2 + b\phi_1.$$

Thus, the probability to get a detection of light in D_1 and an electron at x is the absolute value squared of Equation (2.1):

$$\text{Prob}(\text{Detection in } D_1, \text{electron at } x) = |a\phi_1 + b\phi_2|^2.$$

If we design our experiment well, then $b = 0$, and

$$\text{Prob}(\text{Detection in } D_1, \text{electron at } x) = |a\phi_1|^2,$$

so that up to a silly prefactor ($|a|^2$), a detection of light in D_1 corresponds to an electron passing through slit 1. Now here comes the crux. We want to find the interference pattern at the detector, in other words, we want to find the probability that an electron ends up at x, **regardless of which detector it scatters light into**. Should we add some amplitudes? The answer is NO. **We never add amplitudes for distinguishable final states** (rule 4). Detection of photons in one device is completely independent of detection of photons in another. Thus, in this case, the probabilities add:

$$\begin{aligned}\text{Prob}(&\text{Detection in } D_1 \text{ OR } D_2, \text{electron at } x) \\ &= \text{Prob}(\text{Detection in } D_1, \text{electron at } x) + \text{Prob}(\text{Detection in } D_1, \text{electron at } x) \\ &= |a\phi_1|^2 + |a\phi_2|^2 = P_1' + P_2'.\end{aligned}$$

2.4 The mathematical postulates of quantum mechanics

In the first part of this chapter, we have used thought experiments, based on real experiments, to derive some rules for quantum-mechanical behaviour. While very intuitive, it is also quite long. That discussion can be distilled into the following few mathematical rules. You will probably not understand all of them yet; that is the purpose of Chapters 3–9. However, after studying these chapters, you should try to reconcile the following laws with Feynman's more intuitive explanations.

Mathematical postulates of quantum mechanics:

1. Each physical system is associated with a (separable) Hilbert space $\mathcal{H}$. Norm-one vectors in $\mathcal{H}$ are associated with states of the system.

The norm can be re-constituted by pairing with elements in the dual space (Riesz Representation Theorem). Norm-one vectors that differ only by a phase represent the same state.

2. The Hilbert space of a composite system is the Hilbert-space tensor-product of the state spaces associated with the component systems.

3. Physical symmetries act on $\mathcal{H}$ through unitary or conjugate-unitary operators.

4. A physical observable is represented by a Hermitian operator on $\mathcal{H}$; the only allowed results of measurement of the physical observable are the eigenvalues of the operator.

   Quantum mechanics is probabilistic: the probability that a system prepared in a state $|x\rangle \in \mathcal{H}$ is measured to be in an eigenstate $|x_a\rangle$ of the observable A is given by

   $$|\langle x_a|x\rangle|^2,$$

   where $\langle x_a| \in \mathcal{H}^*$ is dual to $|x_a\rangle \in \mathcal{H}$ and $\langle x_a|x\rangle$ is the pairing induced by the scalar product.

Notes: The Schrödinger equation is not a fundamental postulate of quantum mechanics; it can be derived from these laws. the same is true for Heisenberg's uncertainty principle. To understand these concepts, we shall need to recall the concept of a vector space over the complex numbers.

Chapter 3

Complex vector spaces

It is left as homework to study this chapter; Reading material for this chapter: Simms 211.

Definition 3.1 *A set $\mathcal{H}$ is called a* **complex vector space** *if the following properties hold:*

1. *An operation*

$$\begin{aligned} \mathcal{H} \times \mathcal{H} &\rightarrow \mathcal{H}, \\ (\boldsymbol{x}, \boldsymbol{y}) &\rightarrow \boldsymbol{x} + \boldsymbol{y} \end{aligned}$$

   is given, called **addition of vectors**, *such that*

   (a) *The addition is associative:* $(\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z} = \boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z})$;

   (b) *The addition is commutative:* $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$;

   (c) *There is an additive identity:* $\boldsymbol{x} + 0 = \boldsymbol{x}$;

   (d) *There are inverses:* $\boldsymbol{x} + (-\boldsymbol{x}) = 0$.

   These properties make the vector space into an **abelian group**.

2. *An operation*

$$\begin{aligned} \mathbb{C} \times \mathcal{H} &\rightarrow \mathcal{H}, \\ (\lambda, \boldsymbol{x}) &\rightarrow \lambda \boldsymbol{x} \end{aligned}$$

   is given, called **scalar multiplication**, *which satisfies*

   (a) *The multiplication is distributive:* $\lambda(\boldsymbol{x} + \boldsymbol{y}) = \lambda \boldsymbol{x} + \lambda \boldsymbol{y}$;

(b) Distributivity: $(\lambda + \mu)\boldsymbol{x} = \lambda\boldsymbol{x} + \mu\boldsymbol{x}$*;*

(c) Distributivity: $(\lambda\mu)\boldsymbol{x} = \lambda(\mu\boldsymbol{x})$*;*

(d) $1 \in \mathbb{C}$ *is a multiplicative identity* $1\boldsymbol{x} = \boldsymbol{x}$*,*

for all $\lambda, \mu \in \mathbb{C}$ *and* $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathcal{H}$*. The elements of* $\mathcal{H}$ *are called* **vectors** *and the elements of* $\mathbb{C}$ *are, in this context, called* **scalars**.

Examples:

1. The set
$$\mathbb{C}^n = \{(z_1, \cdots z_n) \,|z_1, \cdots z_n \in \mathbb{C}\,\}$$
is a complex vector space.

2. The set $\mathcal{H}_\Omega$ of all complex-valued functions of a real variable,
$$\mathcal{H}_\Omega = \{f | f : (\Omega \subset \mathbb{R}) \to \mathbb{C}\,\}$$
is a vector space, with vector addition
$$(f + g)(x) = f(x) + g(x),$$
and scalar multiplication
$$(\lambda f)(x) = \lambda f(x),$$
for all $x \in \Omega$, all $f, g \in \mathcal{H}_\Omega$, and $\lambda \in \mathbb{C}$. Note that the addition operation is called **pointwise** because it is defined with reference to each point $x \in \Omega$.

3. The set of all solutions of the equation
$$\frac{d^2u}{dx^2} + u = 0$$
is a vector space.

Definition 3.2 *Let* $\mathcal{G} \subset \mathcal{H}$ *and let* $\mathcal{H}$ *be a complex vector space. Then* $\mathcal{G}$ *is called a* **vector subspace of** $\mathcal{H}$ *if it is non-empty, and if,*

1. *Closure under addition:* $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{G} \implies \boldsymbol{x} + \boldsymbol{y} \in \mathcal{G}$*;*

2. *Closure under scalar multiplication:* $\lambda \in \mathbb{C}\, \boldsymbol{x} \in \mathcal{G} \implies \lambda\boldsymbol{x} \in \mathcal{G}$*.*

Thus, $\mathcal{G}$ is itself a complex vector space.

Example: Let $\mathcal{H}_\Omega$ be the set of all complex-valued functions of a single real variable, with domain Ω. Let $f \in \mathcal{H}_\Omega$. Define the L^2 norm of f:

$$\|f\|_2 = \sqrt{\int_\Omega \mathrm{d}x |f(x)|^2}.$$

The set

$$\mathcal{G}_\Omega = \{f \in \mathcal{H}_\Omega | \|f\|_2 < \infty\}$$

is closed under addition and scalar multiplication, and is therefore a vector subspace of $\mathcal{H}_\Omega$.

Definition 3.3 *Let $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$ be vectors in the complex vector space $\mathcal{H}$, and let $\lambda_1, \cdots \lambda_r$ be scalars. Then the vector*

$$\lambda_1 \boldsymbol{x}_1 + \cdots + \lambda_r \boldsymbol{x}_r$$

is called a **linear superposition** *of $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$. We write*

$$\mathcal{S}(\boldsymbol{x}_1, \cdots \boldsymbol{x}_r) = \{\lambda_1 \boldsymbol{x}_1 + \cdots \lambda_r \boldsymbol{x}_r | \lambda_1, \cdots, \lambda_r \in \mathbb{C}\}$$

to denote the set of all linear combinations of $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$. $\mathcal{S}(\boldsymbol{x}_1, \cdots \boldsymbol{x}_r)$ is a vector subspace of $\mathcal{H}$, and is called the subspace **spanned** *by $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$.*

If $\mathcal{S}(\boldsymbol{x}_1, \cdots \boldsymbol{x}_r) = \mathcal{H}$, we say that $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$ **span** *the whole space $\mathcal{H}$. Then, for each $\boldsymbol{x} \in \mathcal{H}$, there exist scalars $\lambda_1, \cdots, \lambda_r$ such that*

$$\boldsymbol{x} = \lambda_1 \boldsymbol{x}_1 + \cdots \lambda_r \boldsymbol{x}_r.$$

Examples:

1. The vectors

$$\begin{aligned} \boldsymbol{e}_1 &= (1,0,0), \\ \boldsymbol{e}_2 &= (0,1,0), \\ \boldsymbol{e}_3 &= (0,0,1) \end{aligned}$$

span $\mathbb{C}^3$ because for any $\boldsymbol{x}$ in $\mathbb{C}^3$,

$$\begin{aligned} \boldsymbol{x} &= (z_1, z_2, z_3), \\ &= z_1 \boldsymbol{e}_1 + z_2 \boldsymbol{e}_2 + z_3 \boldsymbol{e}_3. \end{aligned}$$

2. The functions

$$\mathrm{e}^{\mathrm{i}x}, \qquad \mathrm{e}^{-\mathrm{i}x}$$

span the space of solutions of the equation

$$\frac{d^2u}{dx^2} + u = 0,$$

because for any $u(x)$ in the solution space,

$$u(x) = \lambda_1 \mathrm{e}^{\mathrm{i}x} + \lambda_2 \mathrm{e}^{-\mathrm{i}x}.$$

Definition 3.4 *Let $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$ be vectors in a complex vector space $\mathcal{H}$. Then,*

1. *$\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$ are* **linearly dependent** *if there exist scalars $\lambda_1, \cdots, \lambda_r$ such that*

$$\lambda_1 \boldsymbol{x}_1 + \cdots \lambda_r \boldsymbol{x}_r = 0.$$

2. *They are* **linearly independent** *if*

$$\lambda_1 \boldsymbol{x}_1 + \cdots \lambda_r \boldsymbol{x}_r = 0$$

implies that $\lambda_1 = \cdots = \lambda_r = 0$.

Example: $\mathrm{e}^{\mathrm{i}x}$ and $\mathrm{e}^{-\mathrm{i}x}$ are linearly independent functions in $\mathcal{H}_\Omega$. For, let us solve

$$\lambda \mathrm{e}^{\mathrm{i}x} + \mu \mathrm{e}^{-\mathrm{i}x} = 0, \qquad \text{for all } x \in \mathbb{C}.$$

Since this expression must be true for all $x \in \mathbb{R}$, set $x = 0$. Since $\mathrm{e}^0 = 1$ we have

$$\mu = -\lambda.$$

Thus, we have

$$0 = \lambda \mathrm{e}^{\mathrm{i}x} + \mu \mathrm{e}^{-\mathrm{i}x} = \lambda \left(\mathrm{e}^{\mathrm{i}x} - \mathrm{e}^{-\mathrm{i}x} \right) = 2\mathrm{i}\lambda \sin(x).$$

The only way for this to be *identically* zero is for $\lambda = 0$.

Note: If $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$ are linearly independent, with

$$\lambda_1 \boldsymbol{x}_1 + \cdots \lambda_r \boldsymbol{x}_r = 0,$$

and $\lambda_1 \neq 0$ (say), then

$$\boldsymbol{x}_1 = -\frac{1}{\lambda_1} \left(\lambda_2 \boldsymbol{x}_2 + \cdots \lambda_r \boldsymbol{x}_r \right).$$

Thus, $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$ **are linearly dependent iff one of them is a linear combination of the others.**

Note: Let $\boldsymbol{x} \neq 0$ be a vector in $\mathcal{H}$. Then the pair $\{0, \boldsymbol{x}\}$ are linearly dependent, because

$$1.0 + 0.\boldsymbol{x} = 0.$$

Thus, a list of linearly independent vectors can never contain the zero vector.

Definition 3.5 *A sequence of vectors $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$ in a real vector space $\mathcal{H}$ is called a* **basis** *for $\mathcal{H}$ if,*

1. *$\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$ are linearly independent;*
2. *$\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$ span $\mathcal{H}$.*

Thus, given a basis $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$, we can write a vector $\boldsymbol{x} \in \mathcal{H}$ as

$$\boldsymbol{x} = \alpha_1 \boldsymbol{x}_1 + \cdots \alpha_n \boldsymbol{x}_n, \qquad \alpha_1, \cdots \alpha_n \in \mathbb{C}.$$

The α_i's are called the **coordinates** of the vector.

Definition 3.6 *Let $\mathcal{H}$ be a complex vector space spanned by a finite number of vectors. The minimal number of vectors required to span the vector is called the* **dimsension** *of the space. The number of elements in a basis is equal to the dimension of the space.*

A vector spaces that is spanned by a finite number of vectors is called **finite-dimensional vector spaces**. Examples:

1. The vectors

$$\begin{aligned} \boldsymbol{e}_1 &= (1,0,0), \\ \boldsymbol{e}_2 &= (0,1,0), \\ \boldsymbol{e}_3 &= (0,0,1) \end{aligned}$$

   are a basis for $\mathbb{C}^3$. They certainly span $\mathbb{C}^3$:

$$\begin{aligned} \boldsymbol{x} &= (z_1, z_2, z_3), \\ &= z_1 \boldsymbol{e}_1 + z_2 \boldsymbol{e}_2 + z_3 \boldsymbol{e}_3. \end{aligned}$$

They are also linearly independent:

$$\begin{aligned} a\boldsymbol{e}_1 + b\boldsymbol{e}_2 + c\boldsymbol{e}_3 &= 0, \\ (a, b, c) &= 0, \\ a = b = c &= 0. \end{aligned}$$

2. The functions $\mathrm{e}^{\mathrm{i}x}$ and $\mathrm{e}^{-\mathrm{i}x}$ form a basis for the solution space of the differential equation

$$\frac{d^2u}{dx^2} + u = 0.$$

3. The $m \times n$ matrices

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix},$$

form a basis for $\mathbb{C}^{m\times n}$ as a complex vector space.

Chapter 4

Scalar products

Reading material for this chapter: Simms 211

4.1 The definition

Definition 4.1 *Let $\mathcal{H}$ be a complex vector space. A scalar product on $\mathcal{H}$ is a map*

$$\mathcal{H} \times \mathcal{H} \to \mathbb{C},$$
$$(\boldsymbol{x}, \boldsymbol{y}) \to \langle \boldsymbol{x} | \boldsymbol{y} \rangle,$$

that is conjugate-linear and conjugate-symmetric:

1. $\langle \lambda \boldsymbol{x} + \mu \boldsymbol{y} | \boldsymbol{z} \rangle = \lambda^* \langle \boldsymbol{x} | \boldsymbol{z} \rangle + \mu^* \langle \boldsymbol{y} | \boldsymbol{z} \rangle$,
2. $\langle \boldsymbol{x} | \lambda \boldsymbol{y} + \mu \boldsymbol{z} \rangle = \lambda \langle \boldsymbol{x} | \boldsymbol{y} \rangle + \mu \langle \boldsymbol{x} | \boldsymbol{z} \rangle$,
3. $\langle \boldsymbol{x} | \boldsymbol{y} \rangle = \langle \boldsymbol{y} | \boldsymbol{x} \rangle^*$,

for all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathcal{H}$ and $\lambda, \mu \in \mathbb{C}$.

4.2 The dot product on $\mathbb{C}^n$

Consider the usual basis on $\mathbb{C}^n$:

$$\begin{aligned} \boldsymbol{e}_1 &= (1, 0, \cdots, 0), \\ \boldsymbol{e}_2 &= (0, 1, \cdots, 0), \\ \vdots &= \vdots, \\ \boldsymbol{e}_n &= (0, 0, \cdots, 1). \end{aligned}$$

Define the **dot product** of two basis vectors:

$$\langle \boldsymbol{e}_i | \boldsymbol{e}_j \rangle = \delta_{ij}$$

where δ_{ij} is the Kronecker delta. Extend this definition by linearity two arbitrary vectors in $\mathbb{C}^n$:

$$\begin{aligned} \boldsymbol{a} &= a_1\boldsymbol{e}_1 + \cdots + a_n\boldsymbol{e}_n, \\ \boldsymbol{b} &= b_1\boldsymbol{e}_1 + \cdots + b_n\boldsymbol{e}_n, \\ \langle \boldsymbol{a}|\boldsymbol{b}\rangle &= \langle a_1\boldsymbol{e}_1 + \cdots + a_n\boldsymbol{e}_n | b_1\boldsymbol{e}_1 + \cdots + b_n\boldsymbol{e}_n\rangle, \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i^*\delta_{ij}b_j, \\ &= a_1^*b_1 + \cdots a_n^*b_n. \end{aligned}$$

Note: $\langle \boldsymbol{a}|\boldsymbol{b}\rangle = \sum_i a_i^* b_i$, and

$$\langle \boldsymbol{b}|\boldsymbol{a}\rangle = \sum_i a_i b_i^* = \left(\sum_i a_i^* b_i\right)^* = \langle \boldsymbol{a}|\boldsymbol{b}\rangle^*.$$

Moreover, for a vector $\boldsymbol{a} \in \mathbb{C}^n$, we define its norm, $|\boldsymbol{a}|$:

$$\begin{aligned} \boldsymbol{a} &= a_1\boldsymbol{e}_1 + \cdots + a_n\boldsymbol{e}_n, \\ |\boldsymbol{a}| := \sqrt{\langle \boldsymbol{a}|\boldsymbol{a}\rangle} &= \sqrt{|a_1|^2 + \cdots + |a_n|^2}. \end{aligned}$$

Theorem 4.1 *The dot product on $\mathbb{C}^n$ satisfies the Cauchy–Schwartz inequality:*

$$|\langle \boldsymbol{a}|\boldsymbol{b}\rangle| \leq |\boldsymbol{a}||\boldsymbol{b}|.$$

Proof: Consider

$$F(x) := \langle x\mathrm{e}^{\mathrm{i}\theta}\boldsymbol{a} + \boldsymbol{b} | x\mathrm{e}^{\mathrm{i}\theta}\boldsymbol{a} + \boldsymbol{b}\rangle,$$

where x is a **real** variable and θ is an arbitrary parameter which we shall fix. Since $\langle \boldsymbol{a}|\boldsymbol{a}\rangle \geq 0$ for all $\boldsymbol{a} \in \mathbb{C}^n$, we have $F(x) \geq 0$, for all x real. We have

$$\begin{aligned} F(x) &= x^2|\boldsymbol{a}|^2 + x\mathrm{e}^{-\mathrm{i}\theta}\langle \boldsymbol{a}|\boldsymbol{b}\rangle + x\mathrm{e}^{\mathrm{i}\theta}\langle \boldsymbol{b}|\boldsymbol{a}\rangle + |\boldsymbol{b}|^2, \\ &= x^2|\boldsymbol{a}|^2 + x\mathrm{e}^{-\mathrm{i}\theta}\langle \boldsymbol{a}|\boldsymbol{b}\rangle + x\mathrm{e}^{\mathrm{i}\theta}\left(\langle \boldsymbol{a}|\boldsymbol{b}\rangle\right)^* + |\boldsymbol{b}|^2, \end{aligned}$$

But θ is arbitrary. We choose it such that

$$\langle \boldsymbol{a}|\boldsymbol{b}\rangle = |\langle \boldsymbol{a}|\boldsymbol{b}\rangle|\, \mathrm{e}^{\mathrm{i}\theta}.$$

Hence,

$$\begin{aligned} F(x) &= x^2|\boldsymbol{a}|^2 + x\mathrm{e}^{-\mathrm{i}\theta}\left(|\langle \boldsymbol{a}|\boldsymbol{b}\rangle|\, \mathrm{e}^{\mathrm{i}\theta}\right) + \mathrm{e}^{\mathrm{i}\theta}\left(\left(|\langle \boldsymbol{a}|\boldsymbol{b}\rangle|\, \mathrm{e}^{\mathrm{i}\theta}\right)\right)^* + |\boldsymbol{b}|^2, \\ &= x^2|\boldsymbol{a}|^2 + 2x\,|\langle \boldsymbol{a}|\boldsymbol{b}\rangle| + |\boldsymbol{b}|^2. \end{aligned}$$

This is a quadratic function in x, with real coefficients, and with roots

$$x_\pm = \frac{|\langle \boldsymbol{a}|\boldsymbol{b}\rangle| \pm \sqrt{|\langle \boldsymbol{a}|\boldsymbol{b}\rangle|^2 - |\boldsymbol{a}|^2|\boldsymbol{b}|^2}}{|\boldsymbol{a}|^2}.$$

But $F(x) \geq 0$, the quadratic function has at most one real root, so

$$|\langle \boldsymbol{a}|\boldsymbol{b}\rangle|^2 - |\boldsymbol{a}|^2|\boldsymbol{b}|^2 \leq 0,$$

or

$$|\langle \boldsymbol{a}|\boldsymbol{b}\rangle| \leq |\boldsymbol{a}||\boldsymbol{b}|,$$

as required.

Note: The Cauchy–Schwartz inequality is true for any scalar product with the positive-definite property $\langle \boldsymbol{x}|\boldsymbol{x}\rangle > 0$ for $\boldsymbol{x} \neq 0$.

Definition 4.2 *Since* $|\langle \boldsymbol{a}|\boldsymbol{b}\rangle| \leq |\boldsymbol{a}||\boldsymbol{b}|$*, we define the* **angle between vectors** $\boldsymbol{a}$ *and* $\boldsymbol{b}$ *(up to a sign):*

$$|\cos\theta| = \frac{|\langle \boldsymbol{a}|\boldsymbol{b}\rangle|}{|\boldsymbol{a}||\boldsymbol{b}|}.$$

Definition 4.3 *Two vectors are* **orthogonal** *if the angle between them is zero:*

$$\langle \boldsymbol{a}|\boldsymbol{b}\rangle = 0.$$

4.3 Spaces of functions

For a Lebesgue-measurable set $\Omega \subset \mathbb{R}$, consider the set $\mathcal{H}_\Omega$ of all complex-valued measurable functions,

$$\mathcal{H}_\Omega = \{f | f : (\Omega \subset \mathbb{R}) \to \mathbb{C}\}.$$

This is a vector space, with pointwise operations of addition and scalar multiplication.

Definition 4.4 *The set*

$$L^2(\Omega) = \left\{ f \in \mathcal{H}_\Omega \middle| \int_\Omega |f(x)|^2 \mathrm{d}x < \infty \right\}$$

is a vector subspace of $\mathcal{H}_\Omega$ called the space of **square-integrable functions**.

Theorem 4.2 *The map*

$$\langle \cdot | \cdot \rangle : L^2(\Omega) \times L^2(\Omega) \to \mathbb{C},$$
$$(f, g) \to \int_\Omega f^*(x) g(x) \mathrm{d}x$$

is a scalar product on the vector space $L^2(\Omega)$.

The proof is easy: all you do is show biconjugate-linearity, e.g.

$$\begin{aligned} (\lambda f + \mu g, h) &= \lambda^* (g, h) + \mu^* (g, h), \\ (f, \lambda g + \mu h) &= \lambda (f, g) + \mu (f, h), \end{aligned}$$

for functions $f, g, h \in L^2(\Omega)$ and scalars λ and μ.

Definition 4.5 *Let $f \in L^2(\Omega)$. Then the* **norm of the function** *f is denoted by $\|f\|_2$, and is defined by*

$$\|f\|_2^2 := \langle f | f \rangle = \int_\Omega |f(x)|^2 \mathrm{d}x.$$

Definition 4.6 *Let $f, g \in L^2(\Omega)$. These functions are* **orthogonal** *if*

$$\langle f | g \rangle = \int_\Omega f^*(x) g(x) \mathrm{d}x = 0.$$

Example: Let $\Omega = [-\pi, \pi]$. The length of the function $\mathrm{e}^{\mathrm{i}x}$ is given by

$$\|\mathrm{e}^{\mathrm{i}x}\|_2^2 = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}x} \mathrm{e}^{-\mathrm{i}x} \mathrm{d}x = \int_{-\pi}^{\pi} \mathrm{d}x = 2\pi.$$

The functions $\mathrm{e}^{\mathrm{i}x}$ and $\mathrm{e}^{-\mathrm{i}x}$ are orthogonal because

$$\langle \mathrm{e}^{\mathrm{i}x} | \mathrm{e}^{-\mathrm{i}x} \rangle = \int_{-\pi}^{\pi} \mathrm{e}^{-2\mathrm{i}x} = \frac{-1}{2\mathrm{i}} e^{-2\mathrm{i}x} \Big|_{-\pi}^{\pi} = 0.$$

Now we can define what a Hilbert space is:

Definition 4.7 *A* **Hilbert space** $\mathcal{H}$ *is a complex vector space endowed with a positive-definite scalar product.*

Moreover, the space is a complete metric space with respect to the norm induced by the scalar product.

The second part of the definition is not important for our purposes, and already follows from the first part for finite-dimensional vector spaces. We also have the following definition and theorems that are not important for our purposes:

Definition 4.8 (Denseness, Separability) *Let* $\mathcal{H}$ *be a Hilbert space with norm* $\|\cdot\|$*. A subset* $D \subset \mathcal{H}$ *is* **dense** *in* $\mathcal{H}$ *if for each* $\boldsymbol{x} \in \mathcal{H}$ *and each* $\epsilon > 0$*, there exists* $\boldsymbol{d} \in D$ *such that* $\|\boldsymbol{d} - \boldsymbol{x}\| < \epsilon$*. The Hilbert space* $\mathcal{H}$ *is called* **separable** *if it contains a countable dense set.*

Also,

Theorem 4.3 *A Hilbert space is separable if and only if it has an orthonormal basis with a countable number of elements.*

Chapter 5

Linear forms and duality

Reading material for this chapter: Simms211

5.1 The definition

Definition 5.1 *Let $\mathcal{H}$ be a complex vector space. A linear form f is a map*

$$\begin{aligned} f : \mathcal{H} &\rightarrow \mathbb{C}, \\ \boldsymbol{x} &\rightarrow f \cdot \boldsymbol{x}, \end{aligned}$$

that satisfies the following linearity properties:

1. $f \cdot (\lambda \boldsymbol{x} + \mu \boldsymbol{y}) = \lambda (f \cdot \boldsymbol{x}) + \mu (f \cdot \boldsymbol{y})$,
2. $(\lambda f + \mu g) \cdot \boldsymbol{x} = \lambda (f \cdot \boldsymbol{x}) + \mu (g \cdot \boldsymbol{y})$,

for all linear forms f and g, vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, and scalars λ and μ.

5.2 Coordinate functions

Let $\mathcal{H}_n$ be a finite-dimensional vector space of dimension n, with basis $\boldsymbol{b}_1, \cdots \boldsymbol{b}_n$. Thus, for each vector $\boldsymbol{x} \in \mathcal{H}_n$, we have

$$\boldsymbol{x} = \sum_{i=1}^{n} \lambda_i \boldsymbol{b}_i, \qquad \lambda_i \in \mathbb{C}.$$

Define the i^{th} coordinate map f_i by its action on $\boldsymbol{x}$:

$$\begin{aligned} f_i : \mathcal{H}_n &\rightarrow \mathbb{C}, \\ f_i \cdot \boldsymbol{x} &= \lambda_i. \end{aligned}$$

This map is a linear form. We can write

$$\boldsymbol{x} = \sum_{i=1}^{n} (f_i \cdot \boldsymbol{x})\, \boldsymbol{b}_i$$

(note that $f_i \cdot \boldsymbol{b}_j = \delta_{ij}$). We can also define a new vector space of functions:

$$\mathcal{H}_n^* = \mathcal{S}\left(f_1, \cdots, f_n\right).$$

The coordinate maps f_i are linearly independent. For, consider the map

$$\alpha f_i + \beta f_j,$$

where α and β are scalars and $i \neq j$. The addition of maps is defined in a pointwise way. Thus, let us examine

$$(\alpha f_i + \beta f_j) \cdot \boldsymbol{x} = \alpha\,(f_i \cdot \boldsymbol{x}) + \beta\,(f_j \cdot \boldsymbol{x}) = \alpha\lambda_i + \beta\lambda_j.$$

The only way for this to be identically zero (for all possible values of λ_i and λ_j) is for α and β both to be zero. Thus, the coordinate maps $\{f_1, \cdots, f_n\}$ are linearly independent, and $\mathcal{H}_n^*$ is n-dimensional as a complex vector space. This motivates the following definition:[1]

Definition 5.2 *The space $\mathcal{H}_n^*$ is called the* **dual space to** $\mathcal{H}_n$.

Example: Let

$$\boldsymbol{x} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{C}^2.$$

Consider the vector

$$f = \left(w_1^*, w_2^*\right).$$

Taking the matrix product of these two elements, we have

$$f\boldsymbol{x} = w_1^* z_1 + w_2^* z_2.$$

[1] The dual space is also where mathematicians go to sort out their differences in a violent way. A bad joke, I know.

Thus, f is a linear form on $\mathbb{C}^2$.

The coordinate functions with respect to the usual basis

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

are

$$f_1 = (1, 0), \qquad f_2 = (0, 1).$$

Hence, given the vector $\boldsymbol{x}$,

$$z_1 = f_1 \cdot \boldsymbol{x}, \qquad z_2 = f_2 \cdot \boldsymbol{y},$$

and

$$\boldsymbol{x} = (f_1 \cdot \boldsymbol{x})\boldsymbol{e}_1 + (f_2 \cdot \boldsymbol{x})\boldsymbol{e}_2.$$

5.3 A special scalar product

There is a bijective map between $\mathcal{H}_n$ and $\mathcal{H}_n^*$. For, let $\boldsymbol{x} \in \mathcal{H}_n$, such that

$$\boldsymbol{x} = \sum_{i=1}^{n} \lambda_i \boldsymbol{b}_i.$$

Then, we can write down a corresponding linear form:

$$f_x = \sum_{i=1}^{n} \lambda_i^* f_i,$$

where f_i is the i^{th} coordinate map, $f_i \cdot \boldsymbol{x} = \lambda_i$. Thus, for each $\boldsymbol{x}$ there is an f_x, and for each f_x there is a $\boldsymbol{x}$. Moreover, let us take

$$\begin{aligned} f_x \cdot \boldsymbol{x} &= \left(\sum_{i=1}^{n} \lambda_i^* f_i \right) \left(\sum_{i=1}^{n} \lambda_i \boldsymbol{b}_j \right), \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i^* \lambda_j f_i \boldsymbol{b}_j, \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i^* \lambda_j \delta_{ij}, \\ &= \sum_{i=1}^{n} |\lambda_i|^2, \end{aligned}$$

which is suggestive of the norm on $\mathbb{C}^n$!!! Indeed, it suggests a recipe for constructing a scalar product on any finite-dimensional vector space $\mathcal{H}_n$.

- Choose a basis for $\mathcal{H}_n$, $\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n\}$, say.
- Write vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ as $\boldsymbol{x} = \sum_i \lambda_i \boldsymbol{b}_i$ and $\boldsymbol{y} = \sum_i \mu_i \boldsymbol{b}_j$.
- Identify the dual-space elements $f_x = \sum_i \lambda_i^* f_i$ and $f_y = \sum_i \mu_i^* f_i$.
- Define the scalar product of $\boldsymbol{x}$ and $\boldsymbol{y}$:

$$\langle \boldsymbol{x} | \boldsymbol{y} \rangle := f_x \cdot \boldsymbol{y} = \sum_{i=1}^{n} \lambda_i^* \mu_i.$$

Because the dual-space scalar product is 'special', we introduce some special notation:

- The vector $\boldsymbol{y} \in \mathcal{H}_n$ will be re-written as $|y\rangle$ and called a 'ket';
- Similarly, the vector $\boldsymbol{x} \in \mathcal{H}_n$ is written as $|x\rangle$. Its corresponding dual element in $\mathcal{H}_n^*$ will be re-written as $\langle x|$ and called a 'bra'.
- The scalar product constructed by uniting $\langle x|$ with $|y\rangle$ will be written in standard form as

$$\langle x | y \rangle.$$

  Thus, the 'bra' and 'ket' are united into one 'bracket'. Where the missing 'c' has gone is a mystery yet to be solved by quantum mechanics.

Example: Take $\mathbb{C}^2$ again. Form the vector

$$\boldsymbol{x} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{C}^2.$$

Its dual element is

$$f_x = (z_1^*, z_2^*)\,,$$

or, treating $\boldsymbol{x}$ as a 2×1 matrix, $f_x = \boldsymbol{x}^{*T}$. Hence,

$$f_x \cdot \boldsymbol{x} = (z_1^*, z_2^*) \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = |z_1|^2 + |z_2|^2,$$

which is the usual dot product on $\mathbb{C}^2$.

Remarkably, the prescription for creating the natural pairing is independent of the basis $\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n\}$ used to formulate the scalar product, because of the following theorem:

Theorem 5.1 *Let $\{\boldsymbol{a}_i\}_{i=1}^n$ and $\{\boldsymbol{b}_i\}_{i=1}^n$ be two bases for $\mathbb{C}^n$, connected by a unitary transformation,*

$$\boldsymbol{b}_i = \sum_{j=1}^n Q_{ji}\boldsymbol{a}_j, \qquad Q^\dagger Q = \mathbb{I}, \qquad \left(Q^\dagger\right)_{ij} = Q^*_{ji}.$$

Then, the natural scalar product is the same in the $\boldsymbol{a}$- and $\boldsymbol{b}$-bases.

Before proving this theorem, I want to admit that it seems weird. But consider the following analogous statement for real vector spaces:

Let $\{\boldsymbol{a}_i\}_{i=1}^n$ and $\{\boldsymbol{b}_i\}_{i=1}^n$ be two orthonormal bases for $\mathbb{R}^n$, connected by a rotation,

$$\boldsymbol{b}_i = \sum_{j=1}^n R_{ji}\boldsymbol{a}_j, \qquad R^T T = \mathbb{I}.$$

Then, the usual dot product is the same in both bases:

$$\begin{aligned}\boldsymbol{x}\cdot\boldsymbol{x} = (\lambda_1\boldsymbol{b}_1 + \cdots \lambda_n\boldsymbol{b}_n)\cdot(\lambda_1\boldsymbol{b}_1 + \cdots \lambda_n\boldsymbol{b}_n) \\ = \left(\widetilde{\lambda_1}\boldsymbol{a}_1 + \cdots + \widetilde{\lambda_n}\boldsymbol{a}_n\right)\cdot\left(\widetilde{\lambda_1}\boldsymbol{a}_1 + \cdots + \widetilde{\lambda_n}\boldsymbol{a}_n\right).\end{aligned}$$

Now we prove the theorem:

$$\begin{aligned}\boldsymbol{x} &= \sum_{i=1}^n \lambda_i\boldsymbol{b}_i, \\ &= \sum_{i=1}^n \lambda_i\left(\sum_{j=1}^n Q_{ji}\boldsymbol{a}_j\right), \\ &= \sum_{j=1}^n \left(\sum_{i=1}^n Q_{ji}\lambda_i\right)\boldsymbol{a}_j, \\ &= \sum_{j=1}^n \widetilde{\lambda}_j\boldsymbol{a}_j, \qquad \widetilde{\lambda}_j = \sum_{k=1}^n Q_{jk}\lambda_k.\end{aligned}$$

Similarly, $\boldsymbol{y} = \sum_{i=1}^{n} \widetilde{\mu}_j \boldsymbol{a}_j$. Therefore, in the $\boldsymbol{a}$-basis,

$$\begin{aligned}
\langle x|y\rangle_a &= \sum_{j=1}^{n} \widetilde{\lambda}_j^{\,*} \widetilde{\mu}_j, \\
&= \sum_{j=1}^{n} \left(\sum_{k=1}^{n} Q_{jk} \lambda_k \right)^* \left(\sum_{\ell=1}^{n} Q_{j\ell} \mu_\ell \right), \\
&= \sum_{j=1}^{n} \left(\sum_{k=1}^{n} \sum_{\ell=1}^{n} Q_{jk}^* Q_{j\ell} \right) \lambda_i^* \mu_k,
\end{aligned}$$

$$\begin{aligned}
\langle x|y\rangle_a &= \sum_{k=1}^{n} \sum_{\ell=1}^{n} \left(\sum_{j=1}^{n} (Q^{*T})_{kj} Q_{j\ell} \right) \lambda_i^* \mu_k, \\
&= \sum_{k=1}^{n} \sum_{\ell=1}^{n} \delta_{k\ell} \lambda_i^* \mu_k, \\
&= \sum_{k=1}^{n} \lambda_k^* \mu_k, \\
&= \langle x|y\rangle_b.
\end{aligned}$$

Notes:

- From this proof, it follows that here is nothing arbitrary about the scalar product just defined.
- Because it has the positive-definite property, it makes $\mathcal{H}_n$ into a Hilbert space.
- The proof relies on the transformation matrix Q being unitary; we discuss this in more detail now.

5.4 Unitary matrices

Consider a ket $|x\rangle$ in $\mathbb{C}^n$. Let's act on the ket with a unitary matrix Q:

$$|x\rangle \to Q|x\rangle.$$

We know from the example in $\mathbb{C}^2$ that the transformed bra is

$$\begin{aligned}
(Q|x\rangle)^{*T} &= (|x\rangle)^{*T} Q^{*T}, \\
&= \langle x|Q^\dagger.
\end{aligned}$$

Let's take the norm of the transformed variable:

$$\langle x|Q^\dagger Q|x\rangle = \langle x|\mathbb{I}|x\rangle = \langle x|x\rangle.$$

Thus, **unitary transformations** preserve the norm of vectors.

This is very similar to rotations in $\mathbb{R}^n$. Consider a vector $\boldsymbol{x} \in \mathbb{R}^n$. If we rotate the vector, we act on it with a real symmetric matrix R, $R^T R = \mathbb{I}$. The norm of the vector is

$$\boldsymbol{x} \cdot \boldsymbol{x} = \boldsymbol{x}^T \boldsymbol{x},$$

and the norm of the rotated vector is

$$(R\boldsymbol{x})^T (R\boldsymbol{x}) = \boldsymbol{x}^T R^T R \boldsymbol{x} = \boldsymbol{x}^T \mathbb{I} \boldsymbol{x}.$$

The quantity $\boldsymbol{x} \cdot \boldsymbol{x}$ is therefore a **scalar**, and we formulate physical theories based on such scalars. Thus, if we are working with $\mathbb{C}^n$, instead of $\mathbb{R}^n$, it is natural to formulate a physical theory based on quantities that are norm-invariant. Natural transformations are therefore those that preserve the norm – or unitary matrices.

Now we can make some more sense of Postulate 3 of Quantum mechanics. Consider two states of $|\phi\rangle$ and $|\psi\rangle$ of a physical system. To characterise the system, we need information about probabilities that certain states are realised. Such information is contained in pairings like

$$\langle \phi|\psi\rangle.$$

A symmetry of the system must not change this information. Thus, consider a unitary operator $\hat{U}$. The pairing

$$\langle \phi|\hat{U}^\dagger \hat{U}|\psi\rangle = \langle \phi|\mathbb{I}|\psi\rangle = \langle \phi|\psi\rangle$$

contains exactly the same information as $\langle \phi|\psi\rangle$. Thus, the system is effectively unchanged when

$$|\phi\rangle \to \hat{U}|\phi\rangle, \qquad |\psi\rangle \to \hat{U}|\psi\rangle.$$

This is a justification of the third postulate.

5.5 On the induced scalar product versus the prescribed one

So far we have worked with a finite-dimensional Hilbert space $\mathcal{H}_n$. This means that there is a definite scalar product that is **prescribed** or given to us. On the other hand, we have described a

process of pairing elements in $\mathcal{H}_n$ with elements in $\mathcal{H}_n^*$ which **induces** a scalar product on $\mathcal{H}_n$. It will be helpful (especially in the infinite-dimensional case) to know that these two scalar products agree. The following results give a condition that guarantees that these two scalar products agree:

Lemma 5.1 *Let $\{|e_i\rangle\}_{i=1}^n$ be a basis for $\mathcal{H}_n$ that is orthonormal with respect to the* **given** *scalar product:*

$$\langle e_i|e_j\rangle = \delta_{ij}$$

Then, the scalar product induced by pairing is the same as the given scalar product:

$$\langle e_i|e_j\rangle = \langle e_i|e_j\rangle_{eP} = \delta_{ij},$$

where the subscript eP here denotes the scalar product got by pairing with respect to the $|e_i\rangle$-basis.

Proof: By definition, $f_i|e_j\rangle = \delta_{ij}$, where f_i is the i^{th} coordinate function with respect to the $|e_i\rangle$-basis. In other words, $\langle e_i|e_j\rangle_{eP} = \delta_{ij}$, and the result is shown.

Theorem 5.2 (Agreement between the prescribed scalar product and the induced one) *Let $\{|e_i\rangle\}_{i=1}^n$ be a basis for $\mathcal{H}_n$ that is orthonormal with respect to the* **given** *scalar product:*

$$\langle e_i|e_j\rangle = \delta_{ij}$$

and let $\{|b\rangle_i\}_{i=1}^n$ be another basis, connected to the $|e_i\rangle$-basis via a unitary transformation:

$$|b_i\rangle = \sum_{j=1}^{n} Q_{ji}|e_j\rangle, \qquad Q^\dagger Q = \mathbb{I}.$$

Then, the scalar product induced by pairing with respect to the $|b_i\rangle$-basis is the same as the given scalar product:

$$\langle b_i|b_j\rangle = \langle b_i|b_j\rangle_{bP} = \langle e_i|e_j\rangle = \delta_{ij},$$

where the subscript bP here denotes the scalar product got by pairing with respect to the $|b_i\rangle$-basis.

Proof: We have

$$\delta_{ij} = \langle b_i|b_j\rangle_{bP},$$

Since the $|e_i\rangle$- and $|b_i\rangle$-bases are connected via a unitary transformation, by Theorem 5.1 we have that

$$\delta_{ij} = \langle b_i|b_j\rangle_{bP} = \langle b_i|b_j\rangle_{eP}$$

and by Lemma 5.1 we get

$$\delta_{ij} = \langle b_i|b_j\rangle_{bP} = \langle b_i|b_j\rangle_{eP} = \langle b_i|b_j\rangle,$$

and the theorem is shown.

5.6 Riesz representation theorem

The correspondence between the Hilbert space and its dual extends to infinite-dimensional spaces, where it is called the **Riesz representation theorem:**

Theorem 5.3 *If $\mathcal{H}$ is a Hilbert space with prescribed scalar product $\langle\cdot|\cdot\rangle$, then for any continuous linear form $f : \mathcal{H} \to \mathbb{C}$, there exists a unique element $|u\rangle \in \mathcal{H}$ such that*

$$f|x\rangle = \langle u|x\rangle, \qquad \forall |x\rangle \in \mathcal{H}.$$

In this way – just as in the finite-dimensional case, an arbitrary linear form can f can be identified with a vector $|u\rangle$, and we would write $f \equiv f_u = \langle u|$. However, this theorem relies for its proof on the topological properties of the Hilbert space induced by the prescribed scalar product, and we cannot in this case simply start with a pairing operation and construct a scalar product – we must proceed in the reverse order. These are technical points whose elucidation is well beyond the scope of this module.

Chapter 6

Operators

Reading material for this chapter: Simms211; LandauLifshitz, Chapter 1

6.1 Linear operators

Definition 6.1 *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be complex vector spaces. A linear operator $\hat{\mathrm{A}}$ is a map*

$$\begin{aligned} \hat{\mathrm{A}} : \mathcal{H}_1 &\rightarrow \mathcal{H}_2, \\ |x\rangle &\rightarrow \hat{\mathrm{A}}|x\rangle \end{aligned}$$

such that

$$\begin{aligned} \hat{\mathrm{A}}\left(|x\rangle + |y\rangle\right) &= \hat{\mathrm{A}}|x\rangle + \hat{\mathrm{A}}|y\rangle, \\ \hat{\mathrm{A}}(\lambda|x\rangle) &= \lambda\hat{\mathrm{A}}|x\rangle, \end{aligned}$$

for all $|x\rangle, |y\rangle \in \mathcal{H}_1$ and $\lambda \in \mathbb{C}$. Examples:

- An $n \times n$ matrix is a linear operator on $\mathbb{C}^n$, and maps $\mathbb{C}^n$ to itself.
- Let $C^r(\Omega)$ be the space of all complex-valued functions of a single real variable that are r-times continuously differentiable on the open interval $\Omega \subset \mathbb{R}$. Then the usual derivative operation is a linear operator:

$$\begin{aligned} d/dx : C^r(\Omega) &\rightarrow C^{r-1}(\Omega), \\ f(x) &\rightarrow (df/dx), \end{aligned}$$

since

$$\begin{aligned}(d/dx)\left[f(x)+g(x)\right] &= (df/dx)+(dg/dx),\\ (d/dx)\left[\lambda f(x)\right] &= \lambda(df/dx),\end{aligned}$$

for all $f(x), g(x) \in C^r(\Omega)$ and $\lambda \in \mathbb{C}$.

Definition 6.2 *Let $\hat{\mathrm{A}}$ be a linear operator that maps the Hilbert space $\mathcal{H}$ to itself. The* **adjoint** *of $\hat{\mathrm{A}}$, $\hat{\mathrm{A}}^\dagger$ is an operator acting on $\mathcal{H}^*$, defined as follows:*

- *Identify $|x\rangle$ and $\hat{\mathrm{A}}|x\rangle$ in $\mathcal{H}$.*
- *Pair $\hat{\mathrm{A}}|x\rangle$ with an element $\langle y|$ in the dual space.*
- *Call $\langle y| := \langle x|\hat{\mathrm{A}}^\dagger$.*

Example: Let $\mathcal{H} = \mathbb{C}^n$ with the usual basis $\{\boldsymbol{e}_i\}_{i=1}^n$. Consider a matrix $A \in \mathbb{C}^{n\times n}$. This can be made into an operator on $\mathcal{H}$ by defining the action of A on the usual basis elements:

$$\hat{\mathrm{A}}\boldsymbol{e}_i := \sum_{j=1}^{n} A_{ji}\boldsymbol{e}_j, \qquad A_{ij} \in \mathbb{C},$$

(NOTE THE ORDER!) This can be extended by linearity to the whole space:

$$\begin{aligned}\boldsymbol{x} &= \sum_{i=1}^{n}\lambda_i\boldsymbol{e}_i,\\ \hat{\mathrm{A}}\boldsymbol{x} &= \hat{\mathrm{A}}\left(\sum_{i=1}^{n}\lambda_i\boldsymbol{e}_i\right),\\ &= \sum_{i=1}^{n}\lambda_i\left(\hat{\mathrm{A}}\boldsymbol{e}_i\right),\\ &= \sum_{i=1}^{n}\lambda_i\left(\sum_{j=1}^{n}A_{ji}\boldsymbol{e}_j\right),\\ &= \sum_{j=1}^{n}\left(\sum_{i=1}^{n}A_{ji}\lambda_i\right)\boldsymbol{e}_j,\\ &= \sum_{j=1}^{n}\widetilde{\lambda}_j\boldsymbol{e}_j = \sum_{i=1}^{n}\widetilde{\lambda}_i\boldsymbol{e}_i.\end{aligned}$$

In 'bra'-'ket' notation,

$$\begin{aligned}
|x\rangle &= \sum_{i=1}^{n} \lambda_i |e_i\rangle, \\
\hat{\mathrm{A}}|x\rangle &= \hat{\mathrm{A}} \left(\sum_{i=1}^{n} \lambda_i |e_i\rangle \right), \\
&= \sum_{i=1}^{n} \lambda_i \left(\hat{\mathrm{A}} |e_i\rangle \right), \\
&= \sum_{i=1}^{n} \lambda_i \left(\sum_{j=1}^{n} A_{ji} |e_j\rangle \right), \\
&= \sum_{j=1}^{n} \left(\sum_{i=1}^{n} A_{ji} \lambda_i \right) |e_j\rangle, \\
&= \sum_{j=1}^{n} \widetilde{\lambda}_j |e_j\rangle = \sum_{i=1}^{n} \widetilde{\lambda}_i |e_i\rangle.
\end{aligned}$$

Note also,

$$\begin{aligned}
\hat{\mathrm{A}}|e_i\rangle &= \sum_{j=1}^{n} A_{ji} |e_j\rangle, \\
\langle e_k | \hat{\mathrm{A}} | e_i \rangle &= A_{ki}
\end{aligned}$$

(NOTE THE ORDER!!). We call A_{ki} the **components** of the operator $\hat{\mathrm{A}}$ w.r.t. the usual basis.

Now let us work out what the action of the adjoint is on basis elements:

$$\hat{\mathrm{A}}|e_i\rangle = \sum_{j=1}^{n} A_{ji} |e_j\rangle \sim \sum_{j=1}^{n} A^*_{ji} \langle e_j| := \langle e_i | \hat{\mathrm{A}}^\dagger.$$

To work out the components of $\langle e_i | A^\dagger$, we pair it with $|e_k\rangle$:

$$\begin{aligned}
\langle e_i | A^\dagger &= \sum_{j=1}^{n} A^*_{ji} \langle e_j|, \\
\langle e_i | A^\dagger | e_k \rangle &= \left(\sum_{j=1}^{n} A^*_{ji} \langle e_j| \right) |e_k\rangle, \\
&= A^*_{ki}, \\
&= \left(A^{T*} \right)_{ik}.
\end{aligned}$$

In conclusion, we have the following identifications

- $\hat{\mathrm{A}} \to A$, where A is a matrix with components

$$A_{ij} = \langle e_i | \hat{\mathrm{A}} | e_j \rangle,$$

- $\hat{\mathrm{A}}^\dagger \to A^\dagger$, where $A^\dagger$ is the matrix

$$A^\dagger = A^{T*} = A^{*T}.$$

Definition 6.3 *A matrix (or an operator)* $\hat{\mathrm{A}}$ *is called* **Hermitian** *if*

$$\hat{\mathrm{A}}^\dagger = \hat{\mathrm{A}}.$$

Definition 6.4 *A matrix (or an operator)* $\hat{U}$ *is called* **unitary** *if*

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger \hat{U} = \mathbb{I}.$$

Example: Consider the matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

These matrices are Hermitian. For example,

$$\sigma_y^\dagger = \begin{pmatrix} 0 & +\mathrm{i} \\ -\mathrm{i} & 0 \end{pmatrix}^T = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} = \sigma_y.$$

They are also unitary. Again,

$$\sigma_y^\dagger \sigma_y = \sigma_y^2 = \begin{pmatrix} 0 & +\mathrm{i} \\ -\mathrm{i} & 0 \end{pmatrix} \begin{pmatrix} 0 & +\mathrm{i} \\ -\mathrm{i} & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

6.2 The spectral theorem

Theorem 6.1 *Let* $\hat{\mathrm{A}}$ *be a Hermitian operator on a Hilbert space* $\mathcal{H}$*. Then the eigenvalues of* $\hat{\mathrm{A}}$ *are necessarily real.*

Proof: Let $|x\rangle$ be an eigenvector of $\hat{\mathrm{A}}$ with eigenvalue λ. By definition,

$$\hat{\mathrm{A}}|x\rangle = \lambda|x\rangle.$$

The adjoint operator acting on the $\langle x|$ is obtained from the duality identification:

$$\langle x|\hat{\mathrm{A}}^\dagger = \lambda^*\langle x|.$$

Pair up both expressions:

$$\begin{aligned}\langle x|\hat{\mathrm{A}}|x\rangle &= \lambda\langle x|x\rangle,\\ \langle x|\hat{\mathrm{A}}^\dagger|x\rangle &= \lambda^*\langle x|x\rangle.\end{aligned}$$

The operator is Hermitian, hence $\hat{\mathrm{A}} = \hat{\mathrm{A}}^\dagger$, and thus

$$\lambda\langle x|x\rangle = \lambda^*\langle x|x\rangle.$$

By definition, an eigenvector is non-zero, hence

$$\lambda = \lambda^*,$$

and $\lambda \in \mathbb{R}$.

Theorem 6.2 *Let $\hat{\mathrm{A}}$ be a Hermitian operator on a Hilbert space $\mathcal{H}$. Then the eigenvectors of $\hat{\mathrm{A}}$ corresponding to distinct eigenvalues are necessarily orthogonal.*

Proof: Consider two distinct eigenvector-eigenvalue pairs:

$$\begin{aligned}\hat{\mathrm{A}}|x\rangle &= \lambda|x\rangle,\\ \hat{\mathrm{A}}|y\rangle &= \mu|y\rangle.\end{aligned}$$

Take the scalar product of the first equation with $|y\rangle$ and the scalar product of the second equation with $|x\rangle$:

$$\begin{aligned}\langle y|\hat{\mathrm{A}}|x\rangle &= \lambda\langle y|x\rangle,\\ \langle x|\hat{\mathrm{A}}|y\rangle &= \mu\langle x|y\rangle.\end{aligned}$$

But

$$\langle y|\hat{\mathrm{A}}|x\rangle = \langle x|\hat{\mathrm{A}}|y\rangle^* = \mu^*\langle x|y\rangle^* = \mu\langle y|x\rangle$$

Hence,

$$\begin{aligned}\langle y|\hat{\mathrm{A}}|x\rangle &= \lambda\langle y|x\rangle,\\ \langle y|\hat{\mathrm{A}}|x\rangle &= \mu\langle y|x\rangle.\end{aligned}$$

Subtracting gives

$$(\lambda-\mu)\langle y|x\rangle = 0,$$

and since $\lambda \neq \mu$, $\langle y|x\rangle = 0$.

Note: These theorems give information about the properties of eigenvalues and eigenvectors of Hermitian operators. However, they do not guarantee that such eigenvectors form a basis for the space. We therefore turn to a theorem that guarantees such an outcome:

Theorem 6.3 (The spectral theorm) *Let $\hat{\mathrm{A}}$ be a Hermitian operator on a finite-dimensional Hilbert space $\mathcal{H}$. Then the eigenvectors of $\hat{\mathrm{A}}$ form an orthogonal basis for $\mathcal{H}$.*

The result extends to infinite-dimensional spaces if the Green's function of $\hat{\mathrm{A}}$ is bounded and continuous.

Conseqeunces of the spectral theorem; the problem of measurement

The spectral theorem is stated without proof but is of fundamental importance to quantum mechanics. Although it is stated only for finite-dimensional Hilbert spaces, it extends to other cases, provided $\hat{\mathrm{A}}$ satisfies certain technical conditions (the examples considered in this module fall into this 'well-behaved' category, namely operators whose Green's function is bounded and continuous).

1. By postulate (4), if we can compute the eigenvalues of $\hat{\mathrm{A}}$, then we can **predict** all possible observed (measured) states of a system with respect to the property $\hat{\mathrm{A}}$.

2. Given an operator $\hat{\mathrm{A}}$ on a Hilbert space $\mathcal{H}$, we formulate the so-called **completeness relation**. Let $\{|x_i\rangle\}$ be the complete orthonormal basis of $\hat{\mathrm{A}}$ from the spectral theorem, $\hat{\mathrm{A}}|x\rangle_i = a_i|x\rangle_i$. It is called complete because **any** vector $|x\rangle \in \mathcal{H}$ can be written as a superposition of basis elements:

   $$|x\rangle = \sum_i \lambda_i |x_i\rangle,$$

   Here the coordinate λ_i is given by pairing the coordinate function ('bra') $\langle x_i|$ with $|x\rangle$:

   $$\lambda_i = \langle x_i|x\rangle.$$

Thus,

$$\begin{aligned} |x\rangle &= \sum_i \lambda_i |x_i\rangle, \\ &= \sum_i \langle x_i|x\rangle |x_i\rangle, \\ &= \sum_i |x_i\rangle\langle x_i|x\rangle. \end{aligned}$$

We now define an operator $\mathcal{I}$ on $\mathcal{H}$:

$$\begin{aligned} \mathcal{I} : \mathcal{H} &\rightarrow \mathcal{H}, \\ |x\rangle &\rightarrow \mathcal{I}|x\rangle := \sum_i |x_i\rangle\langle x_i|x\rangle. \end{aligned}$$

In other words,

$$\mathcal{I} = \sum_i |x_i\rangle\langle x_i|.$$

But

$$\begin{aligned} |x\rangle &= \sum_i |x_i\rangle\langle x_i|x\rangle, \\ &= \mathcal{I}|x\rangle, \end{aligned}$$

hence

$$\mathcal{I} = \mathbb{I},$$

and we have the following **completeness relation**:

$$\mathbb{I} = \sum_i |x_i\rangle\langle x_i|.$$

3. Consider again the observable $\hat{\mathrm{A}}$ with orthonormal basis $\{|x_i\rangle\}$. Suppose that the system is prepared in a state $|y\rangle$. By completeness,

$$|y\rangle = \sum_i \langle x_i|y\rangle |x_i\rangle.$$

By postulate 4, the only outcome of a measurement of property A is an eigenvalue of $\hat{\mathrm{A}}$. Thus, measurement forces the system into an eigenstate,

$$|y\rangle \rightarrow_{\text{measurement}} |x_i\rangle.$$

This is called the **collapse of the wavefunction**. By postulate 4 again, the **probability amplitude** that the measurement forces the system into the eigenstate $|x_i\rangle$ (with eigenvalue a_i) is equal to,

$$\langle x_i|y\rangle,$$

and the probability that the measurement forces the system into the eigenstate $|x_i\rangle$ is equal to

$$\text{Prob}(y \to x_i) = |\langle x_i|y\rangle|^2.$$

If the spectrum is **non-degenerate** (each eigenspace is one-dimensional), then this is the probability that measurement of the observable A yields the value a_i:

$$\text{Prob}(a = a_i) = |\langle x_i|y\rangle|^2.$$

If the spectrum is degenerate, and the eigenvectors $|x_{a1}\rangle, \cdots |x_{ag}\rangle$ are linearly independent and share the common eigenvalue a_i, then

$$\text{Prob}(a = a_i) = |\langle x_{a1}|y\rangle|^2 + \cdots + |\langle x_{ag}|y\rangle|^2$$

(Never add amplitudes for distinguishable final states!).

4. In this interpretation, we may also view $|\langle x_i|y\rangle|^2$ as the probability that the system is in a state $|x_i\rangle$ given it is also in a state $|y\rangle$. This point of view, which has just been given for eigenstates, holds generally: if $|i\rangle$ and $|f\rangle$ are two normed states, we interpret $\langle f|i\rangle$ as the probability amplitude that the system when in the state $|i\rangle$ is also in the state $|f\rangle$.

5. Thus, the reason for the requirement of unit-norm states is clear: The quantity $|\langle x|x\rangle|^2$ is the probability that the system is in a state $|x\rangle$ given that it is in a state $|x\rangle$, which must necessarily be unity.

6. The statement $\langle x|y\rangle = 0$ is also the statement of mutual exclusivity: that it is impossible for the system to be in two mutually exclusive states at once. For example, suppose a particle is measured to have energy E_i. Then it is in a state $|E_i\rangle$. It is impossible for the particle simulatneously to occupy a state with another, diferent energy, E_j. Thus, $\langle E_i|E_j\rangle = 0$.

7. It is still not clear what measurement is. Landau and Lifshitz define it as an interaction between a quantum-mechanical system and a detector that obeys classical physics (Copenhagen interpretation). For example, a current of electrons in a circuit can be measured by an ammeter – a very classical device. Thus, the theory of quantum-mechanical measurement relies for its formulation on the classical limit. This is rather unsatisfactory, but is a bearable oddity.

8. Rather more serious is the silence of quantum mechanics on the actual dynamics of measurement: does the wavefunction collapse instantaneously? If so, how can causality be respected? How can the continuous nature of time changes be respected? Is time continuous at all? Obviously, this leads to much discussion. The standard description of measurement is given by the **Copenhagen interpretation**. Other, self-consistent but bizarre descriptions are possible, such as the **many-worlds interpretation**. Such abstruse discussions are beyond the scope of this course.

Example: Consider three elements enclosed in a sealed box: a sealed container of noxious gas, a radioactive source, and a cat that is alive just before the box is sealed. The radioactive source decays and emits decay products, with probability $1/2$. The sealed container is connected to a device such that the seal is broken when struck with the decay products, thus killing the cat. The cat's wavefunction is given by

$$|\text{cat}\rangle = \tfrac{1}{\sqrt{2}}|\text{alive}\rangle + \tfrac{1}{\sqrt{2}}|\text{dead}\rangle,$$

where the prefactors are chosen such that

$$\begin{aligned}
\langle\text{cat}|\text{cat}\rangle &= \left(\tfrac{1}{\sqrt{2}}\langle\text{alive}| + \tfrac{1}{\sqrt{2}}\langle\text{dead}|\right)\left(\tfrac{1}{\sqrt{2}}|\text{alive}\rangle + \tfrac{1}{\sqrt{2}}|\text{dead}\rangle\right), \\
&= \tfrac{1}{2}\left(\langle\text{alive}|\text{alive}\rangle + \langle\text{alive}|\text{dead}\rangle + \langle\text{dead}|\text{alive}\rangle + \langle\text{dead}|\text{dead}\rangle\right), \\
&= = \tfrac{1}{2}\left(1 + 0 + 0 + 1\right) = 1,
\end{aligned}$$

since the amplitudes $\langle\text{dead}|\text{dead}\rangle$ and $\langle\text{alive}|\text{alive}\rangle$ must be equal to one (the probability that the cat is alive given that it is alive must be 1!). We wait some time and measure the system (shake the box!). This forces the system into an eigenstate. The amplitude for the cat to be alive given it is initially in the mixed state is

$$\langle\text{alive}|\text{cat}\rangle = \tfrac{1}{\sqrt{2}},$$

with probability $1/2$. Similarly, the probability that measurement yields a dead cat is $1/2$. However, until the measurement is made, there cat is regarded as neither alive nor dead – its state is indeterminate.

Chapter 7

Commutation relations; Time evolution and the Schrödinger equation

Reading material for this chapter: Mandl, Chapter 3; Landau-Lifshitz, Chapter 2

7.1 Commutation relations

Let $\hat{\mathrm{A}}$ and $\hat{\mathrm{B}}$ be linear operators on a Hilbert space $\mathcal{H}$, and let $|x\rangle \in \mathcal{H}$. Consider the compositions

$$\hat{\mathrm{A}} \circ \hat{\mathrm{B}}|x\rangle = \hat{\mathrm{A}}\left(\hat{\mathrm{B}}|x\rangle\right) := \hat{\mathrm{A}}\hat{\mathrm{B}}|x\rangle,$$

and

$$\hat{\mathrm{B}} \circ \hat{\mathrm{A}}|x\rangle = \hat{\mathrm{B}}\left(\hat{\mathrm{A}}|x\rangle\right) := \hat{\mathrm{B}}\hat{\mathrm{A}}|x\rangle,$$

There is no reason that these should yield the same answer. We define the **commutator** of $\hat{\mathrm{A}}$ and $\hat{\mathrm{B}}$ on the vector $|x\rangle$ as follows:

$$\left[\hat{\mathrm{A}}, \hat{\mathrm{B}}\right]|x\rangle := \left(\hat{\mathrm{A}}\hat{\mathrm{B}} - \hat{\mathrm{B}}\hat{\mathrm{A}}\right)|x\rangle.$$

This relation exists entirely independently of the vector $|x\rangle$, thus, we consider

$$\left[\hat{\mathrm{A}}, \hat{\mathrm{B}}\right] := \hat{\mathrm{A}}\hat{\mathrm{B}} - \hat{\mathrm{B}}\hat{\mathrm{A}},$$

where the multiplication is regarded as the multiplication of operators. For example, let $\mathcal{H} = \mathbb{C}^2$, and let

$$\hat{\mathrm{A}} = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \hat{\mathrm{B}} = \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}.$$

Then

$$\sigma_x\sigma_y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix} = \mathrm{i}\sigma_z,$$

and

$$\sigma_y\sigma_z = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -\mathrm{i} & 0 \\ 0 & \mathrm{i} \end{pmatrix} = -\mathrm{i}\sigma_z,$$

hence

$$\sigma_x\sigma_y - \sigma_y\sigma_x = 2\mathrm{i}\sigma_z,$$

or

$$[\sigma_x, \sigma_y] = 2\mathrm{i}\sigma_z.$$

We have another important theorem, almost as important as the spectral theorem:

Theorem 7.1 *Let* $\hat{\mathrm{A}}$ *and* $\hat{\mathrm{B}}$ *be two commuting Hermitian operators on a separable Hilbert space* $\mathcal{H}$*, i.e.* $[\hat{\mathrm{A}}, \hat{\mathrm{B}}] = 0$*. Then there exists an orthonormal basis for* $\mathcal{H}$ *whose elements are simultaneous eigenvectors of* $\hat{\mathrm{A}}$ *and* $\hat{\mathrm{B}}$*.*

No proof is given here. Put another way, the theorem says that the commutation relation

$$\left[\hat{\mathrm{A}}, \hat{\mathrm{B}}\right] = 0,$$

implies a basis $\{|x_i\rangle\}$ for $\mathcal{H}$:

$$|x\rangle = \sum_i \lambda_i |x_i\rangle,$$

such that

$$\hat{\mathrm{A}}|x_i\rangle = a_i|x_i\rangle,$$

and such that

$$\hat{\mathrm{B}}|x_i\rangle = b_i|x_i\rangle.$$

Example: Consider again the matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Define the matrix

$$\sigma^2 := \sigma_x^2 + \sigma_y^2 + \sigma_z^2 = 3\mathbb{I}.$$

We have

$$\left[\sigma_x, \sigma^2\right] = 3\left[\sigma_x, \mathbb{I}\right] = 3\left(\sigma_x \mathbb{I} - \mathbb{I}\sigma_x\right) = 3\left(\sigma_x - \sigma_x\right) = 0.$$

Thus, σ_x and σ^2 are simultaneously diagonalisable. The eigenvalues of σ_x are ± 1, with corresponding (orthonormal) eigenvectors

$$|x_-\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \qquad |x_+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

such that

$$\sigma_x |x_-\rangle = (-1)\,|x_-\rangle, \qquad \sigma^2 |x_-\rangle = (+3)\,|x_-\rangle,$$

and such that

$$\sigma_x |x_+\rangle = (+1)\,|x_+\rangle, \qquad \sigma^2 |x_+\rangle = (+3)\,|x_+\rangle.$$

It is not possible to find a simultaneous eigenbasis for $\{\sigma^2, \sigma_x, \sigma_y\}$ because σ_x and σ_y do not commute: $[\sigma_x, \sigma_y] = 2\mathrm{i}\sigma_z$. Note also:

$$\begin{aligned} \langle x_-| &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}^{*T} = \frac{1}{\sqrt{2}}\,(1, -1)\,, \\ \langle x_+| &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}^{*T} = \frac{1}{\sqrt{2}}\,(1, 1)\,, \end{aligned}$$

hence

$$\begin{aligned} \langle x_-|x_-\rangle &= \left[\frac{1}{\sqrt{2}}\,(1,-1)\right]\left[\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right] = \frac{1+1}{2} = 1, \\ \langle x_-|x_+\rangle &= \left[\frac{1}{\sqrt{2}}\,(1,-1)\right]\left[\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ +1 \end{pmatrix}\right] = \frac{1-1}{2} = 0, \\ \langle x_+|x_+\rangle &= \left[\frac{1}{\sqrt{2}}\,(1,1)\right]\left[\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ +1 \end{pmatrix}\right] = \frac{1+1}{2} = 1, \end{aligned}$$

which should give the reader some more familiarity with the 'bras' and 'kets'.

7.2 Time evolution and the Schrödinger equation

Consider a physical system that is invariant under time translation. Postulate 1 says that there is a Hilbert space $\mathcal{H}$ that describes the system, and that vectors in the space describe states of the system. Consider one such state, $|\phi(0)\rangle$, where the 0 denotes the state at time $t = 0$. The system

is invariant under time translation. Therefore, by Postulate 2, the state of the system a short time later is given by

$$|\phi(\Delta t)\rangle = \hat{U}(\Delta t)|\phi(0)\rangle,$$

where $\hat{U}(\Delta t)$ is a unitary operator. We assume that the increment Δt is small. Thus, the unitary operator $\hat{U}(\Delta t)$ can be written as

$$\hat{U}(\Delta t) = \mathbb{I} - \frac{\mathrm{i}\hat{H}\Delta t}{\hbar},$$

where $\hat{H}$ is a **hermitian** operator, $\hat{H}^\dagger = \hat{H}$, since then

$$\begin{aligned}
\hat{U}(\Delta t)^\dagger \hat{U}(\Delta t) &= \left(\mathbb{I} - \frac{\mathrm{i}\hat{H}\Delta t}{\hbar}\right)^\dagger \left(\mathbb{I} - \frac{\mathrm{i}\hat{H}\Delta t}{\hbar}\right), \\
&= \left(\mathbb{I} + \frac{\mathrm{i}\hat{H}\Delta t}{\hbar}\right)\left(\mathbb{I} - \frac{\mathrm{i}\hat{H}\Delta t}{\hbar}\right), \\
&= \mathbb{I} + \frac{\mathrm{i}\hat{H}\Delta t}{\hbar} - \frac{\mathrm{i}\hat{H}\Delta t}{\hbar} + O\left(\Delta t^2\right), \\
&= \mathbb{I} + O\left(\Delta t^2\right),
\end{aligned}$$

which is unitary as $\Delta t \to 0$. The operator $\hat{H}$ is independent of time because the system is assumed to be invariant under time translation. However, this assumption is not necessary.

Now consider the system at a much later time $t_n = n\Delta t$:

$$\begin{aligned}
|\phi(t_n)\rangle &= \Pi \left(\mathbb{I} - \frac{\mathrm{i}\hat{H}\Delta t}{\hbar}\right)^n |\phi(0)\rangle, \\
&= \mathrm{e}^{-\mathrm{i}\hat{H}n\Delta t/\hbar}|\phi(0)\rangle, \\
&= \mathrm{e}^{-\mathrm{i}\hat{H}t_n/\hbar}|\phi(0)\rangle.
\end{aligned}$$

Letting $\Delta t \to 0$ and $n \to \infty$ keeping $t := n\Delta t$ fixed, we obtain

$$|\phi(t)\rangle = \mathrm{e}^{-\mathrm{i}\hat{H}t/\hbar}|\phi(0)\rangle$$

Next, we differentiate the state $|\phi(t)\rangle$ to see how it evolves in time:

$$\begin{aligned}
\frac{\partial}{\partial t}|\phi(t)\rangle &= \lim_{\delta t\to 0}\left(\frac{\mathrm{e}^{-\mathrm{i}\hat{H}(t+\delta t)/\hbar}-\mathrm{e}^{-\mathrm{i}\hat{H}/\hbar}}{\delta t}\right)|\phi(0)\rangle,\\
&= \lim_{\delta t\to 0}\left[\frac{\sum_{n=0}^{\infty}\left(\frac{-\mathrm{i}\hat{H}(t+\delta t)}{\hbar}\right)^n\frac{1}{n!}-\sum_{n=0}^{\infty}\left(\frac{-\mathrm{i}\hat{H}t^n}{\hbar}\right)^n\frac{1}{n!}}{\delta t}\right]|\phi(0)\rangle,\\
&= \lim_{\delta t\to 0}\left[\frac{\sum_{n=0}^{\infty}\left(\frac{-\mathrm{i}\hat{H}}{\hbar}\right)^n\frac{(t+\delta t)^n}{n!}-\sum_{n=0}^{\infty}\left(\frac{-\mathrm{i}\hat{H}}{\hbar}\right)^n\frac{t^n}{n!}}{\delta t}\right]|\phi(0)\rangle,\\
&= \lim_{\delta t\to 0}\left[\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{-\mathrm{i}\hat{H}}{\hbar}\right)^n\frac{(t+\delta t)^n-t^n}{\delta t}\right]|\phi(0)\rangle,\\
&= \left[\sum_{n=0}^{\infty}\frac{1}{n!}nt^{n-1}\left(\frac{-\mathrm{i}\hat{H}}{\hbar}\right)^n\right]|\phi(0)\rangle,\\
&= \sum_{n=1}^{\infty}\frac{1}{(n-1)!}t^{n-1}\left(\frac{-\mathrm{i}\hat{H}}{\hbar}\right)^{n-1}\left(\frac{-\mathrm{i}\hat{H}}{\hbar}\right)|\phi(0)\rangle,\\
&= \left(\frac{-\mathrm{i}\hat{H}}{\hbar}\right)\sum_{n=1}^{\infty}\frac{1}{(n-1)!}t^{n-1}\left(\frac{-\mathrm{i}\hat{H}}{\hbar}\right)^{n-1}|\phi(0)\rangle,\\
&= \left(\frac{-\mathrm{i}\hat{H}}{\hbar}\right)e^{-\mathrm{i}\hat{H}t/\hbar}|\phi(0)\rangle,\\
&= \left(\frac{-\mathrm{i}\hat{H}}{\hbar}\right)|\phi(t)\rangle.
\end{aligned}$$

Multiplying both sides by $\mathrm{i}\hbar$ gives the celebrated **Schrödinger equation**:

$$\mathrm{i}\hbar\frac{\partial}{\partial t}|\phi(t)\rangle = \hat{H}|\phi(t)\rangle.$$

The operator $\hat{H}$ is called the **Hamiltonian** operator and has dimensions of energy. It is identified with the (non-relativistic) energy of the system. If the system in question is a single particle experiencing a potential $\mathcal{U}(\boldsymbol{x})$, then

$$\hat{H} = \frac{1}{2m}\hat{\mathbf{p}}^2 + \widehat{\mathcal{U}(\boldsymbol{x})},$$

where $\hat{\mathbf{p}}$ is the momentum. However, the left-hand side is an operator, therefore, the momentum squared $\hat{\mathbf{p}}^2$, and the potential energy $\widehat{\mathcal{U}(\boldsymbol{x})}$, must be promoted to operator status.

7.3 The Heisenberg picture

In the so-called **Schrödinger picture** of quantum mechanics, states of the system evolve in time:

$$|\phi(0)\rangle \to e^{-\mathrm{i}\hat{H}t/\hbar}|\phi(0)\rangle,$$

while observables stay constant, $\hat{\mathrm{A}} \to_t \hat{\mathrm{A}}$. On the other hand, in the so-called **Heisenberg picture** of quantum mechanics, the states of the system are regarded as constant:

$$|\phi(0)\rangle \to_t |\phi(0)\rangle,$$

while observables change over time:

$$\hat{\mathrm{A}} \to_t \hat{\mathrm{A}}_t := \hat{U}^\dagger \hat{\mathrm{A}} \hat{U}, \qquad \hat{U} = \mathrm{e}^{-\mathrm{i}\hat{H}t/\hbar}.$$

However, both pictures are equivalent, because the Heisenberg expectation value

$$\langle\phi(0)|\hat{\mathrm{A}}_t|\phi(0)\rangle$$

is equal to the Schrödinger expectation value:

$$\begin{aligned} \langle\phi(0)|\hat{\mathrm{A}}_t|\phi(0)\rangle &= \langle\phi(0)|\hat{U}^\dagger\hat{\mathrm{A}}\hat{U}|\phi(0)\rangle, \\ &= \langle\phi(t)|\hat{\mathrm{A}}|\phi(t)\rangle. \end{aligned}$$

Just as Schrödinger states satisfy a time-evolution equation, so too do Heisenberg observables:

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\hat{\mathrm{A}}_t = \left[\hat{\mathrm{A}}_t, \hat{H}\right].$$

Proof: We work with expectation values:

$$\begin{aligned} \mathrm{i}\hbar\langle\phi(0)|\hat{\mathrm{A}}_t|\phi(0)\rangle_H &= \mathrm{i}\hbar\langle\phi(t)|\hat{\mathrm{A}}|\phi(t)\rangle_S, \\ \mathrm{i}\hbar\frac{\partial}{\partial t}\langle\phi(0)|\hat{\mathrm{A}}_t|\phi(0)\rangle_H &= \mathrm{i}\hbar\frac{\partial}{\partial t}\langle\phi(t)|\hat{\mathrm{A}}|\phi(t)\rangle_S, \end{aligned}$$

$$
\begin{aligned}
\langle\phi(0)|\mathrm{i}\hbar\frac{\partial\hat{\mathrm{A}}_t}{\partial t}|\phi(0)\rangle_H &= \left[\mathrm{i}\hbar\left(\frac{\partial}{\partial t}\langle\phi(t)|\right)\hat{\mathrm{A}}|\phi(t)\rangle + \langle\phi(t)|\hat{\mathrm{A}}\left(\mathrm{i}\hbar\frac{\partial}{\partial t}|\phi(t)\rangle\right)\right]_S, \\
&= \left[-\langle\phi(t)|\hat{H}\hat{\mathrm{A}}|\phi(t)\rangle + \langle\phi(t)|\hat{\mathrm{A}}\hat{H}|\phi(t)\rangle\right]_S, \\
&= \langle\phi(t)|\left(\hat{\mathrm{A}}\hat{H} - \hat{H}\hat{\mathrm{A}}\right)|\phi(t)\rangle_S, \\
&= \langle\phi(0)|\left(\hat{\mathrm{A}}\hat{H} - \hat{H}\hat{\mathrm{A}}\right)\hat{U}(t)|\phi(0)\rangle_H, \\
&= \langle\phi(0)|\hat{U}(t)^\dagger\hat{\mathrm{A}}\hat{H}\hat{U}(t)|\phi(0)\rangle_H - \langle\phi(0)|\hat{U}(t)^\dagger\hat{H}\hat{\mathrm{A}}\hat{U}(t)|\phi(0)\rangle_H, \\
&= \langle\phi(0)|\hat{U}(t)^\dagger\hat{\mathrm{A}}\hat{U}(t)\hat{H}|\phi(0)\rangle_H - \langle\phi(0)|\hat{H}\hat{U}(t)^\dagger\hat{\mathrm{A}}\hat{U}(t)|\phi(0)\rangle_H, \\
&= \langle\phi(0)|\hat{\mathrm{A}}_t\hat{H}|\phi(0)\rangle_H - \langle\phi(0)|\hat{H}\hat{\mathrm{A}}_t|\phi(0)\rangle_H,
\end{aligned}
$$

for all initial states $|\phi(0)\rangle$. Hence,

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\hat{\mathrm{A}}_t = \left[\hat{\mathrm{A}}_t, \hat{H}\right].$$

Unless otherwise stated, we shall use the Schrödinger picture in this module.

Chapter 8

Expectation values and uncertainty

Reading material for this chapter: Mandl, Chapters 1 and 3

8.1 Introduction

According to postulate 4,

> A physical observable is represented by a Hermitian operator on $\mathcal{H}$; the only allowed results of measurement of the physical observable are the eigenvalues of the operator.
>
> Quantum mechanics is probabilistic: the probability that a system prepared in a state $|x\rangle \in \mathcal{H}$ is measured to be in an eigenstate $|x_a\rangle$ of the observable A is given by
>
> $$|\langle x_a|x\rangle|^2,$$
>
> where $\langle x_a| \in \mathcal{H}^*$ is dual to $|x_a\rangle \in \mathcal{H}$ and $\langle x_a|x\rangle$ is the natural pairing.

In this section carry out some calculations based on this postulate. As always, let $\mathcal{H}$ be the Hilbert space of some physical system, and let $\hat{\mathrm{A}}$ and $\hat{\mathrm{B}}$ be observables.

8.2 Expectation values

Let the operator $\hat{\mathrm{A}}$ be equipped with the complete (possibly degenerate) orthonormal basis $\{|x_i\rangle\}$, with (possibly) degenerate eigenvalues a_i. Then, for any state $|\phi\rangle$,

$$|\phi\rangle = \sum_j |x_j\rangle\langle x_j|\phi\rangle.$$

The amplitude that a measurement forces the system into the eigenstate $|x_i\rangle$ is

$$\langle x_i|\phi\rangle,$$

with probability

$$|\langle x_i|\phi\rangle|^2.$$

Thus, we may view the observable A as a random variable that takes certain values a_i with probability $P_i := |\langle x_i|\phi\rangle|^2$. We know how to find the average value of such random variables – it is called the **expectation value**:

$$\text{average value of a} = \sum_i P_i a_i := \langle \hat{\mathrm{A}} \rangle_\phi,$$

where the probabilities sum to unity: $\sum_i P_i = 1$. This latter condition is guaranteed provided $\langle\phi|\phi\rangle = 1$, since

$$\begin{aligned}
\mathbb{I} &= \sum_i |x_i\rangle\langle x_i|, \\
\langle\phi|\phi\rangle &= \sum_i \langle\phi|x_i\rangle\langle x_i|\phi\rangle, \\
&= \sum_i |\langle x_i|\phi\rangle|^2, \\
&= \sum_i P_i.
\end{aligned}$$

Let's look at the expression for $\langle \hat{\mathrm{A}} \rangle_\phi$ again:

$$\begin{aligned}
\langle \hat{\mathrm{A}} \rangle_\phi &= \sum_i P_i a_i, \\
&= \sum_i |\langle x_i|\phi\rangle|^2 a_i, \\
&= \sum_i \langle\phi|x_i\rangle\langle x_i|\phi\rangle a_i, \\
&= \sum_i \langle\phi|x_i\rangle\langle a_i x_i|\phi\rangle, \qquad \ldots \qquad a_i \in \mathbb{R}, \\
&= \sum_i \langle\phi|x_i\rangle \left(\langle x_i|\hat{\mathrm{A}}\right)|\phi\rangle, \\
&= \sum_i \langle\phi|x_i\rangle\langle x_i| \left(\hat{\mathrm{A}}|\phi\rangle\right), \\
&= \langle\phi| \left(\sum_i |x_i\rangle\langle x_i|\right) \left(\hat{\mathrm{A}}|\phi\rangle\right), \\
&= \langle\phi|\hat{\mathrm{A}}|\phi\rangle.
\end{aligned}$$

Thus, the **expectation value** of the observable $\hat{\mathrm{A}}$ in a state $|\phi\rangle$ is given by

$$\langle\hat{\mathrm{A}}\rangle = \langle\phi|\hat{\mathrm{A}}|\phi\rangle.$$

In a similar manner, we define the **uncertainty** in observations of the observable A as the standard deviation of the observations away from the expected values:

$$\text{uncertainty in } A = \sqrt{\left\langle \left(\hat{\mathrm{A}} - \langle\hat{\mathrm{A}}\rangle\right)^2 \right\rangle} := \Delta A$$

Note: The expectation value and the uncertainty always depend on the state $|\phi\rangle$ in which they are calculated. Thus, we speak of 'the expectation value w.r.t. a particular state'.

8.3 Uncertainty principle

We prove the following theorem:

Theorem 8.1 *Let $\hat{\mathrm{A}}$ and $\hat{\mathrm{B}}$ be Hermitian operators on a Hilbert space $\mathcal{H}$, that satisfy the following commutation relation:*

$$\left[\hat{\mathrm{A}}, \hat{\mathrm{B}}\right] = \alpha\mathbb{I}, \qquad \alpha \in \mathbb{C}.$$

It follows that

$$\Delta A \Delta B \geq |\alpha|/2.$$

Proof: Form the operators

$$\hat{\alpha} = \hat{\mathrm{A}} - \langle\hat{\mathrm{A}}\rangle_\phi \mathbb{I}, \qquad \hat{\beta} = \hat{\mathrm{B}} - \langle\hat{\mathrm{B}}\rangle_\phi \mathbb{I}.$$

Thus,

$$\Delta A^2 = \langle\phi|\hat{\alpha}^2|\phi\rangle = \|\hat{\alpha}\phi\|_2^2, \qquad \Delta B^2 = \langle\phi|\hat{\beta}^2|\phi\rangle = \|\hat{\beta}\phi\|_2^2,$$

and

$$\begin{aligned} \Delta A^2 \Delta B^2 &= \|\hat{\alpha}\phi\|^2 \|\hat{\beta}\phi\|^2, \\ &\geq \left|\langle\phi|\hat{\alpha}\hat{\beta}|\phi\rangle\right|^2, \end{aligned}$$

where the inequality is due to Cauchy–Schwartz. As in that proof, consider the following trick:

$$\begin{aligned}\langle\phi|\hat{\alpha}\hat{\beta}|\phi\rangle &= \left|\langle\phi|\hat{\alpha}\hat{\beta}|\phi\rangle\right| \mathrm{e}^{\mathrm{i}\theta},\\ \langle\phi|\hat{\alpha}\hat{\beta}|\phi\rangle^* &= \left|\langle\phi|\hat{\alpha}\hat{\beta}|\phi\rangle\right| \mathrm{e}^{-\mathrm{i}\theta},\\ &= \langle\phi|\hat{\beta}\hat{\alpha}|\phi\rangle.\end{aligned}$$

Subtract these results:

$$\langle\phi|\hat{\alpha}\hat{\beta}|\phi\rangle - \langle\phi|\hat{\beta}\hat{\alpha}|\phi\rangle = 2\mathrm{i}\sin\theta\left|\langle\phi|\hat{\alpha}\hat{\beta}|\phi\rangle\right|,$$

in other words,

$$\left|\langle\phi|\left[\hat{\alpha},\hat{\beta}\right]|\phi\rangle\right| = 2|\sin\theta|\left|\langle\phi|\hat{\alpha}\hat{\beta}|\phi\rangle\right|.$$

But $|\sin\theta| \leq 1$, hence

$$\left|\langle\phi|\left[\hat{\alpha},\hat{\beta}\right]|\phi\rangle\right|^2 \leq 4|\left|\langle\phi|\hat{\alpha}\hat{\beta}|\phi\rangle\right|^2.$$

Going back to the Cauchy–Schwartz result, we have

$$\begin{aligned}\Delta A^2\Delta B^2 &\geq \left|\langle\phi|\hat{\alpha}\hat{\beta}|\phi\rangle\right|^2,\\ &\geq \frac{1}{4}\left|\langle\phi|\left[\hat{\alpha},\hat{\beta}\right]|\phi\rangle\right|^2.\end{aligned}$$

Specialising to the commutation relation assumed in the theorem, we have

$$\Delta A\Delta B \geq |\alpha|/2.$$

Chapter 9

Representation of Hilbert spaces

Reading material for this chapter: Mandl, Chapter 12; LandauLifshitz, Chapters 2–3.

9.1 Introduction

A Hilbert space is an abstract object. For example, the n-dimensional Hilbert space $\mathcal{H}_n$ is an abstract object defined by the axioms in Ch. 2, and endowed with the natural scalar product obtained by studying the dual space. The set

$$\mathbb{C}^n = \{(z_1, \cdots, z_n) \,|z_1, \cdots, z_n \in \mathbb{C}\}$$

can be thought of as a **realisation** or a **representation** of the abstract vector space, and we might write $\mathcal{H}_n \sim \mathbb{C}^n$. However, there are many possible representations of the n-dimensional Hilbert space. For example, we could take

$$\mathcal{H}_n \sim \mathcal{S}\,(M_1, \cdots, M_n)\,,$$

where $M_1, \cdots, M_n$ are some n linearly independent matrices all of the same size. However, a bijective map between $\mathbb{C}^n$ and $\mathcal{S}\,(M_1, \cdots, M_n)$ exists, which guarantees that these representations are equivalent. A very earthy way of thinking about this is to regard the Hilbert space as being like a cookery book, with many receipes (abstract lists) for delicious pies. Then, the representations of the Hilbert space are like your mother's cooking, where those receipes are embodied in real, solid food.

9.2 The position representation

The position of a particle is a physical quantity, therefore, by Postulate 3, there must be a Hermitian operator associated with it. Call it $\hat{\mathbf{r}}$, the **position operator**. If the particle is at location $\boldsymbol{r}_0$, it can be regarded as having the state $|\boldsymbol{r}_0\rangle$. Then,

$$\hat{\mathbf{r}}|\boldsymbol{r}_0\rangle = \boldsymbol{r}_0|\boldsymbol{r}_0\rangle,$$

and the collection of kets $\{|\boldsymbol{r}\rangle\}_{\boldsymbol{r}\in\mathbb{R}^3}$ is therefore a basis of eigenvectors[1]. We normalise the basis such that

$$\langle\boldsymbol{r}'|\boldsymbol{r}\rangle = \delta(\boldsymbol{r}-\boldsymbol{r}').$$

We assume completeness:

$$\mathbb{I} = \int \mathrm{d}^3 r |\boldsymbol{r}\rangle\langle\boldsymbol{r}|.$$

Thus, an arbitrary state $|\phi\rangle$ has the form

$$|\phi\rangle = \int \mathrm{d}^3 r |\boldsymbol{r}\rangle\langle\boldsymbol{r}|\phi\rangle.$$

But, by postulate 4, the quantity

$$\langle\boldsymbol{r}|\phi\rangle$$

is the amplitude that the state $|\phi\rangle$ is measured to have a position $\boldsymbol{r}$. In a probabilistic setting, the probability to measure a point is zero, thus we assign the quantity

$$|\langle\boldsymbol{r}|\phi\rangle|^2 \mathrm{d}^3 r$$

the probability that the state $|\phi\rangle$ is measured within an small volume $\mathrm{d}^3 r$ of space. We call

$$\phi(\boldsymbol{r}) := \langle\boldsymbol{r}|\phi\rangle$$

the **wave function**.

Let us justify the choice of normalisation $\langle\boldsymbol{r}'|\boldsymbol{r}\rangle = \delta(\boldsymbol{r}-\boldsymbol{r}')$. By the completeness relation, we have

$$\begin{aligned} |\phi\rangle &= \int \mathrm{d}^3 r |\boldsymbol{r}\rangle\langle\boldsymbol{r}|\phi\rangle, \\ \langle\boldsymbol{r}'|\phi\rangle &= \int \mathrm{d}^3 r \langle\boldsymbol{r}'|\boldsymbol{r}\rangle\langle\boldsymbol{r}|\phi\rangle, \\ \phi(\boldsymbol{r}') &= \int \mathrm{d}^3 r \langle\boldsymbol{r}'|\boldsymbol{r}\rangle\phi(\boldsymbol{r}). \end{aligned}$$

[1] This is a uncountable set; however, we do not worry ourselves with the details here, and assume that all relevant results of spectral theory apply.

But $\phi(\boldsymbol{r})$ is an arbitrary function; the only way for this integral equation to hold for all functions is if $\langle \boldsymbol{r}'|\boldsymbol{r}\rangle = \delta(\boldsymbol{r}-\boldsymbol{r}')$, since then

$$\phi(\boldsymbol{r}') = \int \mathrm{d}^3 r \delta(\boldsymbol{r}-\boldsymbol{r}')\phi(\boldsymbol{r}) = \phi(\boldsymbol{r}').$$

In addition,

$$\langle \boldsymbol{r}'|\hat{\mathbf{r}}|\boldsymbol{r}\rangle = \boldsymbol{r}\langle \boldsymbol{r}'|\boldsymbol{r}\rangle = \boldsymbol{r}\delta(\boldsymbol{r}-\boldsymbol{r}'),$$

and, by a Taylor expansion,

$$\langle \boldsymbol{r}'|f(\hat{\mathbf{r}})|\boldsymbol{r}\rangle = f(\boldsymbol{r})\langle \boldsymbol{r}'|\boldsymbol{r}\rangle = f(\boldsymbol{r})\delta(\boldsymbol{r}-\boldsymbol{r}'),$$

By completeness,

$$\begin{aligned}\langle \boldsymbol{r}|f(\hat{\mathbf{r}})|\phi\rangle &= \int \mathrm{d}^3 r\, \langle \boldsymbol{r}|f(\hat{\mathbf{r}})|\boldsymbol{r}'\rangle\langle \boldsymbol{r}'|\phi\rangle, \\ &= f(\boldsymbol{r})\langle \boldsymbol{r}|\phi\rangle = f(\boldsymbol{r})\phi(\boldsymbol{r}).\end{aligned}$$

Thus, in position representation, the operator $f(\hat{\mathbf{r}})$ acting on a state $|\phi\rangle$ corresponds to multiplying the wavefunction $\phi(\boldsymbol{r})$ by $f(\boldsymbol{r})$.

9.3 The scalar product

Take the completeness relation:

$$\mathbb{I} = \int \mathrm{d}^3 r |\boldsymbol{r}\rangle\langle \boldsymbol{r}|,$$

and operate on the identity from both sides with $\langle\phi|$ and $|\psi\rangle$:

$$\langle\phi|\psi\rangle = \int \mathrm{d}^3 r \langle\phi|\boldsymbol{r}\rangle\langle \boldsymbol{r}|\psi\rangle.$$

Thus, the position represenation of the natural pairing is the ordinary scalar product on function spaces:

$$\langle\phi|\psi\rangle = \int \mathrm{d}^3 r \phi^*(\boldsymbol{r})\psi(\boldsymbol{r}),$$

and the norm has the representation

$$\langle\phi|\phi\rangle = \int \mathrm{d}^3 r \phi^*(\boldsymbol{r})\phi(\boldsymbol{r}) = \int \mathrm{d}^3 r |\phi(\boldsymbol{r})|^2.$$

But $\mathrm{d}^3 r|\phi(\boldsymbol{r})|^2$ is a probability, hence

$$\int \mathrm{d}^3 r|\phi(\boldsymbol{r})|^2 = 1,$$

consistent with the requirement that physical states have unit norm:

$$\langle\phi|\phi\rangle = 1.$$

Thus, allowed wavefunctions live in the space

$$L^2(\mathbb{R}^3) = \{\phi|\|\phi\|_2^2 < \infty\}.$$

9.4 Momentum representation

As in the position case, the momentum of a particle is a physical quantity, therefore, by Postulate 3, there must be a Hermitian operator associated with it. Call it $\hat{\mathbf{p}}$, the **momentum operator**. If the particle has momentum $\boldsymbol{p}_0$, it can be regarded as having the state $|\boldsymbol{p}_0\rangle$. Then,

$$\hat{\mathbf{p}}|\boldsymbol{p}_0\rangle = \boldsymbol{p}_0|\boldsymbol{p}_0\rangle,$$

and the collection of kets $\{|\boldsymbol{p}\rangle\}_{\boldsymbol{p}\in\mathbb{R}^3}$ is therefore a basis of eigenvalues. We normalise the basis such that

$$\langle\boldsymbol{p}'|\boldsymbol{p}\rangle = \delta(\boldsymbol{p}-\boldsymbol{p}').$$

We assume completeness:

$$\mathbb{I} = \int \mathrm{d}^3 p|\boldsymbol{p}\rangle\langle\boldsymbol{p}|.$$

Thus, an arbitrary state $|\phi\rangle$ has the form

$$|\phi\rangle = \int \mathrm{d}^3 p|\boldsymbol{p}\rangle\langle\boldsymbol{p}|\phi\rangle.$$

But, by postulate 4, the quantity

$$\langle\boldsymbol{p}|\phi\rangle$$

is the amplitude that the state $|\phi\rangle$ is measured to have a momentum $\boldsymbol{p}$. As before, we are forced to assign the quantity

$$|\langle\boldsymbol{p}|\phi\rangle|^2\mathrm{d}^3 p$$

the probability that the state $|\phi\rangle$ is measured within an small volume d^3p of momentum space. We call

$$\widehat{\phi}_{\boldsymbol{p}} := \langle \boldsymbol{p}|\phi\rangle$$

the momentum-space wavefunction. However, by Fourier-transform theory, we know that a spatial signal $\phi(\boldsymbol{r})$ is made up of a sum of plane waves in momentum space:

$$\phi(\boldsymbol{r}) = \int \frac{\mathrm{d}^3p}{(2\pi\hbar)^3} \mathrm{e}^{\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}\widehat{\phi}_{\boldsymbol{p}}, \qquad \boldsymbol{k} = \boldsymbol{p}/\hbar.$$

In other words,

$$\begin{aligned}
\phi(\boldsymbol{r}) &= \langle \boldsymbol{r}|\phi\rangle, \\
&= \int \frac{\mathrm{d}^3p}{(2\pi\hbar)^3} \mathrm{e}^{\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}\widehat{\phi}_{\boldsymbol{p}}, \\
&= \int \frac{\mathrm{d}^3p}{(2\pi\hbar)^3} \mathrm{e}^{\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}\langle \boldsymbol{p}|\phi\rangle,
\end{aligned}$$

Since $|\phi\rangle$ is arbitrary, we take it to be $|\phi\rangle = |\boldsymbol{p}'\rangle$. Thus,

$$\begin{aligned}
\langle \boldsymbol{r}|\phi\rangle = \langle \boldsymbol{r}|\boldsymbol{p}'\rangle &= \int \frac{\mathrm{d}^3p}{(2\pi\hbar)^3} \mathrm{e}^{\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}\langle \boldsymbol{p}|\boldsymbol{p}'\rangle, \\
&= \int \frac{\mathrm{d}^3p}{(2\pi\hbar)^3} \mathrm{e}^{\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}\delta(\boldsymbol{p}-\boldsymbol{p}'), \\
&= \frac{\mathrm{e}^{\mathrm{i}\boldsymbol{p}'\cdot\boldsymbol{r}/\hbar}}{(2\pi\hbar)^3}.
\end{aligned}$$

Thus, we have the following condition:

$$\langle \boldsymbol{r}|\boldsymbol{p}\rangle = \frac{\mathrm{e}^{\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}}{(2\pi\hbar)^3}.$$

As before, we have

$$\langle \boldsymbol{p}|f(\hat{\mathbf{p}})|\phi\rangle = f(\boldsymbol{p})\widehat{\phi}_{\boldsymbol{p}}.$$

However, we are more interested in understanding the action of the momentum operator in position space. Therefore, we compute

$$\begin{aligned}
\langle \boldsymbol{r}|\hat{\mathbf{p}}|\phi\rangle &= \int \mathrm{d}^3p\, \langle \boldsymbol{r}|\hat{\mathbf{p}}|\boldsymbol{p}\rangle\langle \boldsymbol{p}|\phi\rangle, \\
&= \int \mathrm{d}^3p\, \boldsymbol{p}\langle \boldsymbol{r}|\boldsymbol{p}\rangle\langle \boldsymbol{p}|\phi\rangle, \\
&= \int \mathrm{d}^3p\, \boldsymbol{p}\frac{\mathrm{e}^{\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}}{(2\pi\hbar)^3}\langle \boldsymbol{p}|\phi\rangle, \\
&= \int \mathrm{d}^3p\, \boldsymbol{p}\frac{\mathrm{e}^{\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}}{(2\pi\hbar)^3}\widehat{\phi}_{\boldsymbol{p}}, \\
&= -\mathrm{i}\hbar\nabla\int \mathrm{d}^3\boldsymbol{p}\frac{\mathrm{e}^{\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar}}{(2\pi\hbar)^3}\widehat{\phi}_{\boldsymbol{p}}, \\
&= -\mathrm{i}\hbar\nabla\phi(\boldsymbol{r}).
\end{aligned}$$

Thus, in position representation, the operator $\hat{\mathbf{p}}$ acting on a state $|\phi\rangle$ corresponds to acting on the wavefunction $\phi(\boldsymbol{r})$ by $-\mathrm{i}\hbar\nabla$. The same is true for powers of $\hat{\mathbf{p}}$.

Finally, therefore, the operator

$$\hat{H} = \frac{1}{2m}\hat{\mathbf{p}}^2 + \widehat{\mathcal{U}(\boldsymbol{r})},$$

in position space, corresponds to

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + \mathcal{U}(\boldsymbol{r}),$$

and Schrödinger's equation

$$\mathrm{i}\hbar\frac{\partial}{\partial t}|\phi(t)\rangle = \hat{H}|\phi(t)\rangle,$$

is represented by the following operator equation in the wavefunction $\phi(\boldsymbol{r},t)$:

$$\mathrm{i}\hbar\frac{\partial\phi}{\partial t} = \left(-\frac{\hbar^2}{2m}\nabla^2 + \mathcal{U}(\boldsymbol{r})\right)\phi.$$

9.5 Heisenberg uncertainty

We can readily compute the commutation relation betweem $\hat{\mathbf{p}}$ and $\hat{\mathbf{r}}$ in the position representation. For, let $\phi(\boldsymbol{r})$ be an arbitrary function. Then,

$$\begin{aligned}
\hat{\mathrm{p}}_i\left(\hat{\mathrm{r}}_j\phi\right) &= -\mathrm{i}\hbar\frac{\partial}{\partial r_i}\left(r_j\phi\right), \\
&= -\mathrm{i}\hbar\delta_{ij}\phi(\boldsymbol{r}) - \mathrm{i}\hbar r_j\frac{\partial\phi}{\partial r_i}.
\end{aligned}$$

Similarly,

$$\hat{\mathrm{r}}_j \left(\hat{\mathrm{p}}_i \phi\right) = -\mathrm{i}\hbar r_j \frac{\partial \phi}{\partial r_i}.$$

Subtracting gives

$$\begin{aligned} \hat{\mathrm{p}}_i \left(\hat{\mathrm{r}}_j \phi\right) - \hat{\mathrm{r}}_j \left(\hat{\mathrm{p}}_i \phi\right) &= \left[\hat{\mathrm{p}}_i, \hat{\mathrm{r}}_j\right] \phi(\boldsymbol{r}), \\ &= -\mathrm{i}\hbar \delta_{ij} \phi(\boldsymbol{r}). \end{aligned}$$

Since this is true for all functions $\phi(\boldsymbol{r})$, we have

$$\left[\hat{\mathrm{p}}_i, \hat{\mathrm{r}}_j\right] = -\mathrm{i}\hbar \delta_{ij}.$$

Applying the uncertainty theorem from Ch. 8, we have

$$\Delta p_x \Delta x \geq \hbar/2,$$

and similarly for the other directions.

9.6 Conservation law of probability

Let us take the probability density

$$P(\boldsymbol{r}, t) = |\phi(\boldsymbol{r}, t)|^2,$$

differentiate it, and apply the Schrödinger equation:

$$\frac{\partial P}{\partial t} = \phi^* \frac{\partial \phi}{\partial t} + \frac{\partial \phi^*}{\partial t} \phi.$$

We have,

$$\frac{\partial \phi}{\partial t} = \frac{1}{\mathrm{i}\hbar} \hat{H} \phi,$$

and

$$\frac{\partial \phi^*}{\partial t} = \frac{-1}{\mathrm{i}\hbar} \hat{H} \phi^*,$$

since $\hat{H}$ is assumed to be both real-valued and Hermitian. Thus

$$\begin{aligned} \frac{\partial P}{\partial t} &= \frac{1}{\mathrm{i}\hbar} \left(\phi^* \hat{H} \phi - \phi \hat{H} \phi^*\right), \\ &= \frac{1}{\mathrm{i}\hbar} \left(-\frac{\hbar^2}{2m} \phi^* \nabla^2 \phi + \phi^* \mathcal{U}(\boldsymbol{r}) \phi + \frac{\hbar^2}{2m} \phi \nabla^2 \phi^* - \phi \mathcal{U}(\boldsymbol{r}) \phi^*\right), \\ &= -\frac{\hbar^2}{2m} \frac{1}{\mathrm{i}\hbar} \left(\phi^* \nabla^2 \phi - \phi \nabla^2 \phi^*\right). \end{aligned}$$

Now we apply a neat trick: Green's theorem:

$$\phi^*\nabla^2\phi - \phi\nabla^2\phi^* = \nabla\cdot\left(\phi^*\nabla\phi - \phi\nabla\phi^*\right).$$

Calling

$$\boldsymbol{J} := \frac{\hbar}{2m\mathrm{i}}\left(\phi^*\nabla\phi - \phi\nabla\phi^*\right),$$

we have the following **conservation law**:

$$\frac{\partial P}{\partial t} + \nabla\cdot\boldsymbol{J} = 0.$$

The vector field $\boldsymbol{J}$ is called the **probability current**. Integrating the conservation law over a domain Ω gives

$$\begin{aligned}\int_\Omega \mathrm{d}^3r\,\frac{\partial P}{\partial t} &= -\int_\Omega \mathrm{d}^3r\nabla\cdot\boldsymbol{J},\\ \frac{\partial}{\partial t}\int_\Omega \mathrm{d}^3r\,|\phi(\boldsymbol{r},t)|^2 &= -\int_{\partial\Omega}\mathrm{d}\boldsymbol{S}\cdot\boldsymbol{J},\end{aligned}$$

or

$$\frac{\partial}{\partial t}\text{Prob}\left(\text{System in region }\Omega\text{ at time }t\right) = -\int_{\partial\Omega}\mathrm{d}\boldsymbol{S}\cdot\boldsymbol{J}.$$

Taking $\Omega = \mathbb{R}^3$ and assuming that the wavefunction vanishes as $|\boldsymbol{r}|\to\infty$, we have the following law of **conservation of probability**:

$$\frac{\partial}{\partial t}\text{Prob}\left(\text{System somewhere in }\mathbb{R}^3\text{ at time }t\right) = 0. \qquad (*)$$

We normalise the wavefunction:

$$\int_{\mathbb{R}^3}\mathrm{d}^3r\,|\phi(\boldsymbol{r})|^2 = 1;$$

Eq. (*) guarantees that it stays normalised for all time. Note:

> The conservation law of probability is derived from Schrödinger's equation, which in turn is derived from the unitarity assumption (Postulate 3). Thus, unitarity is needed to ensure conservation of probabilities.

Chapter 10

Plane waves, or the free particle

Reading material for this chapter: None recommended

In this section we study the dynamics of a free particle, experiencing no forces, moving in three-dimensional space. We first of all recall how this problem is studied in classical mechanics. The energy of such a system is conserved and given by

$$\frac{\boldsymbol{p}^2}{2m} = E.$$

But $\boldsymbol{p} = m\dot{\boldsymbol{r}}$, hence

$$\tfrac{1}{2}m\dot{\boldsymbol{r}}^2 = E.$$

Differentiating with respect to time gives

$$\dot{\boldsymbol{r}} \cdot \ddot{\boldsymbol{r}} = 0.$$

Now the only way for $\ddot{\boldsymbol{r}} \perp \dot{\boldsymbol{r}} = 0$ for all possible vectors $\dot{\boldsymbol{r}}$ is if $\ddot{\boldsymbol{r}} = 0$, hence

$$\boldsymbol{r}(t) = \boldsymbol{r}_0 + \boldsymbol{v}_0 t,$$

and the particle moves in a straight line. This is the simplest possible mechanical system.

Using quantum mechanics and the position representation, we know how to solve this problem: The momentum $\boldsymbol{p}$ is promoted to operator status:

$$\boldsymbol{p} \to \hat{\mathbf{p}} = -\mathrm{i}\hbar\nabla,$$

and E is identified with the Hamiltonian:

$$E \to \hat{H},$$

and the Schrödinger equation to solve is

$$\begin{aligned} \mathrm{i}\hbar\frac{\partial\Psi}{\partial t} &= \frac{\hat{\mathbf{p}}^2}{2m}\Psi, \\ &= -\frac{\hbar^2}{2m}\nabla^2\Psi. \end{aligned}$$

To solve this equation, we assume that it is prepared in the initial state

$$\Psi(\boldsymbol{r}, t=0) = \psi_0(\boldsymbol{r}),$$

and introduce the Fourier transform:

$$\begin{aligned} \hat{\Psi}_{\boldsymbol{k}}(t) &:= \int \mathrm{d}^3 r \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\Psi(\boldsymbol{r}, t), \\ \Psi(\boldsymbol{r}, t) &= \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\hat{\Psi}_{\boldsymbol{k}}(t) \end{aligned}$$

Next, we operate on both sides of the Schrödinger equation with $\int \mathrm{d}^3 r \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$:

$$\begin{aligned} \int \mathrm{d}^3 r \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\left[\mathrm{i}\hbar\frac{\partial\Psi}{\partial t}\right] &= \int \mathrm{d}^3 r \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\left[-\frac{\hbar^2}{2m}\nabla^2\Psi\right], \\ \mathrm{i}\hbar\frac{\partial}{\partial t}\int \mathrm{d}^3 r \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\Psi(\boldsymbol{r}, t) &= \int \mathrm{d}^3 r \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\left[-\frac{\hbar^2}{2m}\nabla^2\Psi\right]. \end{aligned}$$

Hence,

$$\begin{aligned} \mathrm{i}\hbar\frac{\partial\hat{\Psi}_{\boldsymbol{k}}}{\partial t} &= -\frac{\hbar^2}{2m}\int \mathrm{d}^3 r \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\nabla^2\Psi, \\ &= -\frac{\hbar^2}{2m}\left[\int \mathrm{d}^3 r \left[\nabla\cdot\left(\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\nabla\Psi\right) + \mathrm{i}\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\boldsymbol{k}\cdot\nabla\Psi\right]\right], \\ &= -\frac{\hbar^2}{2m}\int \mathrm{d}S\left(\partial_n\Psi\right)\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} - \frac{\mathrm{i}\hbar^2}{2m}\int \mathrm{d}^3 r \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\boldsymbol{k}\cdot\nabla\Psi, \end{aligned}$$

where $\partial_n\Psi$ is the outward-pointing normal derivative of Ψ at $|\boldsymbol{r}| = \infty$; this assumed to be zero. Thus

$$\begin{aligned} \mathrm{i}\hbar\frac{\partial\hat{\Psi}_{\boldsymbol{k}}}{\partial t} &= -\frac{\mathrm{i}\hbar^2}{2m}\int \mathrm{d}^3 r \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\boldsymbol{k}\cdot\nabla\Psi, \\ &= -\frac{\mathrm{i}\hbar^2}{2m}\left[\int \mathrm{d}^3 r \boldsymbol{k}\cdot\left[\nabla\left(\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\Psi\right) + \mathrm{i}\boldsymbol{k}\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\Psi\right]\right], \\ &= -\frac{\mathrm{i}\hbar^2}{2m}\int \mathrm{d}^3 r \boldsymbol{k}\cdot\nabla\left(\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\Psi\right) - \frac{\mathrm{i}\hbar^2}{2m}\int \mathrm{d}^3 r \mathrm{i}\boldsymbol{k}^2\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\Psi, \end{aligned}$$

and the first term vanishes by Gauss's theorem, hence

$$\mathrm{i}\hbar\frac{\partial\hat{\Psi}_{\boldsymbol{k}}}{\partial t} = \frac{\hbar^2}{2m}\int \mathrm{d}^3 r \boldsymbol{k}^2 \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\Psi = \frac{\hbar^2\boldsymbol{k}^2}{2m}\hat{\Psi}_{\boldsymbol{k}}.$$

Thus, we obtain a **dispersion relation**

$$E = E_{\boldsymbol{k}} = \frac{\hbar^2\boldsymbol{k}^2}{2m} = \frac{\boldsymbol{p}^2}{2m},$$

and we solve the following equation in momentum space:

$$\mathrm{i}\hbar\frac{\partial\hat{\Psi}_{\boldsymbol{k}}}{\partial t} = E_{\boldsymbol{k}}\hat{\Psi}_{\boldsymbol{k}}.$$

But we know the solution to this immediately:

$$\hat{\Psi}_{\boldsymbol{k}}(t) = \hat{\Psi}_{\boldsymbol{k}}(0)\mathrm{e}^{-\mathrm{i}E_{\boldsymbol{k}}t/\hbar} := C_{\boldsymbol{k}}\mathrm{e}^{-\mathrm{i}E_{\boldsymbol{k}}t/\hbar}$$

We plug this back into the Fourier-transform solution:

$$\begin{aligned}\Psi(\boldsymbol{r},t) &= \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\hat{\Psi}_{\boldsymbol{k}}(t),\\ &= \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}C_{\boldsymbol{k}}\mathrm{e}^{-\mathrm{i}E_{\boldsymbol{k}}t/\hbar}.\end{aligned}$$

The weights $C_{\boldsymbol{k}}$ can be obtained from the initial condition:

$$\begin{aligned}\Psi(\boldsymbol{r},t=0) &= \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}C_{\boldsymbol{k}},\\ &= \psi_0(\boldsymbol{r}),\end{aligned}$$

hence

$$C_{\boldsymbol{k}} = \int \mathrm{d}^3 r \mathrm{e}^{-\mathrm{i}\boldsymbol{r}\cdot\boldsymbol{k}}\psi_0(\boldsymbol{r}).$$

Let's experiment with some specific initial data.

For simplicity, we focus on the one-dimensional case. The particle is prepared in the following state:

$$\psi_0(x) = N\mathrm{e}^{-(x-x_0)^2/(4\sigma^2)}, \qquad N = \frac{1}{(2\pi\sigma^2)^{1/4}}.$$

where $\sigma > 0$ is some parameter. Hence,

$$\begin{aligned} C_k &= \int \mathrm{d}x \mathrm{e}^{-\mathrm{i}kx} \psi_0(x), \\ &= N \int \mathrm{d}x \mathrm{e}^{-\mathrm{i}kx} \mathrm{e}^{-(x-x_0)^2/(4\sigma^2)}, \end{aligned}$$

To integrate ths, we complete the square in the following manner. First, call $y = x - x_0$. Then,

$$\begin{aligned} C_k &= \int_{-\infty}^{\infty} \mathrm{d}y \mathrm{e}^{-\mathrm{i}k(y+x_0)} \mathrm{e}^{-y^2/4\sigma^2}, \\ &= e^{-\mathrm{i}kx_0} \int_{-\infty}^{\infty} \mathrm{d}y \mathrm{e}^{-\mathrm{i}ky} \mathrm{e}^{-y^2/4\sigma^2}. \end{aligned}$$

Call $a := 1/4\sigma^2$. Then

$$\begin{aligned} C_k &= e^{-\mathrm{i}kx_0} \int_{-\infty}^{\infty} \mathrm{d}y \mathrm{e}^{-\mathrm{i}ky} \mathrm{e}^{-ay^2}, \\ &= Ne^{-\mathrm{i}kx_0} \int_{-\infty}^{\infty} \mathrm{d}y \mathrm{e}^{-ay^2 - \mathrm{i}ky}, \\ &= Ne^{-\mathrm{i}kx_0} \int_{-\infty}^{\infty} \mathrm{d}y \mathrm{e}^{-a(y^2 + (\mathrm{i}k/a)y)}, \\ &= Ne^{-\mathrm{i}kx_0} \int_{-\infty}^{\infty} \mathrm{d}y \mathrm{e}^{-a[(y^2 + (\mathrm{i}k/2a))^2 + k^2/4a^2]}, \\ &= Ne^{-\mathrm{i}kx_0} \int_{-\infty}^{\infty} \mathrm{d}y \mathrm{e}^{-a(y^2 + (\mathrm{i}k/2a))^2} \mathrm{e}^{-k^2/4a}, \\ &= Ne^{-\mathrm{i}kx_0} \mathrm{e}^{-k^2/4a} \int_{-\infty}^{\infty} \mathrm{d}y \mathrm{e}^{-a(y^2 + (\mathrm{i}k/2a))^2}, \\ &= Ne^{-\mathrm{i}kx_0} \mathrm{e}^{-k^2/4a} \int_{-\infty}^{\infty} \mathrm{d}y \mathrm{e}^{-ay^2}, \\ &= Ne^{-\mathrm{i}kx_0} \mathrm{e}^{-k^2/4a} \int_{-\infty}^{\infty} \mathrm{d}y \mathrm{e}^{-ay^2}, \\ &= \frac{N}{\sqrt{a}} e^{-\mathrm{i}kx_0} \mathrm{e}^{-k^2/4a} \int_{-\infty}^{\infty} \mathrm{d}z \mathrm{e}^{-z^2}, \\ &= N\sqrt{\frac{\pi}{a}} e^{-\mathrm{i}kx_0} \mathrm{e}^{-k^2/4a}. \end{aligned}$$

Restoring $a = 1/4\sigma^2$, this is

$$C_k = N\sqrt{4\pi\sigma^2} e^{-\mathrm{i}kx_0} \mathrm{e}^{-k^2\sigma^2}.$$

At later times,

$$\begin{aligned}
\Psi(x,t) &= \int \frac{\mathrm{d}k}{2\pi}\mathrm{e}^{\mathrm{i}kx}C_k\mathrm{e}^{-\mathrm{i}E_kt/\hbar},\\
&= N\sqrt{4\pi\sigma^2}\int \frac{\mathrm{d}k}{2\pi}\mathrm{e}^{\mathrm{i}kx}e^{-\mathrm{i}kx_0}\mathrm{e}^{-\mathrm{i}E_kt/\hbar}\mathrm{e}^{-k^2/4a},\\
&= N\sqrt{4\pi\sigma^2}\int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi}\mathrm{e}^{\mathrm{i}k(x-x_0)}\mathrm{e}^{-k^2\sigma^2}\mathrm{e}^{-\mathrm{i}k^2\hbar^2t/2m},\\
&= N\sqrt{4\pi\sigma^2}\int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi}\mathrm{e}^{\mathrm{i}kb}\mathrm{e}^{-ck^2},
\end{aligned}$$

where

$$b = x - x_0, \qquad c = \sigma^2 - \mathrm{i}\hbar t/2m.$$

Completing the square again,

$$\begin{aligned}
\int_{-\infty}^{\infty}\mathrm{d}k\,\mathrm{e}^{\mathrm{i}kb}\mathrm{e}^{-ck^2} &= \int_{-\infty}^{\infty}\mathrm{d}k\,\mathrm{e}^{-c(k^2-\mathrm{i}kb/c)},\\
&= \int_{-\infty}^{\infty}\mathrm{d}k\,\mathrm{e}^{-c[(k-\mathrm{i}b/2c)^2+b^2/4c^2]},\\
&= \int_{-\infty}^{\infty}\mathrm{d}k\,\mathrm{e}^{-c(k^2-\mathrm{i}kb/c)^2}\mathrm{e}^{-b^2/4c},\\
&= \mathrm{e}^{-b^2/4c}\int_{-\infty}^{\infty}\mathrm{d}k\,\mathrm{e}^{-c(k^2-\mathrm{i}kb/c)^2},\\
&= \mathrm{e}^{-b^2/4c}\int_{-\infty}^{\infty}\mathrm{d}k\,\mathrm{e}^{-ck^2},\\
&= \mathrm{e}^{-b^2/4c}\frac{1}{\sqrt{c}}\int_{-\infty}^{\infty}\mathrm{d}k\,\mathrm{e}^{-k^2},\\
&= \mathrm{e}^{-b^2/4c}\sqrt{\frac{\pi}{c}}.
\end{aligned}$$

Hence,

$$\Psi(x,t) = N\frac{1}{2\pi}\sqrt{4\pi\sigma^2}\mathrm{e}^{-b^2/4c}\sqrt{\frac{\pi}{c}}$$

Restoring the meaning of the coefficients, we have

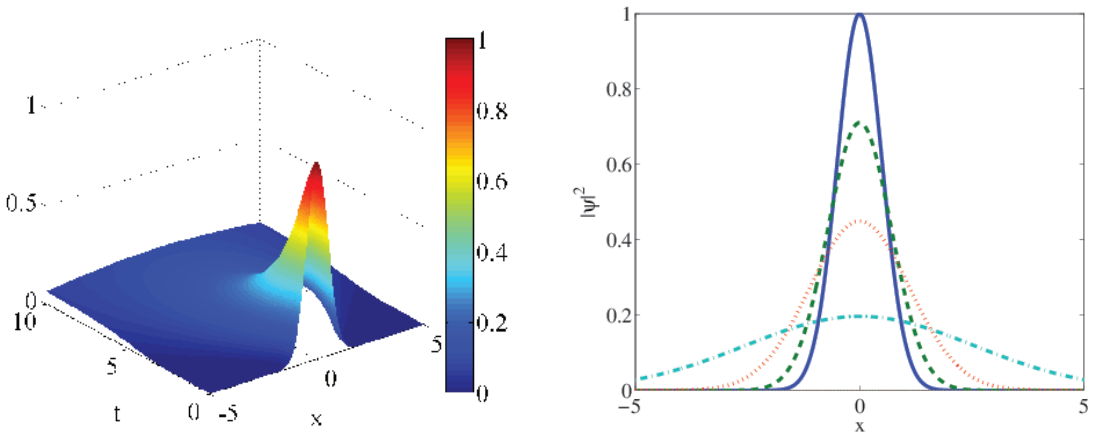
$$\Psi(x,t) = N\frac{1}{2\pi}\sqrt{4\pi\sigma^2}\exp\left[-\frac{(x-x_0)^2/4}{\sigma^2-\mathrm{i}\hbar t/2m}\right]\sqrt{\frac{\pi}{\sigma^2-\mathrm{i}\hbar t/2m}},$$

or

$$\Psi(x,t) = N\sqrt{\frac{\sigma^2}{\sigma^2-\mathrm{i}\hbar t/2m}}\exp\left[-\frac{(x-x_0)^2/4}{\sigma^2-\mathrm{i}\hbar t/2m}\right].$$

Finally,

$$\Psi(x,t) = N\sqrt{\frac{1}{1-\mathrm{i}\hbar t/2m\sigma^2}}\exp\left[-\frac{(x-x_0)^2/4\sigma^2}{1-\mathrm{i}\hbar t/2m\sigma^2}\right].$$



(a) (b)

Figure 10.1: (a) Time evolution of the probability density; (b) The same, for $t = 0, 1, 2, 5, 10$. Here $\sigma = 0.5$ and $\hbar/2m^2 = 1$.

A plot of the probability density associated with this result is shown in Fig. 10.1. Notes:

- The PDF **spreads out** over time – the particle's position becomes less and less certain.

- The mean value of the particle's position stays the same, since the PDF remains centred at zero. In other words,

$$\frac{d}{dt}x_{\mathsf{av}} = 0,$$

or

$$\frac{d}{dt}\langle \hat{\mathrm{x}} \rangle = 0.$$

- This is called **Ehrenfest's theorem** – the expected values of quantum observables obey the classical-mechanical equations. In general,

$$\frac{d}{dt}\langle \widehat{\mathrm{x}} \rangle = \frac{1}{m}\langle \widehat{\mathrm{p}} \rangle,$$
$$\frac{d}{dt}\langle \widehat{\mathrm{p}} \rangle = -\langle \partial_x \mathcal{U} \rangle,$$

  for a particle experiencing a potential $\mathcal{U}(x)$ (proof as homework). This is one way of stating the **correspondnce principle** – that the laws of classical mechanics can be recovered in a certain limit.

- For a system with a discrete spectrum, the correspondence principle states that quantum mechanics reproduces the results of classical mechanics in the limit of large quantum numbers.

10.1 Plane waves

It can be verified that the Fourier-transformed function

$$\widehat{\psi_{\boldsymbol{k}}} = \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}-\mathrm{i}E_{\boldsymbol{k}}t/\hbar}, \qquad E_{\boldsymbol{k}} = \frac{\hbar^2\boldsymbol{k}^2}{2m},$$

satisfies the Schrödinger equation for a free particle. Relabelling, using $\boldsymbol{p} = \hbar\boldsymbol{k}$, we have

$$\begin{aligned} \psi_{\boldsymbol{p}}(\boldsymbol{r}) &:= \widehat{\psi_{\boldsymbol{p}}}, \\ &= \mathrm{e}^{-\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar-\mathrm{i}E_{\boldsymbol{p}}t/\hbar}, \qquad E_{\boldsymbol{p}} = \frac{\boldsymbol{p}^2}{2m}. \end{aligned}$$

This is called the **plane-wave** solution. Note:

- The plane-wave solution is not normalisable ($\|\psi_{\boldsymbol{p}}(\boldsymbol{r})\|_2^2 = \infty$). However, the plane waves do solve the Schrödinger equation, so they must be physical states.
- Thus, we must extend the Hilbert space to include this case. Such an extension is called the **rigged Hilbert space**. We simply note that the extension is required, and do not discuss this functional-analysis topic any further.
- A simple calculation shows that the plane wave is a state of maximum positional uncertainty, $\Delta x = \infty$. However, the momentum is known exactly: $\Delta p = 0$. Such a state does satisfy the Heisenberg uncertainty principle when the calculation is done in a limiting fashion.
- A similar calculation shows that the Gaussian state is a state of minimal uncertainty, where $\Delta x \Delta p$ is exactly $\hbar/2$.

Chapter 11

One-dimensional bound states: Potential wells

Reading material for this chapter: Mandl, Chapter 2

11.1 Particle in a box

The simplest one-dimensional potential well is the so-called *particle-in-a-box*. Here, the particle may only move backwards and forwards along a straight line segment $0 < x < L$ with impenetrable barriers at either end. The walls of the one-dimensional box may be visualised as regions of space with an infinitely large potential energy. Conversely, the interior of the box has a constant, zero potential. Thus, no forces act upon the particle inside the box and it can move freely in that region. However, infinitely large forces repel the particle if it touches the walls of the box, preventing it from escaping. The potential in this model is given as

$$\mathcal{U}(x) = \begin{cases} 0, & 0 < x < L, \\ \infty, & \text{otherwise,} \end{cases}$$

(See Fig. 11.1). We now solve the Schrödinger equation for such a system. To keep with tradition, we henceforth use the symbol $\Psi(x)$ for the wavefunction.

Separation of variables

For a time-independent system, the Schrödinger equation reads

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

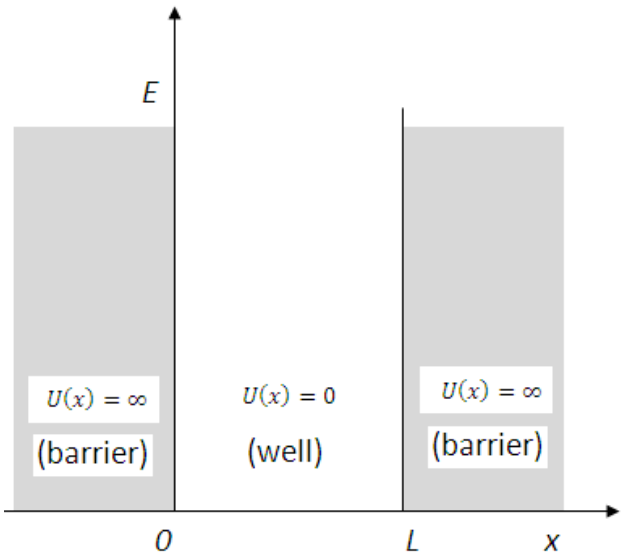


Figure 11.1: Potential well for the particle-in-a-box calculation (infinitely deep well).

Because the right-hand side has no manifest time dependence, we can perform a separation of variables:

$$\Psi(x,t) = T(t)\psi(x).$$

Substitute this trial solution into the Schrödinger equation and divide the result by $T\psi$. The result is

$$\mathrm{i}\hbar\frac{T'(t)}{T} = \frac{-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) + \mathcal{U}(x)\psi(x)}{\psi(x)}.$$

The LHS is a function of t alone, while the RHS is a function of x alone. The only way for this to be true is if LHS $=$ RHS $=$ Const. $:= E$, where E is a constant. Thus,

$$\frac{dT}{dt} = -\mathrm{i}E/\hbar \implies T(t) = T(0)e^{-\mathrm{i}Et/\hbar}.$$

Immediately we see that E has the interpretation of energy. Focus on the space part:

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) + \mathcal{U}(x)\psi(x) = E\psi(x).$$

We have $\hat{H}\psi = E\psi$, which is an **eigenvalue problem**.

An eigenvalue problem

Inside the box, no forces act upon the particle, and $\mathcal{U}(x) = 0$. Thus, in the region $0 < x < L$, we are to solve

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

We take $E > 0$ (shortly we shall see why).

$$k^2 := 2mE/\hbar^2,$$

the solution is

$$\psi(x) = A\sin(kx) + B\cos(kx).$$

The boundary conditions require that the probability current vanish at $x = 0$ and $x = L$, since the probability to find the particle outside of the box is zero. This requirement, together with the continuity of ψ at $x = 0$ and $x = L$, yields $\psi(0) = \psi(L) = 0$, which forces $B = 0$. We are left with

$$\psi(x) = A\sin(kx).$$

However, we still must satisfy

$$\sin(kL) = 0.$$

The only way for this to be true is if $kL = n\pi$, where $n = 1, 2, \cdots$ is a positive integer. Thus, the wavenumber is **quantised**:

$$k^2 = \frac{n^2\pi^2}{L^2} = \frac{2mE_n}{\hbar^2}.$$

Clearly, this quantises the energy, too:

$$E_n = \frac{n^2\pi^2\hbar^2}{2mL^2}.$$

We are left with the following eigenfunctions:

$$\psi_n(x) = A_n \sin\left(\frac{n\pi x}{L}\right),$$

with corresponding eigenvalue $E_n = n^2\pi^2\hbar^2/2mL^2$. The constants A_n are fixed by normalisation:

$$\int_0^L |\psi_n(x)|^2 \mathrm{d}x = 1 \implies A_n = \sqrt{\frac{2}{L}}.$$

In conclusion:

- Allowed states:
$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right);$$
- Corresponding allowed energies:
$$E_n = \frac{n^2\pi^2\hbar^2}{2mL^2};$$

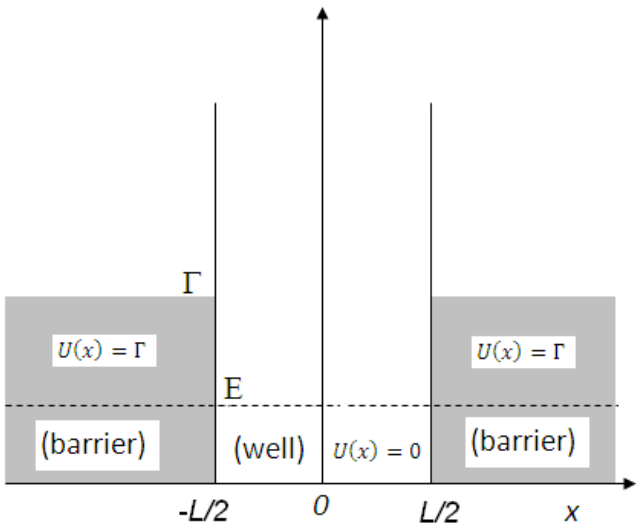


Figure 11.2: Potential well for the particle-in-a-box calculation (finite well depth).

- Full, time-dependent wavefunction:

$$\Psi_n(x,t) = e^{-\mathrm{i}E_n t/\hbar}\psi_n(x).$$

Note: Had we taken $E < 0$, we would have a wavefunction $\psi = A\mathrm{e}^{-|k|x} + B\mathrm{e}^{+|k|x}$, with $k = 2m|E|/\hbar^2$. However, such a choice cannot satisfy the boundary conditions. Similarly, if $k = 0$, then only the trivial solution is possible, $\psi = 0$, which is not normalisable.

11.2 Wells of finite depth

Consider the potential well

$$\mathcal{U}(x) = \begin{cases} 0, & -L/2 < x < L/2, \\ \Gamma, & \text{otherwise,} \end{cases}$$

where Γ is some finite energy (See Fig. 11.2). As before, we solve the eigenvalue problem

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) + \mathcal{U}(x)\psi(x) = E\psi.$$

We break up the solution into three regions:

$$\psi(x) = \begin{cases} \psi_I(x), & \text{if } x < -L/2 \text{ (the region outside the box)} \\ \psi_{II}(x), & \text{if } -L/2 < x < L/2 \text{ (the region inside the box)} \\ \psi_{III}(x) & \text{if } x > L/2 \text{ (the region outside the box).} \end{cases}$$

Let's focus on region II. As before, we solve

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

We take $E > 0$. Hence, the general solution is

$$\psi_{II}(x) = A\sin(kx) + B\cos(kx), \qquad k = \sqrt{2mE}/\hbar.$$

Next, we focus on Region I. Here, $\mathcal{U}(x) = \Gamma$, and Schrödinger's equation reads

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) + \Gamma\psi(x) = E\psi,$$

or

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) = (E - \Gamma)\,\psi.$$

We are going to focus on the **bound state**, where

$$0 < E < \Gamma.$$

Thus, we have the equation

$$\frac{\partial^2}{\partial x^2}\psi(x) = \kappa^2\psi, \qquad \kappa = \sqrt{2m\,(\Gamma - E)\,/\hbar^2} \in \mathbb{R}.$$

The solution is thus

$$\psi_I(x) = \widetilde{C}\mathrm{e}^{-\kappa x} + C\mathrm{e}^{\kappa x}.$$

However, it can not be the case that $\lim_{x\to-\infty}\psi_I(x) = \infty$; we therefore take $\widetilde{C} = 0$ and

$$\psi_I(x) = C\mathrm{e}^{\kappa x}.$$

A similar result holds in Region III:

$$\psi_{III}(x) = D\mathrm{e}^{-\kappa x}.$$

In summary,

$$\psi(x) = \begin{cases} C\mathrm{e}^{\kappa x}, & \text{if } x < -L/2 \\ A\sin(kx) + B\cos(kx), & \text{if } -L/2 < x < L/2 \\ D\mathrm{e}^{-\kappa x} & \text{if } x > L/2, \end{cases}$$

and it remains to fix the constants of integration A, B, C, D. We stipulate that the probability current be everywhere continuous. Thus, ψ and $\partial_x\psi$ must be everywhere continuous; in particular, they must be continuous at $x = \pm L/2$. Thus,

$$\begin{aligned} \psi_I(-L/2) &= \psi_{II}(-L/2), \\ \partial_x\psi_I(-L/2) &= \partial_x\psi_{II}(-L/2), \\ \psi_{II}(L/2) &= \psi_{III}(L/2), \\ \partial_x\psi_{II}(L/2) &= \partial_x\psi_{II}(L/2). \end{aligned}$$

It is tempting solve the resulting four equations by brute force. However, it is simpler to observe that the potential $\mathcal{U}(x)$ is symmetric under $x \to -x$, and thus, the wavefunction must be either even or odd. Let's consider both cases now.

Even case: The even wavefunction satisfies $\psi(x) = \psi(-x)$, and thus $C = D$ and $A = 0$. In other words,

$$\psi(x) = \begin{cases} C\mathrm{e}^{\kappa x}, & \text{if } x < -L/2 \\ B\cos(kx), & \text{if } -L/2 < x < L/2 \\ C\mathrm{e}^{-\kappa x} & \text{if } x > L/2, \end{cases}$$

Matching at $x = -L/2$ gives

$$\begin{aligned} Ce^{-\kappa L/2} &= B\cos(kL/2), \\ \kappa Ce^{-\kappa L/2} &= Bk\sin(kL/2). \end{aligned}$$

Dividing the second equation by the first gives

$$\kappa = k\tan(kL/2),$$

or

$$\kappa(E) - k(E)\tan\left(\frac{k(E)L}{2}\right) = 0,$$

the roots of which give the allowed values of E.

Odd case: The even wavefunction satisfies $\psi(x) = -\psi(-x)$, and thus $D = -C$ and $B = 0$. In other words,

$$\psi(x) = \begin{cases} C\mathrm{e}^{\kappa x}, & \text{if } x < -L/2 \\ A\sin(kx), & \text{if } -L/2 < x < L/2 \\ -C\mathrm{e}^{-\kappa x} & \text{if } x > L/2, \end{cases}$$

Matching at $x = -L/2$ gives

$$\begin{aligned} Ce^{-\kappa L/2} &= -A\sin(kL/2), \\ \kappa Ce^{-\kappa L/2} &= Ak\cos(kL/2). \end{aligned}$$

Dividing the first equation by the second gives

$$\frac{1}{\kappa} = -\frac{1}{k}\tan(kL/2),$$

or

$$\kappa = -k\cot(kL/2).$$

the roots of which give the allowed values of E.

Let us focus in more detail on the **even case**. Call

$$\epsilon := \sqrt{2mE/\hbar^2}\,(L/2)\,, \qquad \gamma := \sqrt{2m\Gamma/\hbar^2}\,(L/2)\,.$$

Then the solvability condition reads

$$\sqrt{\gamma^2 - \epsilon^2} = \epsilon\tan(\epsilon).$$

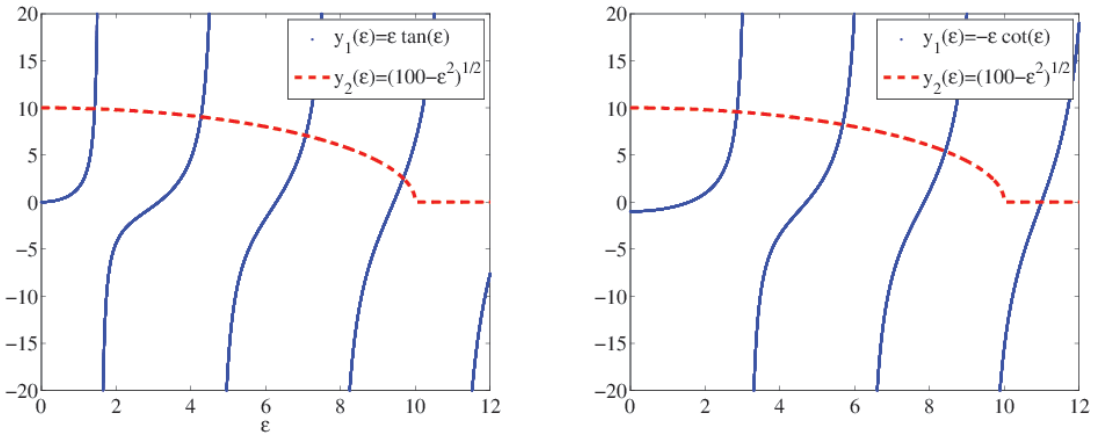
Thus, consider the curves

$$\begin{aligned} y_{1,\mathsf{even}}(\epsilon) &= \epsilon\tan(\epsilon), \\ y_{2,\mathsf{even}}(\epsilon) &= \sqrt{\gamma^2 - \epsilon^2}, \end{aligned}$$

where γ^2 is fixed. The intersection points of these curves give the allowed ϵ-values, hence the allowed values of energy. The corresponding curves for the odd case are

$$\begin{aligned} y_{1,\mathsf{odd}}(\epsilon) &= -\epsilon\cot(\epsilon), \\ y_{2,\mathsf{odd}}(\epsilon) &= \sqrt{\gamma^2 - \epsilon^2}, \end{aligned}$$

These curves are shown in Fig. 11.3. For $\gamma = 10$, there are only seven allowed values of energy.



(a) Even case (b) Odd case

Figure 11.3: Allowed values of energy for a symmetric square well.

Thus, the spectrum is discrete and finite. As γ increases, more and more allowed values of energy become available. However, the spectrum always consists of a discrete number of points.

11.2.1 Asymptotic cases for Fig. 11.3

The case $\gamma \to 0$ In this case, we can show that there is **always at least one bound state**. Focus on the **even case** and Figure 11.3(a). We take $\gamma \to 0$: the circle in the figure has an ever-decreasing radius, while the curve $y_1(\epsilon)$ behaves like a straight line, $y_1 \sim \epsilon$. Such a curve will always intersect in the positive quadrant with a circle centred at the origin, leading to precisely one bound state. On the other hand, taking Fig. 11.3(b), the curve $y_1(\epsilon)$ behaves like $y_1 \sim -1 + (1/2)\epsilon^3 + \mathrm{HOT}$. A quarter-circle of ever-decreasing radius located in the positive quadrant does not intersect this curve, and there is no odd bound state in this limit.

The case $\gamma \to \infty$ Consider the even case first. We are to solve $y_1 = \epsilon \tan\epsilon = y_2 = \infty$. In other words, we must solve $\tan\epsilon = \infty$. This has solution

$$\epsilon = \pi\left(\tfrac{1}{2} + n\right), \qquad n \in \{0, 1, 2, \cdots\}.$$

For the odd case, we must solve $y_1 = -\epsilon\cot\epsilon = y_2 = \infty$. In other words, we must solve $\cot\epsilon = -\infty$, with solution $\epsilon = n\pi$, with $n \in \mathbb{N}$.

The even case now reduces to

$$2\epsilon = \pi\,(1 + 2n) = \pi \times [\text{all odd positive integers}]\,,$$

while the odd case reduces to

$$2\epsilon = 2n\pi = \pi \times [\text{all even positive integers}]\,,$$

Combining both cases, we have

$$2\epsilon = \pi n = \pi \times [\text{all positive integers}]\,,$$

or

$$\epsilon = \tfrac{1}{2}\pi n, \qquad n = 1, 2, \cdots\,,$$

which is precisely the formula for the energy level of a square well of infinite height.

Chapter 12

One-dimensional scattering: Potential barriers

Reading material for this chapter: Mandl, Chapter 2

Consider the situation shown schematically in the figure.

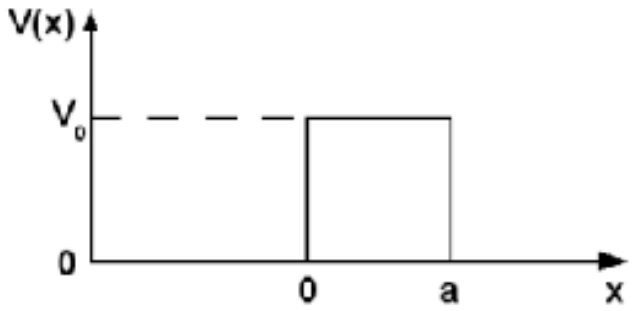


Figure 12.1: Schematic diagram for one-dimensional scattering

Mathematically, this corresponds to a potential-energy landscape

$$\mathcal{U}(x) = \begin{cases} \Gamma, & 0 < x < L, \\ 0, & \text{otherwise}, \end{cases}$$

and we have the following regions:

- Region I, $x < 0$.
- Region II, $0 < x < L$.
- Region III, $x > L$.

We assume that particles are incident on the barrier from $x = -\infty$, and have energy $E < \Gamma$. We determine if any particles are to be found in region III. We must allow for such a scenario *a priori*. Thus, we have the wavefunctions of Fig. 12.2.

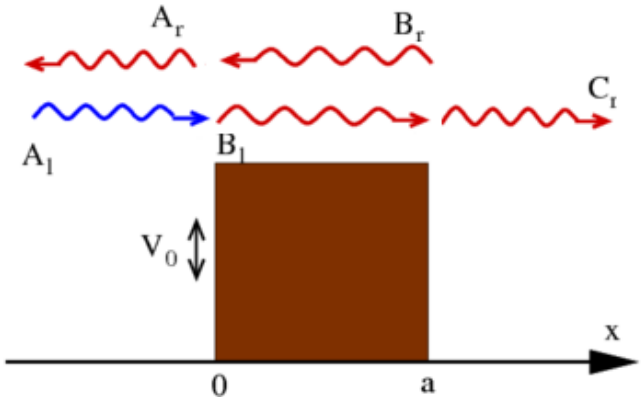


Figure 12.2: Boundary conditions for the scattering problem

It remains to compute these wavefunctions, and to see if the wavefunction in region III $x > L$ is nonzero. We first of all note that there are no left-travelling waves in this region, since no waves are reflected at $x = +\infty$. We now focus on solving the Schrödinger equation.

Region I Here $\mathcal{U} = 0$, and the wavefunction is a plane wave:

$$\psi_I = A_i \mathrm{e}^{\mathrm{i}kx} + A_r \mathrm{e}^{-\mathrm{i}kx},$$

where $k = \sqrt{2mE}/\hbar$.

Region II Here $\mathcal{U} = \Gamma$. The eigenvalue problem reads

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

For $E < \Gamma$, the solution reads

$$\psi_{II}(x) = B_i \mathrm{e}^{\kappa x} + B_r \mathrm{e}^{-\kappa x},$$

where

$$\kappa = \sqrt{2m(\Gamma - E)}/\hbar.$$

Region III Here $\mathcal{U} = 0$ and the wavefunction is again a plane wave. However, there is no possibility of reflection, so any component of the wavefunction reaching this region has to be a right-travelling one:

$$\psi_{III}(x) = C\mathrm{e}^{\mathrm{i}kx}.$$

Instead of computing the coefficients A_i, A_r, B_i, B_r, C, we focus instead on probability currents. In region I, the incident probability current reads

$$\begin{aligned} J_i &= \frac{\hbar}{2m\mathrm{i}}\left[\left(A_i\mathrm{e}^{\mathrm{i}kx}\right)^*\partial_x\left(A_i\mathrm{e}^{\mathrm{i}kx}\right) - \left(A_i\mathrm{e}^{\mathrm{i}kx}\right)\partial_x\left(A_i\mathrm{e}^{\mathrm{i}kx}\right)^*\right], \\ &= \frac{\hbar}{2m\mathrm{i}}\left(\mathrm{i}k|A_i|^2 + \mathrm{i}k|A_i|^2\right), \\ &= \frac{\hbar k}{m}|A_i|^2. \end{aligned}$$

Similarly, the reflected current in region I is

$$J_r = \frac{\hbar k}{m}|A_r|^2,$$

and the transmitted current in region III is

$$J_t = \frac{\hbar k}{m}|C|^2$$

Conservation of probability requires that

$$J_i = J_r + J_t.$$

We define the following **reflection coefficient**

$$R = \frac{J_r}{J_i},$$

and **transmission coefficient**

$$T = \frac{J_t}{J_i};$$

conservation of probability yields $R + T = 1$. We compute these coefficients now.

First, we match ψ and its derivative at $x = L$:

$$\begin{aligned} \psi_{II}(L) &= \psi_{III}(L), \\ \partial_x\psi_{II}(L) &= \partial_x\psi_{III}(L), \end{aligned}$$

In other words,

$$\begin{aligned} B_i \mathrm{e}^{\kappa L} + B_r \mathrm{e}^{-\kappa L} &= C \mathrm{e}^{\mathrm{i}kL}, \\ \kappa B_i \mathrm{e}^{\kappa L} - \kappa B_r \mathrm{e}^{-\kappa L} &= \mathrm{i}k C \mathrm{e}^{\mathrm{i}kL}. \end{aligned}$$

But this is a matrix problem:

$$\begin{pmatrix} \mathrm{e}^{\kappa L} & \mathrm{e}^{-\kappa L}, \\ \kappa \mathrm{e}^{\kappa L} & -\kappa \mathrm{e}^{-\kappa L} \end{pmatrix} \begin{pmatrix} B_i \\ B_r \end{pmatrix} = C \mathrm{e}^{\mathrm{i}kL} \begin{pmatrix} 1 \\ \mathrm{i}k \end{pmatrix}.$$

The matrix has determinant -2κ and hence,

$$\begin{pmatrix} B_i \\ B_r \end{pmatrix} = \frac{-1}{2\kappa} \begin{pmatrix} -\kappa \mathrm{e}^{-\kappa L} & -\mathrm{e}^{-\kappa L}, \\ -\kappa \mathrm{e}^{\kappa L} & \mathrm{e}^{\kappa L} \end{pmatrix} C \mathrm{e}^{\mathrm{i}kL} \begin{pmatrix} 1 \\ \mathrm{i}k \end{pmatrix},$$

or

$$\begin{pmatrix} B_i \\ B_r \end{pmatrix} = \frac{C}{2\kappa} \mathrm{e}^{\mathrm{i}kL} \begin{pmatrix} \kappa \mathrm{e}^{-\kappa L} & \mathrm{e}^{-\kappa L}, \\ \kappa \mathrm{e}^{\kappa L} & -\mathrm{e}^{\kappa L} \end{pmatrix} \begin{pmatrix} 1 \\ \mathrm{i}k \end{pmatrix}.$$

In other words,

$$\begin{aligned} B_i &= \frac{C \mathrm{e}^{\mathrm{i}kL}}{2\kappa} \mathrm{e}^{-\kappa L} \left(\kappa + \mathrm{i}k\right), \\ B_r &= \frac{C \mathrm{e}^{\mathrm{i}kL}}{2\kappa} \mathrm{e}^{+\kappa L} \left(\kappa - \mathrm{i}k\right) \end{aligned}$$

Next, we match ψ and its derivative at $x = 0$:

$$\begin{aligned} \psi_I(0) &= \psi_{II}(0), \\ \partial_x \psi_I(0) &= \partial_x \psi_{II}(0), \end{aligned}$$

In other words,

$$\begin{aligned} A_i + A_r &= B_i + B_r, \\ \mathrm{i}kA_i - \mathrm{i}kA_r &= \kappa B_i - \kappa B_r. \end{aligned}$$

Multiply the first equation by $\mathrm{i}k$ and add the two resulting equations. Obtain

$$2\mathrm{i}kA_i = B_i \left(\mathrm{i}k + \kappa\right) + B_r \left(\mathrm{i}k - \kappa\right).$$

But we know what B_i and B_r are:

$$2k\mathrm{i}A_i = C\mathrm{e}^{\mathrm{i}kL}\frac{\mathrm{i}k+\kappa}{2\kappa}e^{-\kappa L}\left(\kappa+\mathrm{i}k\right) + C\mathrm{e}^{\mathrm{i}kL}\frac{\mathrm{i}k-\kappa}{2\kappa}e^{\kappa L}\left(\kappa-\mathrm{i}k\right)$$

Hence,

$$\begin{aligned}
\mathrm{i}\frac{A_i}{C}\mathrm{e}^{-\mathrm{i}kL} &= \frac{1}{4k\kappa}e^{-\kappa L}\left(\kappa+\mathrm{i}k\right)^2 - \frac{1}{4k\kappa}e^{\kappa L}\left(\kappa-\mathrm{i}k\right)^2, \\
\mathrm{i}\frac{A_i}{C}4\kappa k\mathrm{e}^{-\mathrm{i}kL} &= \kappa^2\left(\mathrm{e}^{-\kappa L}-\mathrm{e}^{\kappa L}\right) + 2\mathrm{i}k\kappa\left(\mathrm{e}^{\kappa L}+\mathrm{e}^{-\kappa L}\right) + k^2\left(\mathrm{e}^{\kappa L}-\mathrm{e}^{-\kappa L}\right), \\
&= 2\sinh(\kappa L)\left(k^2-\kappa^2\right) + 4\mathrm{i}k\kappa\cosh(\kappa L).
\end{aligned}$$

Hence,

$$\begin{aligned}
\left|\frac{A_i}{C}\right|^2\left(4k\kappa\right)^2 &= 4\sinh^2(\kappa L)\left(k^2-\kappa^2\right)^2 + \left(4k\kappa\right)^2\cosh^2(\kappa L), \\
&= 4\sinh^2(\kappa L)\left(k^2-\kappa^2\right)^2 + \left(4k\kappa\right)^2\left[1+\sinh^2\left(\kappa L\right)\right], \\
&= 4\sinh^2(\kappa L)\left(k^2+\kappa^2\right)^2 + \left(4k\kappa\right)^2.
\end{aligned}$$

Inverting,

$$\begin{aligned}
\left|\frac{C}{A_i}\right|^2 &= \frac{\left(4k\kappa\right)^2}{4\sinh^2(\kappa L)\left(k^2+\kappa^2\right)^2 + \left(4k\kappa\right)^2}, \\
&= \frac{\left(2k\kappa\right)^2}{\sinh^2(\kappa L)\left(k^2+\kappa^2\right)^2 + \left(2k\kappa\right)^2}, \\
&= T.
\end{aligned}$$

Thus,

$$T = \frac{\left(2k\kappa\right)^2}{\sinh^2(\kappa L)\left(k^2+\kappa^2\right)^2 + \left(2k\kappa\right)^2}.$$

The reflection coefficient can be calculated in a similar way, or using $R = 1 - T$. In any case,

$$R = \frac{\sinh^2(\kappa L)\left(k^2+\kappa^2\right)^2}{\sinh^2(\kappa L)\left(k^2+\kappa^2\right)^2 + \left(2k\kappa\right)^2}$$

The non-dimensional form for T, using $\gamma = L\sqrt{2m\Gamma}/\hbar$ and $\epsilon = L\sqrt{2mE}/\hbar$ reads

$$T(\epsilon;\gamma) = \frac{4|\gamma^2-\epsilon^2|\epsilon^2}{\sinh^2\left(\sqrt{\gamma^2-\epsilon^2}\right)\gamma^4 + 4|\gamma^2-\epsilon^2|\epsilon^2}.$$

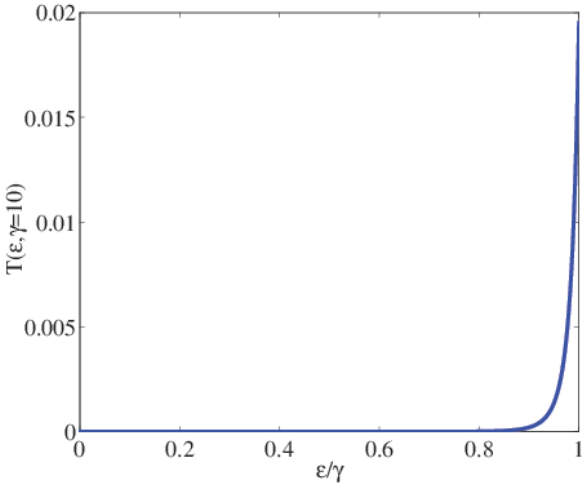


Figure 12.3: Transmission coefficient for particle energies less than the barrier height. The result is non-zero: a portion of the particles pass through the barrier.

A plot of $T(\epsilon; \gamma = 10)$ is shown in Fig. 12.3. The result is NOT identically zero: particles pass through the barrier.

These results call for some discussion. Consider a stream of classical particles incident on the barrier from $x = -\infty$. In region I, we have

$$\frac{p_\infty^2}{2m} = \tfrac{1}{2}mv_\infty^2 = E_I \geq 0.$$

We are interested in the case $E = E_I < \Gamma$. Suppose that the particle enters region II, with velocity v_0. Then

$$\tfrac{1}{2}mv_0^2 + \Gamma = E \implies \tfrac{1}{2}mv_0^2 = E - \Gamma < 0,$$

which is impossible. Therefore, we conclude that all the particles remain in region I, and are thus reflected off the potential barrier. In other words, no particles are transmitted into region III, and

$$T_{\text{classical}}\,(E < \Gamma) = 0.$$

For $E = E_I > \Gamma$, no such restriction exists, and all the particles pass into region III:

$$T_{\text{classical}}\,(E > \Gamma) = 1.$$

Thus, the fact that $T_{\text{QM}}\,(E < \Gamma) \neq 0$ is a remarkable, anti-classical result. Particles, ghostlike, pass through a barrier. This is called **quantum tunnelling**.

Chapter 13

The harmonic oscillator

Reading material for this chapter: Mandl, Chapter 2

We study the Schrödinger equation for the celebrated potential

$$\mathcal{U}(x) = \tfrac{1}{2}k_0 x^2,$$

in one dimension. The Hamiltonian is

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \tfrac{1}{2}k_0 x^2,$$

which is manifestly time independent. Thus, we can separate out the space and time dependence, and solve the eigenvalue problem

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \tfrac{1}{2}k_0 x^2\right)\psi(x) = E\psi(x).$$

13.1 Asymptotic solution

Before attempting a solution, we study the asymptotic behaviour, $|x| \to \infty$. Then, the ODE to solve looks like

$$\frac{\hbar^2}{2m}\psi'' = \tfrac{1}{2}k_0 x^2\psi := \tfrac{1}{2}m\omega^2 x^2\psi.$$

In other words,

$$\psi'' = \frac{m^2\omega^2}{\hbar^2}x^2\psi.$$

This implies a standard unit of length in the problem,

$$a = \sqrt{\frac{\hbar}{m\omega}},$$

and the asymptotic problem can be conveniently re-written as

$$\psi'' \sim \frac{x^2}{a^4}\psi.$$

This form suggests that we re-write the trial solution as

$$\psi(x) = h(x)e^{-x^2/2a^2},$$

since then,

$$\psi''(x) = \frac{x^2}{a^4}h(x)e^{-x^2/2a^2} + \left(h''(x) - \frac{2x}{a^2}h'(x) - \frac{h(x)}{a^2}\right)e^{-x^2/2a^2},$$

and we have captured the leading-order term in the approximation.

13.2 The solution

We substitute

$$\psi(x) = h(x)e^{-x^2/2a^2}$$

into the full eigenvalue problem

$$-\psi''(x) + \frac{x^2}{a^4}\psi(x) = k^2\psi(x), \qquad k^2 = \frac{2mE}{\hbar^2}.$$

We obtain the expression

$$-\left[\frac{x^2}{a^4}h(x) + h''(x) - \frac{2x}{a^2}h'(x) - \frac{h(x)}{a^2}\right]\mathrm{e}^{-x^2/2a^2} + \frac{x^2}{a^2}h(x)\mathrm{e}^{-x^2/2a^2} = k^2h(x)\mathrm{e}^{-x^2/2a^2}.$$

Tidying up, we have the following differential equation:

$$h''(x) - \frac{2x}{a^2}h'(x) + \left(k^2 - \frac{1}{a^2}\right)h(x) = 0.$$

We introduce the non-dimensional variable $s = x/a$. In terms of this variable, the differential equation to solve reads

$$\frac{d^2h}{ds^2} - 2s\frac{dh}{ds} + 2nh(s) = 0, \qquad 2n = k^2a^2 - 1$$

(This ODE is considered on p. 820 of Arfken and Weber). We propose a power-series solution for

this equation:

$$h(s) = \sum_{p=0}^{\infty} a_p x^p.$$

Substituting into the ODE, we obtain

$$\sum_{p=0}^{\infty} p\,(p-1)\,a_p x^{p-2} - \sum_{p=0}^{\infty} 2p a_p x^p + 2n \sum_{p=0}^{\infty} a_p x^p = 0,$$

or

$$\sum_{p=2}^{\infty} p\,(p-1)\,a_p x^{p-2} - \sum_{p=1}^{\infty} 2p a_p x^p + 2n \sum_{p=0}^{\infty} a_p x^p = 0.$$

We call $q = p - 2$ in the first sum, hence $p = q + 2$.

$$\sum_{q=0}^{\infty} (q+2)(q+1) a_{q+2} x^q - \sum_{p=0}^{\infty} (2n - 2p)\, a_p x^p = 0.$$

However, q is a dummy variable, so we let $q \to p$, and we end up with

$$\sum_{p=0}^{\infty} (p+2)(p+1) a_{p+2} x^p - \sum_{p=0}^{\infty} 2\,(n-p)\, a_p x^p = 0,$$

or

$$\sum_{p=0}^{\infty} \left[(p+2)(p+1) a_{p+2} + 2(n-p) a_p\right] x^p = 0.$$

The power series is identically zero, so each term must be zero. In other words,

$$(p+2)(p+1) a_{p+2} + 2(n-p) a_p = 0, \qquad p = 0, 1, \cdots,$$

hence

$$a_{p+2} = \frac{2(p-n)}{(p+2)(p+1)} a_p.$$

This splits into odd and even cases.

Even case: We take a_0 to be a constant of integration and $a_1 = 0$. Then $a_1 = a_3 = a_5 = \cdots = 0$, and

$$a_{p+2} = 2a_p \frac{p-n}{(p+1)(p+2)}, \qquad (*)$$

hence

$$h_{\text{even}}(s) = a_0 \left[1 + \frac{2(-n)s^2}{2!} + \frac{2^2(-n)(2-n)s^4}{4!} + \cdots\right].$$

Similarly for the **odd case**, where we take a_1 to be a constant of integration and $a_0 = 0$. Thus, $a_0 = a_2 = a_4 = \cdots = 0$, and

$$a_{p+2} = 2a_p \frac{p+1-n}{(p+2)(p+3)}, \qquad (**)$$

hence

$$h_{\text{odd}}(s) = a_1 \left[1 + \frac{2(1-n)s}{3!} + \frac{2^2(1-n)(3-n)s^3}{5!} + \cdots \right].$$

We must examine the asymptotic behaviour of these solutions. For large indices, the ratio of successive terms in the even solution is

$$\frac{a_{p+2}}{a_p} \sim \frac{2}{p},$$

suggesting that $h_{\text{even}}(s) \sim e^{2s^2}$ as $|s| \to \infty$. Similarly, the odd solution behaves as se^{2s^2} as $|s| \to \infty$. THIS IS A DISASTER, since the solution must vanish as $|x| \to \infty$. However, help is at hand: if we insist that n be an integer, then the recursions (*) and (**) terminate, and $h(s)$ is a polynomial. These are called the **Hermite polynomials**, the first few of which are given in standard form here:

$$\begin{aligned}
H_0(s) &= 1, \\
H_1(s) &= 2s, \\
H_2(s) &= 4s^2 - 2, \\
H_3(s) &= 8s^3 - 12s, \\
H_4(s) &= 16s^4 - 48s^2 + 12, \\
H_5(s) &= 32s^5 - 160s^3 + 120s, \\
H_6(s) &= 64s^6 - 480s^4 + 720s^2 - 120, \\
H_7(s) &= 128s^7 - 1344s^5 + 3360s^3 - 1680s, \\
H_8(s) &= 256s^8 - 3584s^6 + 13440s^4 - 13440s^2 + 1680, \\
H_9(s) &= 512s^9 - 9216s^7 + 48384s^5 - 80640s^3 + 30240s, \\
H_{10}(s) &= 1024s^{10} - 23040s^8 + 161280s^6 - 403200s^4 + 302400s^2 - 30240,
\end{aligned}$$

and are plotted in Fig. 13.1. These polynomials can be generated from the following property:

$$H_n(s) = (-1)^n e^{s^2} \frac{d^n}{ds^n} e^{-s^2} = e^{s^2/2} \left(s - \frac{d}{ds} \right)^n e^{-s^2/2}$$

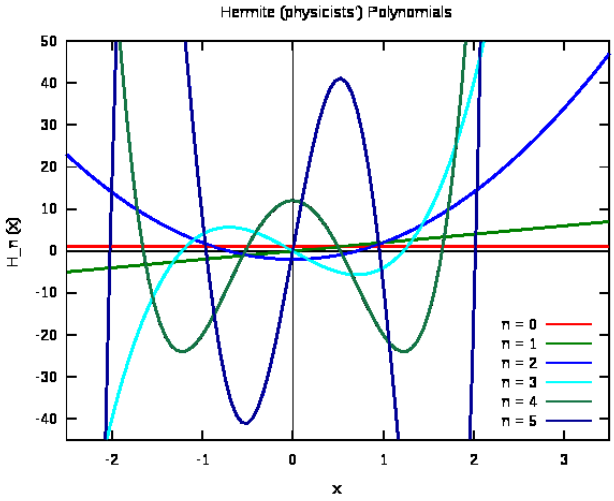


Figure 13.1: The first six Hermite polynomials

Happily, these functions are orthogonal with respect to a weight:

$$\begin{aligned}\int_{-\infty}^{\infty} H_n(s)H_m(s)\,\mathrm{e}^{-s^2}\,\mathrm{d}s &= 0, \qquad n \neq m,\\ \int_{-\infty}^{\infty} H_n(s)H_n(s)\,\mathrm{e}^{-s^2}\,\mathrm{d}s &= n!\,2^n\sqrt{\pi}.\end{aligned}$$

Putting these results together, the **normalised** wavefunctions of the Schrödinger equation are

$$\psi_n(x) = \frac{1}{\sqrt{n!2^n}\pi^{1/4}a^{1/2}} H_n(x/a)e^{-x^2/2a^2}, \qquad a = \sqrt{\hbar/m\omega}$$

where $n = 0, 1, 2, \cdots$ are necessarily integers.

But recall

$$\begin{aligned}2n+1 &= a^2k^2,\\ 2n+1 &= \frac{\hbar}{m\omega}\frac{2mE}{\hbar^2},\end{aligned}$$

hence

$$E_n = \hbar\omega\left(n+\tfrac{1}{2}\right), \qquad n = 0, 1, 2, \cdots,$$

and the energy is quantised.

The first few normalised wavefunctions are shown in Figure 13.2.

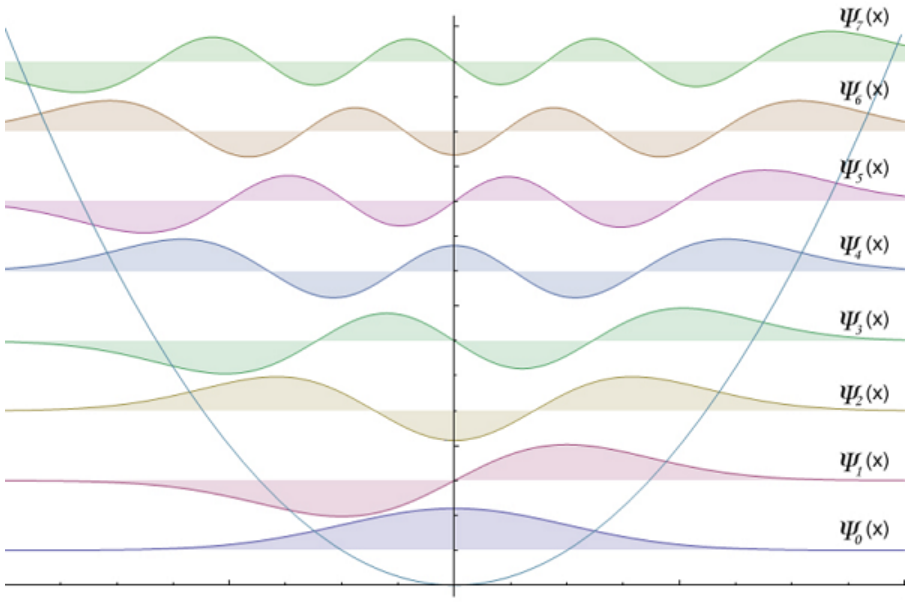


Figure 13.2: Normalised states of the harmonic oscillator.

From now on the useful shorthand

$$N_n := \frac{1}{\sqrt{n!2^n}\pi^{1/4}a^{1/2}}$$

will be used, such that

$$\psi_n(x) = N_n H_n(x/a)e^{-x^2/2a^2}.$$

13.3 Creation and annihilation

We define the operators

$$\hat{\mathrm{a}}_+ = \frac{1}{\sqrt{2}a}\left(\hat{x} - \frac{\mathrm{i}}{m\omega}\hat{p}\right),$$

and its adjoint

$$\hat{\mathrm{a}}_- = \frac{1}{\sqrt{2}a}\left(\hat{x} + \frac{\mathrm{i}}{m\omega}\hat{p}\right),$$

(NOTE THE SIGNS!) such that

$$\hat{\mathrm{a}}_- = \hat{\mathrm{a}}_+^\dagger.$$

In the position representation, $\widehat{\mathrm{p}} = -\mathrm{i}\hbar\partial_x$, and

$$\hat{\mathrm{a}}_+ = \frac{1}{\sqrt{2}a}\left(x - \frac{\hbar}{m\omega}\partial_x\right),$$
$$\hat{\mathrm{a}}_- = \frac{1}{\sqrt{2}a}\left(x + \frac{\hbar}{m\omega}\partial_x\right),$$

We act on the eigenstate $\psi_n(x)$ with the annihilation operator. The following **recursion relation** is useful (Arfken and Weber, p. 817)

$$\partial_s H_n(s) = 2nH_{n-1}(s).$$

Thus,

$$\begin{aligned}
\hat{\mathrm{a}}_-\psi_n(x) &= \frac{1}{\sqrt{2}a}\left(x + a^2\partial_x\right)\psi_n(x), \\
&= \frac{1}{\sqrt{2}a}x\psi_n(x) + \frac{1}{\sqrt{2}a}a^2N_n\partial_x\left[H_n(x/a)e^{-x^2/2a^2}\right], \\
&= \frac{1}{\sqrt{2}a}x\psi_n(x) + \frac{1}{\sqrt{2}a}a^2N_nH_n(x/a)\left(-x/a^2\right)e^{-x^2/2a^2} + \frac{1}{\sqrt{2}a}a^2N_ne^{-x^2/2a^2}\frac{1}{a}\partial_s H_n(s), \\
&= \frac{1}{\sqrt{2}a}N_ne^{-x^2/2a^2}2nH_{n-1}(s), \\
&= \frac{1}{\sqrt{2}}\frac{2nN_n}{N_{n-1}}N_{n-1}H_{n-1}(s)e^{-x^2/2a^2}, \ \ldots\ldots \ \frac{N_n}{N_{n-1}} = \tfrac{1}{\sqrt{2}\sqrt{n}}, \\
&= \sqrt{n}\psi_{n-1}(x).
\end{aligned}$$

Similarly, using

$$H_{n+1}(s) = 2sH_n(s) - 2nH_{n-1}(s)$$

we obtain

$$\hat{\mathrm{a}}_+\psi_n(x) = \sqrt{n+1}\psi_{n+1}(x).$$

We can also show that

$$\hat{\mathrm{a}}_+\hat{\mathrm{a}}_-\psi_n(x) = n\psi_n(x).$$

Thus,

$$\widehat{\mathrm{N}} := \hat{\mathrm{a}}_+\hat{\mathrm{a}}_-$$

is the number operator, and tells us how many quanta of energy are in the system. (Note the order: we MUST act with the annihilation operator BEFORE acting with the creation operator) We have the following interpretations:

- The operator $\hat{\mathrm{a}}_+$ 'creates' a quantum of energy;

- The operator $\hat{\mathrm{a}}_{-}$ 'destroys' a quantum of energy;
- The operator $\widehat{\mathrm{N}}$ tells us how many quanta of energy are in the system.

Chapter 14

The Schrödinger equation of the hydrogen atom

Reading material for this chapter: Mandl, Chapter 2

We study the Schrödinger equation for a system of two particles that attract each other under the potential

$$\mathcal{U}(\boldsymbol{r}) = -\frac{k_0 e^2}{|\boldsymbol{r}|},$$

in three dimensions. We assume that one particle (the positively charged 'proton', charge $+e$) is much more massive than the other particle (the negatively charged 'electron', charge $-e$), and treat the proton as a fixed force centre, with origin O. This assumption can be made rigorous using the definition of **reduced mass**, although we do not pursue this here. The positive constant k_0 is introduced to take care of any prefactors arising from particular choices of physical units (e.g. $k_0 = 1/4\pi\epsilon_0$ in SI). The Hamiltonian for the electron reads

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 - \frac{k_0 e^2}{r},$$

where $\boldsymbol{r} = (x, y, z)$ is the electron's position and $r = |\boldsymbol{r}|$. This is manifestly time independent, and we therefore separate out the space and time dependence, and solve the eigenvalue problem

$$\left(-\frac{\hbar^2}{2m}\nabla^2 - \frac{k_0 e^2}{r}\right)\psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}).$$

This is accomplished using the separation-of-variables technique.

14.1 Separation of variables

First, we perform some scaling arguments. We multiply the Schrödinger equation by $2m/\hbar^2$ to obtain

$$-\nabla^2\psi - \frac{2mk_0e^2}{\hbar^2}\frac{1}{r}\psi(\boldsymbol{r}) = -k^2\psi(\boldsymbol{r}), \qquad -k^2 = 2mE/\hbar^2.$$

We identify a lengthscale

$$a_0 := \frac{\hbar^2}{mk_0e^2},$$

and focus on solving

$$-\nabla^2\psi - \frac{2}{a_0r}\psi = -k^2\psi. \qquad (*)$$

We are searching for bound states ($E < 0$), and it is therefore useful to write the RHS as $-k^2$, an inherently negative quantity.

Because the potential is spherically symmetric, we solve the problem in spherical polar coordinates. In this system, the Laplacian has the form

$$\nabla^2\psi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\varphi^2},$$

where θ is the polar angle and φ is the azimuthal angle (Figure 14.1)[1].

[1]Note the convention: θ is for the polar angle, and φ is for the azimuthal angle. Some authors use the opposite convention. Pay no attention to them!

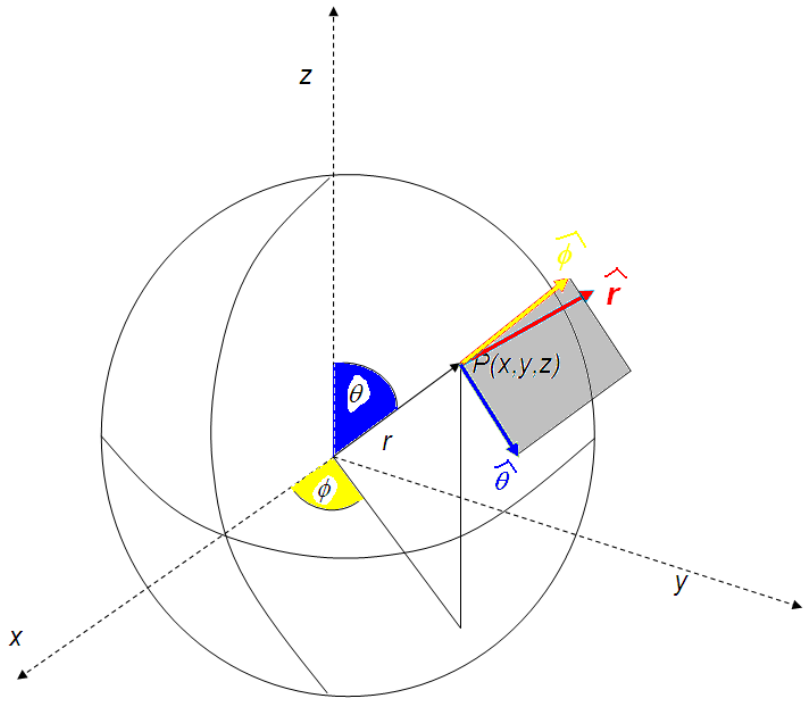


Figure 14.1: Spherical polar coordinates

Substituting this form into the Schrödinger equation (*), we obtain

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\varphi^2}+\frac{2}{a_0 r}\psi=k^2\psi.$$

We attempt a separation-of-variables solution:

$$\psi(r,\theta,\varphi)=R(r)\Theta(\theta)\Phi(\varphi).$$

We substitute this trial solution into the PDE and divide the answer by $R(r)\Theta(\theta)\Phi(\varphi)$. Obtain

$$\frac{1}{R}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right)+\frac{2}{a_0 r}+\frac{1}{\Theta}\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)+\frac{1}{\Phi}\frac{1}{r^2\sin^2\theta}\frac{\partial^2\Phi}{\partial\varphi^2}=k^2.$$

Multiply up by r^2:

$$\frac{1}{R}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right)+\frac{2r}{a_0}-k^2r^2+\frac{1}{\Theta}\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)+\frac{1}{\Phi}\frac{1}{\sin^2\theta}\frac{\partial^2\Phi}{\partial\varphi^2}=0.$$

In other words,

$$\frac{1}{R}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right)+\frac{2r}{a_0}-k^2r^2=-\left[\frac{1}{\Theta}\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)+\frac{1}{\Phi}\frac{1}{\sin^2\theta}\frac{\partial^2\Phi}{\partial\varphi^2}\right].$$

Now the LHS is a function of r alone and the RHS is a function of angles. The only way for this to hold is if both sides are constant:

$$\begin{aligned} \frac{1}{R}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right)+\frac{2r}{a_0}-k^2r^2 &= L^2, \\ \frac{1}{\Theta}\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)+\frac{1}{\Phi}\frac{1}{\sin^2\theta}\frac{\partial^2\Phi}{\partial\varphi^2} &= -L^2, \end{aligned}$$

where L^2 is a constant whose sign is yet to be determined.

Let's take the second of these equations and multiply it by $\sin^2\theta$. The result is

$$\frac{1}{\Theta}\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)+L^2\sin^2\theta+\frac{1}{\Phi}\frac{\partial^2\Phi}{\partial\varphi^2}=0.$$

Re-arranging,

$$\frac{1}{\Theta}\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)+L^2\sin^2\theta=-\frac{1}{\Phi}\frac{\partial^2\Phi}{\partial\varphi^2}.$$

This forces $LHS = RHS =$ Const., which we call $-m^2$:

$$\begin{aligned} \frac{1}{\Theta}\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)+L^2\sin^2\theta &= m^2, \\ -\frac{1}{\Phi}\frac{\partial^2\Phi}{\partial\varphi^2} &= m^2. \end{aligned}$$

Let's take the second of these equations:

$$\frac{\partial^2\Phi}{\partial\varphi^2}+m^2\Phi=0,$$

with solution

$$\Phi=\mathrm{e}^{\pm\mathrm{i}\varphi}.$$

We require a continuous, single-valued wavefunction:

$$\Phi(\varphi)=\Phi(\varphi+2\pi),$$

hence

$$m\in\mathbb{Z}.$$

We return to

$$\frac{1}{\Theta}\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)+L^2\sin^2\theta=m^2,$$

or

$$\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)+\left(L^2\sin^2\theta-m^2\right)\Theta=0.$$

We are going to introduce a new variable

$$x = \cos\theta.$$

Hence,

$$\frac{d}{d\theta} = \frac{dx}{d\theta}\frac{d}{dx} = -\sin\theta\frac{d}{dx},$$

and the ODE becomes

$$\begin{aligned}
\sin\theta\left(-\sin\theta\frac{d}{dx}\right)\left[\sin\theta\left(-\sin\theta\frac{d}{dx}\right)\Theta\right] + \left(L^2\sin^2\theta - m^2\right)\Theta &= 0,\\
\sin^2\theta\frac{d}{dx}\left(\sin^2\theta\frac{d\Theta}{dx}\right) + \left(L^2\sin^2\theta - m^2\right)\Theta &= 0,\\
\frac{d}{dx}\left(\sin^2\theta\frac{d\Theta}{dx}\right) + \left(L^2 - \frac{m^2}{\sin^2\theta}\right)\Theta &= 0,\\
\frac{d}{dx}\left[\left(1-x^2\right)\frac{d\Theta}{dx}\right] + \left(L^2 - \frac{m^2}{1-x^2}\right)\Theta &= 0;
\end{aligned}$$

finally,

$$\left(1-x^2\right)\frac{d^2\Theta}{dx^2} - 2x\frac{d\Theta}{dx} + \left(L^2 - \frac{m^2}{1-x^2}\right)\Theta = 0. \tag{14.1}$$

14.2 Polynomial solutions

As in the harmonic oscillator case, we implement a power series solution for Eq. (14.1) and impose normalisatbility. The result is a set of polynomials called the **associated Legendre polynomials**. The quantity L must be quantised to get a normalisable solution:

$$\begin{aligned}
L^2 &= \ell(\ell+1),\\
\ell &= 0, 1, 2, \cdots,\\
m &= -\ell, -\ell+1, \cdots, \ell-1, \ell.
\end{aligned}$$

Thus, the Legendre polynomials depend on two integer indices, ℓ and m. The first few such polynomials are shown below: (Arfken and Weber, p. 733):

$$P_0^0(x) = 1,$$

$$\begin{aligned} P_1^{-1}(x) &= -\tfrac{1}{2}P_1^1(x), \\ P_1^0(x) &= x, \\ P_1^1(x) &= (1-x^2)^{1/2}, \end{aligned}$$

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

Happily, there is a general expression for these polynomials (A&W, p. 782):

$$P_\ell^m(x) = \frac{1}{2^\ell \ell!}(1-x^2)^{m/2}\,\frac{d^{\ell+m}}{dx^{\ell+m}}(x^2-1)^\ell,$$

and they are orthogonal (A&W, p. 776):

$$\int_{-1}^{1} P_p^m(x)P_\ell^m(x)\mathrm{d}x = \frac{2}{2\ell+1}\frac{(\ell+m)!}{(\ell-m)!}\delta_{p\ell}, \qquad 0 \le m \le \ell$$

(same m-value in each polynomial to be integrated!). Combining the polar and the azimuthal dependencies, we have the following solutions for the angular part of the wavefunction:

$$\Theta(\theta)\Phi(\varphi) = P_\ell^m(\cos\theta)e^{\mathrm{i}m\varphi},$$

where

$$\begin{aligned} \ell &= 0, 1, 2, \cdots, \\ m &= -\ell, -\ell+1, \cdots, \ell-1, \ell. \end{aligned}$$

For convenience, these angular functions are combined together as **spherical harmonics** (A&W, p. 788):

$$Y_\ell^m(\theta,\varphi) = (-1)^m\sqrt{\frac{(2\ell+1)}{4\pi}\frac{(\ell-m)!}{(\ell+m)!}}\,P_\ell^m(\cos\theta)\,\mathrm{e}^{\mathrm{i}m\varphi},$$

which satisfy the **orthogonality relation**

$$\int_{\theta=0}^{\pi}\int_{\varphi=0}^{2\pi} Y_\ell^m(\theta,\varphi)\, Y_{\ell'}^{m'*}(\theta,\varphi)\,\mathrm{d}\Omega = \delta_{\ell\ell'}\,\delta_{mm'},$$

where

$$\mathrm{d}\Omega = \sin\theta \mathrm{d}\theta\,\mathrm{d}\varphi$$

is the element of solid angle. This is the thing to remember!

Finally, we revisit the radial equation:

$$\begin{aligned}\frac{1}{R}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right)+\frac{2r}{a_0}-k^2r^2 &= L^2=\ell(\ell+1),\\ \frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right)+\left[\frac{2r}{a_0}-k^2r^2-\ell(\ell+1)\right]R(r) &= 0.\end{aligned}$$

We try one more trick:

$$u(r)=R(r)r \implies R(r)=u(r)/r.$$

Then

$$\begin{aligned}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) &= \frac{d}{dr}\left(r^2\frac{d}{dr}\frac{u}{r}\right),\\ &= \frac{d}{dr}\left[r^2\left(\frac{u'}{r}-\frac{u}{r^2}\right)\right],\\ &= \frac{d}{dr}\left(u'r-u\right),\\ &= ru''r+u'-u',\\ &= ru''.\end{aligned}$$

Thus

$$r\frac{d^2u}{dr^2}+\left[\frac{2r}{a_0}-k^2r^2-\ell(\ell+1)\right]\frac{u}{r}=0,$$

or

$$\frac{d^2u}{dr^2}=\left[k^2-\frac{2}{a_0r}+\frac{\ell(\ell+1)}{r^2}\right]u.$$

We are interested in bounds states ($E<0$, $k^2>0$). Thus, define a new dimensionless measure of distance,

$$s=kr.$$

The ODE then reads

$$\frac{d^2u}{ds^2}=\left[1-\frac{2}{a_0k}\frac{1}{s}+\frac{\ell(\ell+1)}{s^2}\right]u.$$

The form of this equation matches exactly with that considered by Arfken and Weber (p. 834). The power-series method yields a solution

$$u(r) = s^{\ell+1}e^{-s}\mathcal{L}^{2l+1}_{n-l-1}(2s),$$

where

$$\mathcal{L}^{2l+1}_{n-l-1}(2s)$$

is an **associated Laguerre polynomial**. In order for these polynomials to exist – and hence, in order for a normalisable solution to exist – we have the following conditions on ak:

$$n = \frac{1}{ak},$$

where n must be a **positive integer** (normalisability), AND

$$\ell < n.$$

But

$$ak = \frac{\hbar^2}{mk_0e^2}\frac{\sqrt{2m|E|}}{\hbar}.$$

Hence,

$$\begin{aligned} -\frac{2mE}{\hbar^2}\left(\frac{\hbar^2}{mk_0e^2}\right)^2 &= a^2k^2, \\ &= \frac{1}{n^2}, \qquad n = 1, 2, \cdots, \end{aligned}$$

and

$$\begin{aligned} E &= -\frac{1}{n^2}\left(\frac{mk_0e^2}{\hbar^2}\right)^2\frac{\hbar^2}{2m}, \\ &= -\frac{mk_0^2e^4}{2\hbar^2}\frac{1}{n^2}, \\ &= E_n. \end{aligned}$$

In other words,

$$E_n = -\frac{1}{n^2}\frac{mk_0^2e^4}{2\hbar^2}, \qquad n = 1, 2, \cdots,$$

and in SI units,

$$E_n = -\frac{1}{n^2}\frac{1}{(4\pi\epsilon_0)^2}\frac{me^4}{2\hbar^2}, \qquad n = 1, 2, \cdots,$$

which agrees exactly with the Bohr picture.

Note: the associated Laguerre polynomials can be generated according to the relation

$$\mathcal{L}_n^p(x) = \frac{e^x x^{-p}}{n!}\frac{d^n}{dx^n}\left(e^{-x}x^{n+p}\right),$$

and are orthogonal with respect to a weighting function:

$$\int_0^\infty e^{-x}x^p\mathcal{L}_n^p(x)\mathcal{L}_m^p(x)\,\mathrm{d}x = \frac{(n+p)!}{n!}\delta_{nm} \qquad (*)$$

(The same p-value in each polynomial to be integrated!).

There remain some issues to tidy up. Consider again the argument of the radial function:

$$s = kr, \qquad k = \sqrt{2m|E|/\hbar^2}.$$

Now

$$\begin{aligned} |E| &= \frac{1}{n^2}\frac{mk_0^2e^4}{2\hbar^2}, \\ \frac{2m|E|}{\hbar^2} &= \frac{1}{n^2}\frac{m^2k_0^2e^4}{2\hbar^4}, \\ \sqrt{\frac{2m|E|}{\hbar^2}} &= = \frac{1}{n}\frac{mk_0e^2}{\hbar^2}, \\ &= \frac{1}{n}\frac{1}{a_0}. \end{aligned}$$

Hence,

$$\begin{aligned} R(r) &= \frac{u(r)}{r}, \\ &= \frac{1}{r}\left(\frac{r}{na_0}\right)^{\ell+1}\mathrm{e}^{-r/na_0}\mathcal{L}_{n-l-1}^{2\ell+1}\left(\frac{2r}{na_0}\right), \\ &= \frac{1}{na_0}\left(\frac{r}{na_0}\right)^{\ell}\mathrm{e}^{-r/na_0}\mathcal{L}_{n-l-1}^{2\ell+1}\left(\frac{2r}{na_0}\right). \end{aligned}$$

However, this result is only correct up to normalisation. Thus, we introduce

$$R_{n\ell} = N_{n\ell}\left(\frac{2r}{na_0}\right)^{\ell}\mathrm{e}^{-r/na_0}\mathcal{L}_{n-l-1}^{2\ell+1}\left(\frac{2r}{na_0}\right).$$

In view of the orthogonality relation (*) (the orthogonality condition on $(\mathcal{L}_n^p, \mathcal{L}_m^p)$),

$$N_{n\ell} = \sqrt{\frac{(n-l-1)!}{2n(n+l)!}\left(\frac{2}{na_0}\right)^3}.$$

The final solution for the hydrogen atom is

$$\begin{aligned}\psi_{n\ell m}(r,\theta,\varphi) &= \sqrt{\frac{(n-l-1)!}{2n(n+l)!}\left(\frac{2}{na_0}\right)^3}\left(\frac{2r}{na_0}\right)^{\ell} \mathrm{e}^{-r/na_0}\mathcal{L}^{2\ell+1}_{n-l-1}\left(\frac{2r}{na_0}\right)Y^m_\ell(\theta,\varphi),\\ E_n &= -\frac{1}{n^2}\frac{mk_0^2e^4}{2\hbar^2}.\end{aligned}$$

14.3 Notes on the solution

- The energy levels are quantised according to the positive integer $n = 1, 2, \cdots$. The label n is called the **principal quantum number**.
- The quantity $L^2 = \hbar\ell(\ell+1)$ has the interpretation of **angular momentum**. For each n-value, there are $\ell = 0, 1, 2, \cdots, n-1$ allowed values of an angular momentum.
- The quantity m has the interpretation of the **angular momentum in the z-direction**. For each ℓ-value, there are $m = -\ell, \cdots, \ell$ allowed values of the projection.
- Hence, for $n = 1$ there is precisely one possible state with quantum numbers $(n, \ell, m) = (1, 0, 0)$.
- For $n = 2$ there are four possible states, with quantum numbers

$$(n,\ell,m) = (1,0,0),\ (1,1,-1),\ (1,1,0),\ (1,1,-1).$$

- For $n = 3$ there are nine possible states, with quantum numbers

$$(n,\ell,m) = (3,0,0),\ (3,1,-1),\ (3,1,0),\ (3,1,-1),\\ (3,2,-2),\ (3,3,-1),\ (3,2,0),\ (3,2,1)\,(3,2,2).$$

- The **degeneracy** of an energy level E_n refers to the fact that the level accommodates several distinct eigenstates (different values of angular momentum). Each ℓ-value corresponds to $2\ell + 1$ distinct states, and there are $\ell = 0, 1, \cdots, n-1$ possible ℓ-values for a given energy level. Thus, the degree of degeneracy is

$$\text{degeneracy of } E_n = \sum_{\ell=0}^{n-1}(2\ell+1) = n^2.$$

- In spectroscopy, the states are classified according to the angular-momentum eigenvalue ℓ:
  - $\ell = 0$ – s-states;
  - $\ell = 1$ – p-states;
  - $\ell = 2$ – d-states;

  and thereafter alphabetically[2]

- First few radial wavefunctions (normalised):

$$\begin{aligned} R_{10}(r) &= \frac{2}{a_0^{3/2}} e^{-r/a_0}, \\ R_{20}(r) &= \frac{2}{(2a_0)^{3/2}} \left(1 - \frac{r}{2a_0}\right) e^{-r/2a_0}, \\ R_{21}(r) &= \frac{2}{\sqrt{3}(2a_0)^{3/2}} \frac{r}{a_0} e^{-r/2a_0}. \end{aligned}$$

  See Fig. 14.2.

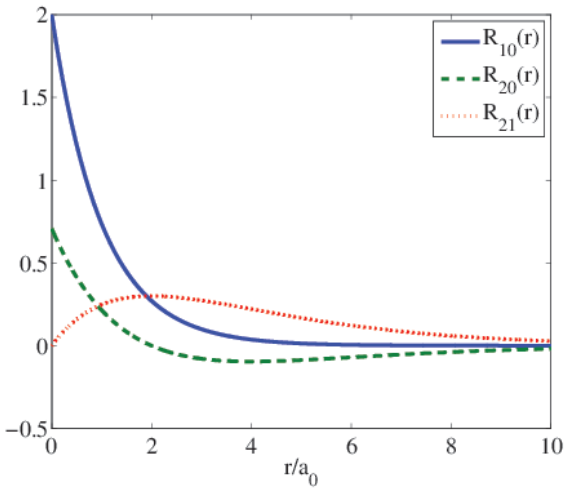


Figure 14.2: First few radial wavefunctions, hydrogen atom.

14.4 Spherical harmonics – visualisations

In Figs. 14.3–14.6 we have plotted the real part of the spherical harmonics corresponding to $\ell = 0, 1, 3, 3$. The Matlab code to do this is given in Appendix A.

[2]Mnemonic: "Silly Physicists Deny Feeling Gravity".

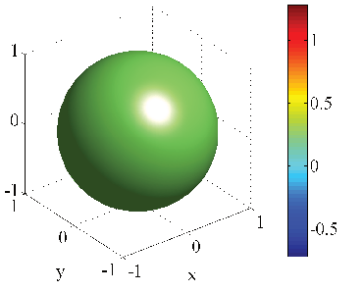
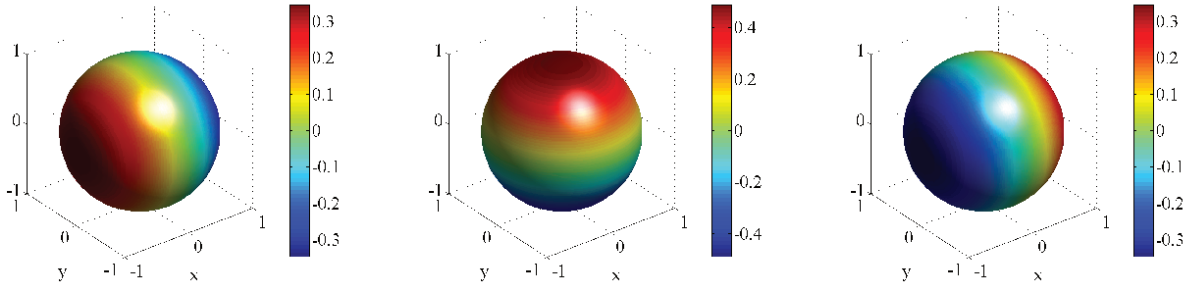
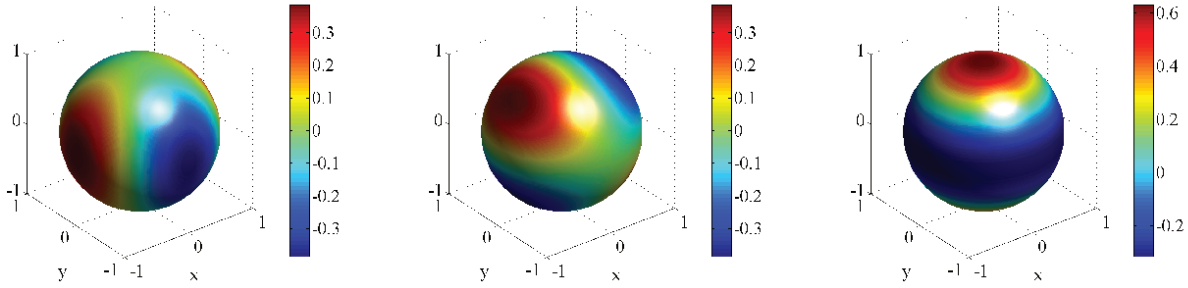


Figure 14.3: The $\ell = 0$ spherical harmonic.

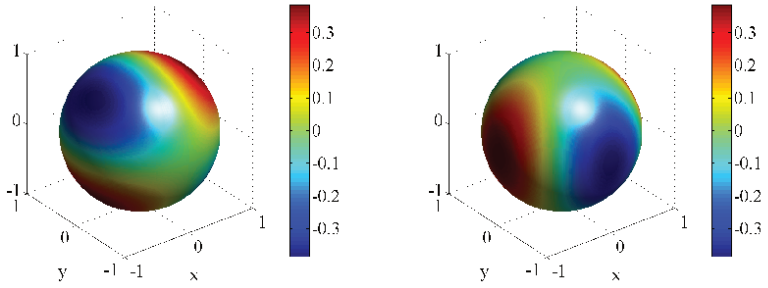


(a) $m = 1$ (b) $m = 0$ (c) $m = -1$

Figure 14.4: Real part of the $\ell = 1$ spherical harmonics.

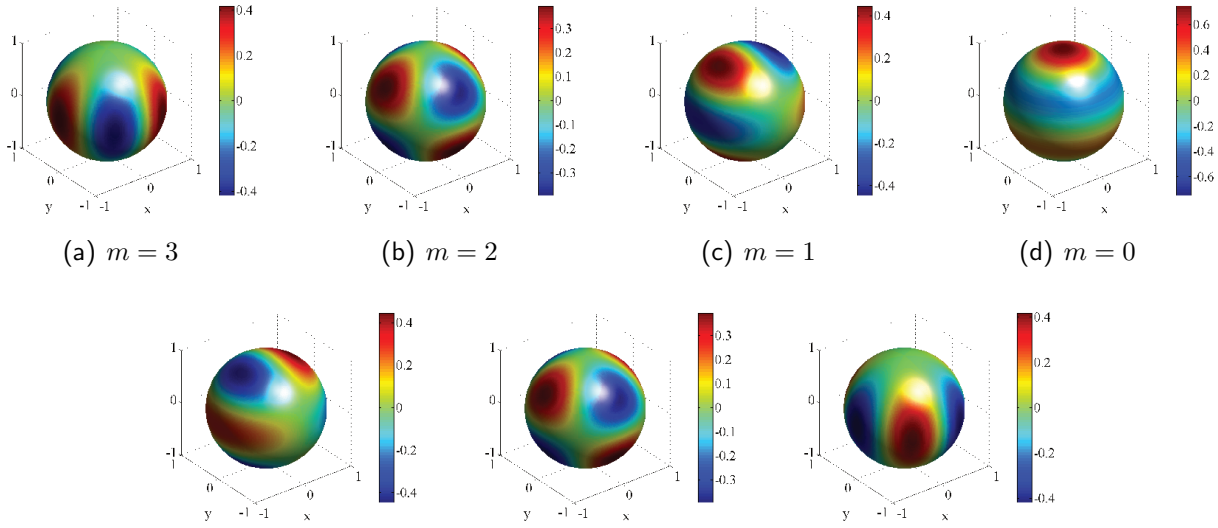


(a) $m = 2$ (b) $m = 1$ (c) $m = 0$



(d) $m = -1$ (e) $m = -2$

Figure 14.5: Real part of the $\ell = 2$ spherical harmonics.



(a) $m = 3$ (b) $m = 2$ (c) $m = 1$ (d) $m = 0$

(e) $m = -1$ (f) $m = -2$ (g) $m = -3$

Figure 14.6: Real part of the $\ell = 3$ spherical harmonics.

Chapter 15

General treatment of central potentials

Reading material for this chapter: None recommended

15.1 Introduction

We study the Schrödinger equation for a particle that experiences a force from the general central potential

$$\mathcal{U}(\boldsymbol{r}) = \mathcal{U}(r).$$

in three dimensions. We are going to recycle much from our experience with the hydrogen atom. The Hamiltonian reads

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + \mathcal{U}(r)$$

This is manifestly time independent, and we therefore separate out the space and time dependence, and solve the eigenvalue problem

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + \mathcal{U}(r)\right)\psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}).$$

This is solved using spherical harmonics.

15.2 The solution

In our treatment of the hydrogen atom, the nature of the potential did not enter into the angular solution. Thus, we propose a solution

$$\psi(\boldsymbol{r}) = R(r)Y_{\ell m}(\theta, \varphi),$$

for the eigenvalue problem. Acting on this solution with the Laplacian gives

$$\nabla^2\psi = \frac{Y_{\ell m}}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) - \frac{\ell(\ell+1)}{r^2}R(r)Y_{\ell m}.$$

Hence, the Schrödinger equation reads

$$-\frac{\hbar^2}{2m}\left[\frac{Y_{\ell m}}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) - \frac{\ell(\ell+1)}{r^2}R(r)Y_{\ell m}\right] + \mathcal{U}(r)R(r)Y_{\ell m} = ER(r)Y_{\ell m}. \qquad (*)$$

Radial part

We can freely divide out by $Y_{\ell m}$ in Eq. (*) to give

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) - \frac{\ell(\ell+1)}{r^2}R(r)\right] + \mathcal{U}(r)R(r) = ER(r).$$

Now we perform the **magic substitution** $R(r) = u(r)/r$:

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) = \frac{1}{r}\frac{\partial^2 u}{\partial r^2}.$$

and the Schrödinger equation reduces to

$$-\frac{\hbar^2}{2m}\left[\frac{u''}{r} - \frac{\ell(\ell+1)}{r^2}\frac{u}{r}\right] + \mathcal{U}(r)\frac{u}{r} = E\frac{u}{r}.$$

Multiplying up by r gives

$$-\frac{\hbar^2}{2m}\left[u'' - \frac{\ell(\ell+1)}{r^2}u\right] + \mathcal{U}(r)u = Eu.$$

Now here is the wonderful trick: We introduce an **effective potential**

$$\mathcal{U}_{\text{eff}}(r) = \mathcal{U}(r) + \frac{\hbar^2}{2m}\frac{\ell(\ell+1)}{r^2}.$$

Then, the Schrödinger equation reduces to a quasi-one-dimensional form:

$$-\frac{\hbar^2}{2m}u'' + \mathcal{U}_{\text{eff}}(r)u = Eu.$$

In conclusion, we have reduced the problem of obtaining the full wavefunction $\phi(r,\theta,\varphi)$, to a comparatively simple, quasi-one-dimensional problem, using the following sequence of steps:

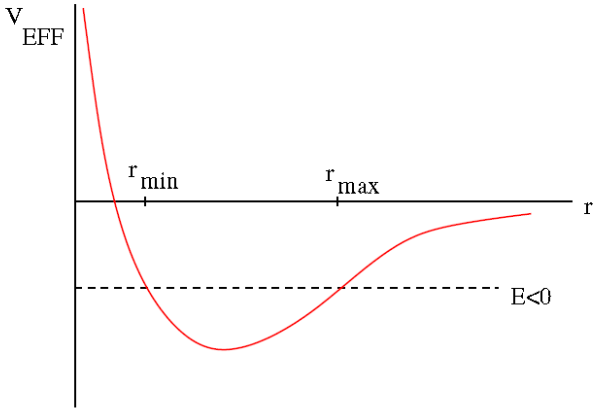


Figure 15.1: Effective potential for hydrogen, $\ell > 0$.

- $\psi(r, \theta, \varphi) = R(r)Y_{\ell m}(\theta, \varphi)$;
- $R(r) = u(r)/r$;
- Effective potential: $\mathcal{U}_{\text{eff}}(r) = \mathcal{U}(r) + (\hbar^2/2m)\ell(\ell+1)/r^2$;
- Quasi-one-dimensional model: $(-\hbar^2/2m)u'' + \mathcal{U}_{\text{eff}}(r)u = Eu$;
- Solve the latter for the eigenvalues of energy.

Note finally the effective potential for hydrogen ($\ell \neq 0$) (Fig. 15.1). For $\ell > 0$, the potential is positive as $r \to 0$ – there is an effective repulsive force as the nucleus is the electron approaches the nucleus. This is sometimes referred to as the **centrifugal barrier**.

Chapter 16

Angular momentum

Reading material for this chapter: Mandl, Chapter 2

16.1 Overview

In previous chapters, we isolated the angular part of the Laplacian, Δ:

$$\frac{1}{r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right].$$

Call the part inside the square brackets Δ_Ω:

$$\Delta_\Omega=\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}.$$

In previous chapters, we solved

$$\Delta_\Omega Y(\theta,\varphi)=-L^2Y(\theta,\varphi),$$

and found that the answer was spherical harmonics:

$$\begin{aligned} Y(\theta,\varphi) &= Y_{\ell m}(\theta,\varphi),\\ L^2 &= \ell(\ell+1),\\ \ell &= 0,1,2,\cdots,\\ m &= -\ell,-\ell+1,\cdots,\ell-1,\ell. \end{aligned}$$

We interpreted these solutions as **eigenfunctions of angular momentum**, with eigenvalues ℓ. The secondary eigenvalues m related to the component of angular momentum measured along the z-axis. We are going to study this interpretation further, and construct an **abstract theory** of angular momentum, independent of the position representation. However, in this chapter, we continue to

work in the position representation to gain and justify our identification of the angular part of the Laplacian with angular momentum.

16.2 The definition

Let $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$ be the usual position and momentum operators with canonical commutation relation

$$[\hat{\mathrm{r}}_i, \hat{\mathrm{p}}_j] = -\mathrm{i}\hbar\delta_{ij}.$$

The angular-momentum operator is **defined** as

$$\widehat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}.$$

Going over to the position representation, this is

$$\widehat{\mathbf{L}} = -\mathrm{i}\hbar \boldsymbol{r} \times \nabla.$$

We use spherical polar coordinates:

$$\begin{aligned} \boldsymbol{r} &= r\hat{\boldsymbol{r}}, \\ \nabla &= \frac{\hat{\boldsymbol{r}}}{h_r}\partial_r + \frac{\hat{\boldsymbol{\theta}}}{h_\theta}\partial_\theta + \frac{\hat{\boldsymbol{\varphi}}}{h_\varphi}\partial_\varphi, \\ \hat{\boldsymbol{r}} \times \hat{\boldsymbol{\theta}} &= \hat{\boldsymbol{\varphi}}, \end{aligned}$$

where $(\hat{\boldsymbol{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\varphi}})$ are an orthonormal triad and

$$(h_r, h_\theta, h_\varphi) = (1, r, r\sin\theta)$$

are the scale factors of the spherical polar coordinate system. Hence,

$$\widehat{\mathbf{L}} = -\mathrm{i}\hbar r \left[\frac{\hat{\boldsymbol{\varphi}}}{h_\theta}\partial_\theta - \frac{\hat{\boldsymbol{\theta}}}{h_\varphi}\partial_\varphi \right].$$

Recall,

$$\begin{aligned} \boldsymbol{x} &= (r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta), \\ \hat{\boldsymbol{\theta}} &= \frac{1}{h_\theta}\frac{\partial \boldsymbol{x}}{\partial\theta}, \\ \hat{\boldsymbol{\varphi}} &= \frac{1}{h_\varphi}\frac{\partial \boldsymbol{x}}{\partial\varphi}. \end{aligned}$$

It follows that

$$\begin{aligned}\hat{\boldsymbol{z}}\cdot\hat{\boldsymbol{\theta}} &= -\sin\theta,\\ \hat{\boldsymbol{z}}\cdot\hat{\boldsymbol{\varphi}} &= 0,\end{aligned}$$

and hence, **the projection of angular momentum on to the z-axis is**

$$\begin{aligned}\widehat{\mathrm{L}}_z &= \hat{\boldsymbol{z}}\cdot\widehat{\mathbf{L}},\\ &= -\mathrm{i}\hbar r\left(-\frac{\hat{\boldsymbol{z}}\cdot\hat{\boldsymbol{\theta}}}{h_\varphi}\partial_\varphi\right),\\ &= -\mathrm{i}\hbar\partial_\varphi.\end{aligned}$$

Now let's compute the action of $\widehat{\mathbf{L}}^2$ on a function $Y(\theta,\varphi)$. We need to be careful here because $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\varphi}}$ depend on space:

$$\begin{aligned}\frac{\widehat{\mathbf{L}}^2 Y}{-\hbar^2 r^2} &= \left(\frac{\hat{\boldsymbol{\varphi}}}{h_\theta}\partial_\theta - \frac{\hat{\boldsymbol{\theta}}}{h_\varphi}\partial_\varphi\right)\cdot\left(\frac{\hat{\boldsymbol{\varphi}}}{h_\theta}\partial_\theta Y - \frac{\hat{\boldsymbol{\theta}}}{h_\varphi}\partial_\varphi Y\right),\\ &= \frac{\hat{\boldsymbol{\varphi}}}{h_\theta}\cdot\partial_\theta\left(\frac{\hat{\boldsymbol{\varphi}}}{h_\theta}\partial_\theta Y\right) - \frac{\hat{\boldsymbol{\varphi}}}{h_\theta}\cdot\partial_\theta\left(\frac{\hat{\boldsymbol{\theta}}}{h_\varphi}\partial_\varphi Y\right) - \frac{\hat{\boldsymbol{\theta}}}{h_\varphi}\cdot\partial_\varphi\left(\frac{\hat{\boldsymbol{\varphi}}}{h_\theta}\partial_\theta Y\right) + \frac{\hat{\boldsymbol{\theta}}}{h_\varphi}\cdot\partial_\varphi\left(\frac{\hat{\boldsymbol{\theta}}}{h_\varphi}\partial_\varphi Y\right),\\ &= \frac{1}{h_\theta}\partial_\theta\left(\frac{\partial_\theta Y}{h_\theta}\right) + \frac{1}{h_\theta^2}\partial_\theta Y\left(\hat{\boldsymbol{\varphi}}\cdot\partial_\theta\hat{\boldsymbol{\varphi}}\right) - \frac{1}{h_\theta h_\varphi}\partial_\varphi Y\left(\hat{\boldsymbol{\varphi}}\cdot\partial_\theta\hat{\boldsymbol{\theta}}\right) - \frac{1}{h_\theta h_\varphi}\partial_\theta Y\left(\hat{\boldsymbol{\theta}}\cdot\partial_\theta\hat{\boldsymbol{\varphi}}\right),\\ &\quad \frac{1}{h_\varphi}\partial_\varphi\left(\frac{\partial_\varphi Y}{h_\varphi}\right) + \frac{1}{h_\varphi^2}\partial_\varphi Y\left(\hat{\boldsymbol{\theta}}\cdot\partial_\varphi\hat{\boldsymbol{\theta}}\right).\end{aligned}$$

A lot of the cross terms can be made to go away. For example,

$$|\hat{\boldsymbol{\theta}}|^2 = 1 \implies \hat{\boldsymbol{\theta}}\cdot\partial_\varphi\hat{\boldsymbol{\theta}} = 0, \qquad |\hat{\boldsymbol{\varphi}}|^2 = 1 \implies \hat{\boldsymbol{\varphi}}\cdot\partial_\theta\hat{\boldsymbol{\varphi}} = 0.$$

The other two cross terms require direct computation:

$$\begin{aligned}\hat{\boldsymbol{\varphi}} &= (-\sin\varphi, \cos\varphi, 0),\\ \hat{\boldsymbol{\theta}} &= (\cos\theta\cos\varphi, cos\theta\sin\varphi, -\sin\theta),\\ \partial_\varphi\hat{\boldsymbol{\varphi}} &= (-\cos\varphi, -\sin\varphi, 0),\\ \partial_\theta\hat{\boldsymbol{\theta}} &= (-\sin\theta\cos\varphi, -\sin\theta\sin\varphi, -\cos\theta),\end{aligned}$$

hence

$$\hat{\boldsymbol{\theta}}\cdot\partial_\varphi\hat{\boldsymbol{\varphi}} = -\cos\theta, \qquad \hat{\boldsymbol{\varphi}}\cdot\partial_\theta\hat{\boldsymbol{\theta}} = 0.$$

Hence,

$$\frac{\widehat{\mathbf{L}}^2 Y}{-\hbar^2 r^2} = \frac{1}{h_\theta}\partial_\theta\left(\frac{\partial_\theta Y}{h_\theta}\right) + \frac{1}{h_\varphi}\partial_\varphi\left(\frac{\partial_\varphi Y}{h_\varphi}\right) - \frac{1}{h_\theta h_\varphi}\left(\hat{\boldsymbol{\varphi}}\cdot\partial_\varphi\hat{\boldsymbol{\theta}}\right)\partial_\theta Y.$$

We substitute:

$$h_\theta = r, \qquad h_\varphi = r\sin\theta, \qquad \hat{\boldsymbol{\theta}}\cdot\partial_\varphi\hat{\boldsymbol{\varphi}} = -\cos\theta,$$

hence

$$\begin{aligned}\frac{\widehat{\mathbf{L}}^2 Y}{-\hbar^2 r^2} &= \frac{1}{r^2}\partial_{\theta\theta}Y + \frac{1}{r^2\sin^2\theta}\partial_{\varphi\varphi}Y + \frac{1}{r^2}\frac{\cos\theta}{\sin\theta}\partial_\theta Y,\\ &= \frac{1}{r^2}\left[\frac{1}{\sin\theta}\partial_\theta\left(\sin\theta\partial_\theta Y\right) + \frac{1}{\sin^2\theta}\partial_{\varphi\varphi}Y\right],\\ &= \frac{1}{r^2}\Delta_\Omega Y.\end{aligned}$$

Thus, starting out with the standard definition of angular momentum, we have shown,

$$\begin{aligned}\widehat{\mathbf{L}} &:= -\mathrm{i}\hbar\boldsymbol{r}\times\nabla,\\ \widehat{\mathbf{L}}^2 &= -\hbar^2\Delta_\Omega, \qquad (=\widehat{\mathrm{L}}_x^2+\widehat{\mathrm{L}}_y^2+\widehat{\mathrm{L}}_z^2)\\ \hat{\boldsymbol{z}}\cdot\widehat{\mathbf{L}} &= -\mathrm{i}\hbar\partial_\varphi, \qquad (=\widehat{\mathrm{L}}_z).\end{aligned}$$

We now return to Cartesian coordinates and derive **commutation relations** between the Cartesian components of the angular-momentum operator $\widehat{\mathbf{L}}$.

16.3 Commutation relations

Define

$$\widehat{\mathbf{L}} = -\mathrm{i}\hbar\boldsymbol{r}\times\nabla.$$

In other words,

$$\begin{aligned}\widehat{\mathrm{L}}_x &= -\mathrm{i}\hbar\left(y\partial_z - z\partial_y\right),\\ \widehat{\mathrm{L}}_y &= -\mathrm{i}\hbar\left(z\partial_x - x\partial_z\right),\\ \widehat{\mathrm{L}}_z &= -\mathrm{i}\hbar\left(x\partial_y - y\partial_x\right).\end{aligned}$$

We compute

$$\begin{aligned}\frac{\widehat{\mathrm{L}}_x\widehat{\mathrm{L}}_y Y}{(-\mathrm{i}\hbar)^2} &= (y\partial_z - z\partial_y)(z\partial_x Y - x\partial_z Y),\\ &= y\partial_z(zY_x) - xyY_{zz} - z^2Y_{yx} + xzY_{yz},\\ &= yY_x + yzY_{xz} - xyY_{zz} - z^2Y_{yx} + xzY_{yz}.\end{aligned}$$

Similarly,

$$\frac{\widehat{\mathrm{L}}_y\widehat{\mathrm{L}}_x Y}{(-\mathrm{i}\hbar)^2} = xY_y + xzY_{zy} - xyY_{zz} - z^2Y_{xy} + xyY_{xz}.$$

Subtracting gives

$$\frac{\widehat{\mathrm{L}}_x\widehat{\mathrm{L}}_y Y - \widehat{\mathrm{L}}_y\widehat{\mathrm{L}}_x Y}{(-\mathrm{i}\hbar)^2} = yY_x - xY_y = -\left(x\partial_y - y\partial_x\right)Y.$$

In other words,

$$\begin{aligned}\frac{\left[\widehat{\mathrm{L}}_x, \widehat{\mathrm{L}}_y\right]}{(-\mathrm{i}\hbar)^2}Y &= -\left(x\partial_y - y\partial_x\right)Y,\\ \left[\widehat{\mathrm{L}}_x, \widehat{\mathrm{L}}_y\right] &= (\mathrm{i}\hbar)(-\mathrm{i}\hbar)\left(x\partial_y - y\partial_x\right)Y,\\ &= \mathrm{i}\hbar\widehat{\mathrm{L}}_z Y.\end{aligned}$$

Performing cyclic permutations on the coordinates gives the following general commutation relation:

$$\left[\widehat{\mathrm{L}}_i, \widehat{\mathrm{L}}_j\right] = \mathrm{i}\hbar\sum_{k=1}^{3}\epsilon_{ijk}\widehat{\mathrm{L}}_k. \qquad (*)$$

Going back to the spherical-polar coordinate representation for an instant, it is readily shown that

$$\left[\Delta_\Omega, \partial_\varphi\right]Y = 0,$$

and thus,

$$\left[\widehat{\mathbf{L}}^2, \widehat{\mathrm{L}}_z\right] = 0.$$

However, there is nothing special about the z-direction – we could just as easily have set up a coordinate system where the polar angle is measured from the x- or y-axis. Thus,

$$\left[\widehat{\mathbf{L}}^2, \widehat{\mathrm{L}}_x\right] = \left[\widehat{\mathbf{L}}^2, \widehat{\mathrm{L}}_y\right] = \left[\widehat{\mathbf{L}}^2, \widehat{\mathrm{L}}_z\right] = 0.$$

Thus, $\{\widehat{\mathbf{L}}^2, \widehat{\mathrm{L}}_z\}$ are **compatible** operators, and there is a simultaneous basis of eigenvectors for them – the spherical harmonics. However, in view of the commutation relation (*), it is not possible to obtain a simultaneous eigenbasis for $\{\widehat{\mathbf{L}}^2, \widehat{\mathrm{L}}_x, \widehat{\mathrm{L}}_z\}$, say. Hence, $\{\widehat{\mathbf{L}}^2, \widehat{\mathrm{L}}_z\}$ is a maximally commuting

set of operators, or a **complete set of commuting observables** (CSCO).

Chapter 17

Angular momentum: abstract setting

Reading material for this chapter: Mandl, Chapter 5

17.1 Abstract setting

In this chapter we define angular momentum in an abstract setting. The spherical harmonics just discussed are then just one **representation** of this abstract formalism. First, a definition:

Definition 17.1 *An Lie algebra* $\mathfrak{g}$ *is a complex vector space endowed with a a binary operation called the Lie bracket:*

$$\begin{aligned} \mathfrak{g} \times \mathfrak{g} &\rightarrow \mathfrak{g}, \\ (a, b) &\rightarrow [a, b], \end{aligned}$$

such that the following axioms hold:

- *Bilinearity:*

$$[\lambda a + \mu b, c] = \lambda [a, b] + \mu [b, c], \qquad [c, \lambda a + \mu b] = \lambda [c, a] + \mu [c, b],$$

- $[a, a] = 0$
- *The Jacobi identity:*

$$[a, [b, c]] + [c, [a, b]] + [b, [c, a]] = 0$$

for all $a, b, b \in \mathfrak{g}$ *and* $\lambda, \mu \in \mathbb{C}$. Note that the properties (1)–(3) imply that the Lie bracket is antisymmetric,

$$[a, b] = -[b, a],$$

for all a and b in $\mathfrak{g}$.

Example

Let $\mathcal{H}$ be a complex vector space. Denote by $\mathcal{L}(\mathcal{H})$ the set of all linear operators from $\mathcal{H}$ to itself. The set $\mathcal{L}(\mathcal{H})$ is itself a complex vector space. Introduce a binary operation on $\mathcal{L}(\mathcal{H})$ using operator composition. This enables us to define the **commutator** on $\mathcal{L}(\mathcal{H})$:

$$\begin{aligned}[\cdot,\cdot] : \mathcal{L}(\mathcal{H}) \times \mathcal{L}(\mathcal{H}) &\to \mathcal{L}(\mathcal{H}), \\ (S,T) &\mapsto [S,T] := ST - TS. \end{aligned} \tag{17.1}$$

We have the following theorem:

Theorem 17.1 *The operator commutator defined in Equation* (17.1) *is bilinear, satisfies* $[S,S] = 0$ *for all* $S \in \mathcal{L}(\mathcal{H})$ *and satisfies the Jacobi identity.*

The proof of the first two properties is straightforward. The proof of the third property is by direct computation:

$$\begin{aligned}[A,[B,C]] &= A[B,C] - [B,C]A, \\ &= A(BC - CB) - (BC - CB)A, \\ &= ABC - ACB - BCA + CBA. \end{aligned}$$

where the brackets $(\cdot)$ are not important here because operator composition is associative. Similarly,

$$[B,[C,A]] = BCA - BAC - CAB + ACB,$$

and

$$[C,[A,B]] = CAB - CBA - ABC + BAC.$$

Add up:

$$[A,[B,C]] + [B,[C,A]] + [C,[A,B]] = 0.$$

Thus, $\mathcal{L}(\mathcal{H})$ is a Lie group with the operator commutator as the Lie bracket.

As a realisation of this concept, consider the set $\mathbb{C}^{n\times n}$. This can be identified as the set $\mathcal{L}(\mathbb{C}^n)$, i.e. the set of all linear operators on complex-valued column vectors. This is a Lie algebra, with Lie bracket given by the matrix commutator,

$$[A,B] = AB - BA.$$

Again, it is immediately obvious that the matrix commutator is bilinear, satisfies $[A, A] = 0$, and satisfies the Jacobi identity. Thus, $\mathbb{C}^{n\times n}$ is a Lie algebra.

17.2 The Lie Algebra of Angular Momentum

Let $\mathcal{H}$ denote the state space of angular momentum. The angular momentum operator is

$$\boldsymbol{J} = J_x\hat{\boldsymbol{x}} + J_y\hat{\boldsymbol{y}} + J_z\hat{\boldsymbol{z}},$$

where $\hat{\boldsymbol{x}}$ is the unit vector in the x direction etc. and J_x is the component of angular momentum in the x-direction etc. The set of all finite operator compositions of J_x, J_y and J_z together with all linear combinations of J_x, J_y and J_z and all linear combinations of finite compositions is a complex vector space, denoted in the present context by $\mathcal{L}(\mathcal{H})$. The set $\mathcal{L}(\mathcal{H})$ by construction is closed under addition of operators, scalar multiplication, and composition of operators. Using the operator composition as a bracket, $\mathcal{L}(\mathcal{H})$ is made into a Lie algebra.

We now consider the vector subspace of $\mathcal{L}(\mathcal{H})$ formed as follows:

$$\mathcal{S}\left(J_x, J_y, J_z\right)$$

and we **impose** the canonical commutation relation

$$[J_i, J_j] = \mathrm{i}\hbar\sum_{k=1}^{3}\epsilon_{ijk}J_k. \tag{17.2}$$

The vector subspace $\mathcal{S}\left(J_x, J_y, J_z\right)$ is closed under linear combinations and scalar multiplication. It is also closed under the canonical commutation relation (17.2). We have:

- $\mathcal{S}\left(J_x, J_y, J_z\right)$ is a complex vector space.
- The bracket in Equation (17.2) takes elements in $\mathcal{S}\left(J_x, J_y, J_z\right)$ and sends them to other elements in $\mathcal{S}\left(J_x, J_y, J_z\right)$.
- The bracket in Equation (17.2) is bilinear, satisfies $[a, a] = 0$ for all $a \in \mathcal{S}\left(J_x, J_y, J_z\right)$ and satisfies the Jacobi identity – these properties are **inherited** from the bracket operating on the full set $\mathcal{L}(\mathcal{H})$.

Thus, $\mathcal{S}\left(J_x, J_y, J_z\right)$ is a Lie algebra in its own right – it is the **Lie algebra of angular momentum**.

In practice, not only are J_x, J_y and J_z of interest, but also $J^2 = J_x^2 + J_y^2 + J_z^2$. Therefore, it is of interest not only to study $\mathcal{S}\left(J_x, J_y, J_z\right)$ but also $\mathcal{L}(\mathcal{H})$, got by forming the closure of $\mathcal{S}\left(J_x, J_y, J_z\right)$

under combinations of operator composition, addition and scalar multiplication. This is called the **enveloping algebra**.

We now prove the following results about the Lie algebra of angular momentum:

Theorem 17.2 *The following results hold for the algebra just defined:*

1. *The operator J^2 commutes with all elements of the algebra.*
2. *The pair $\{J_z, J^2\}$ is a maximally-commuting set, with eigenvalues $(\hbar m, \hbar^2\lambda)$, such that*

$$\lambda \geq 0, \qquad m^2 \leq \lambda.$$

3. *The J_z-eigenspaces are non-degenerate.*
4. *For a given λ-value, there is a maximum and a minimum m-value.*
5. *There is only a finite number eigenvalues of J_z and J^2, in the relation*

$$\begin{aligned} \textit{Eigenvalues of } J^2 &= \hbar^2 j(j+1), \qquad j = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, \cdots, \\ \textit{Eigenvalues of } J_z &= \hbar m, \qquad m = -j, \cdots, j. \end{aligned}$$

Proof:

1. We use the following commutation relation for products:

$$[A, BC] = [A, B]C + B[A, C].$$

   Hence,

$$\begin{aligned} [J_z, J^2] &= [J_z, J_x^2] + [J_z, J_y^2], \\ &= [J_z, J_x]J_x + J_x[J_z, J_x] + [J_z, J_y]J_y + J_y[J_z, J_y], \\ &= \mathrm{i}\hbar J_y J_x + \mathrm{i}\hbar J_x J_y - \mathrm{i}\hbar J_x J_y - \mathrm{i}\hbar J_y J_x, \\ &= 0. \end{aligned}$$

2. The set $\{J^2, J_z\}$ is a commuting set. Forming another set such as $\{J^2, J_z, J_x\}$ leads to a non-commuting triple, since $[J_x, J_z] \neq 0$. Since $\{J_x, J_y, J_z\}$ are assumed to be linearly independent, the maximum possible set of commuting observables is given by $\{J^2, J_z\}$. It follows that there is a basis of simultaneous eigenvectors for this set, or a **complete set of commuting observables** (CSCO).

3. To show that the J_z-eigenspaces are non-degenerate, two approaches can be taken. The first is to note that if the eigenspaces were degenerate, then there would be a third quantum number characterizing the angular momentum. But this is not possible, since J^2 and J_z are a complete set of commuting observables associated with precisely two quantum numbers – the eigenvalues of J^2 and J_z respectively. A second but related approach is to assume that a third quantum number exists (say μ), labelling the supposed degeneracy of the J_z eigenstates, such that the eigenstates are $|\lambda, m, \mu\rangle$. It is possible to go over to the position representation of angular momentum, which is valid for $\lambda = j(j+1)$, with $j \in \{0, 1, 2, \cdots\}$ and $m = -j, \cdots, j$ whereupon we have

$$\langle \theta, \varphi | \lambda, m, \mu \rangle = Y_{j,m}(\theta, \varphi).$$

This is an identity, yet the supposed quantum number μ does not appear on the right-hand side. Thus, we are forced to conclude that the quantum number μ does not exist, and hence that the J_z-eigenspaces are non-degenerate.

4. We show that $|m| \leq \lambda$. Let J^2 have eigenvalues $\hbar^2\lambda$ and let J_z have eigenvalues $\hbar m$.

Let $|\phi\rangle$ be an arbitrary norm-one state and let $|\chi\rangle = J_x|\phi\rangle$. We have,

$$\begin{aligned}
\langle \phi | J_x^2 | \phi \rangle &= \langle \phi | J_x J_x | \phi \rangle, \\
&= \langle \phi | J_x | \chi \rangle, \\
&= \langle \phi | J_x^\dagger | \chi \rangle, \\
&= \langle \chi | \chi \rangle, \\
&\geq 0.
\end{aligned}$$

Similarly for the other components; it follows that $\lambda \geq 0$.

As yet we do not know what the precise values of (λ, m) are. However, we may write down the simultaneous eigenbasis $\{|\lambda, m\rangle\}_{\lambda,m}$, which spans the vector space of states, such that

$$\begin{aligned}
J^2|\lambda, m\rangle &= \hbar^2\lambda|\lambda, m\rangle, \\
J_z|\lambda, m\rangle &= \hbar m|\lambda, m\rangle.
\end{aligned}$$

Hence

$$\begin{aligned}
\langle \lambda, m | J^2 | \lambda, m \rangle &= \hbar^2\lambda, \\
\langle \lambda, m | J_z^2 | \lambda, m \rangle &= \hbar^2 m^2.
\end{aligned}$$

In addition,

$$\langle \lambda, m | J_x^2 + J_y^2 | \lambda, m \rangle = \langle \lambda, m | J^2 - J_z^2 | \lambda, m \rangle \geq 0,$$

hence

$$\lambda - m^2 \geq 0,$$

Thus, the m-eigenvalue is bounded in the sense that $m^2 \leq \lambda$, as required.

5. **Quantisation of** λ: we introduce the **ladder operators**

$$J_\pm = J_x \pm \mathrm{i}J_y,$$

such that

$$[J_z, J_\pm] = \pm\hbar J_\pm, \qquad [J^2, J_\pm] = 0$$

(this is easily shown by direct computation, applying the canonical commutation relation). Hence,

$$\begin{aligned} J_z J_\pm |\lambda, m\rangle &= (J_\pm J_z \pm \hbar J_\pm)|\lambda, m\rangle, \\ &= J_\pm (J_z \pm \hbar)|\lambda, m\rangle, \\ &= \hbar(m \pm 1) J_\pm |\lambda, m\rangle. \end{aligned}$$

Thus, if $|\lambda, m\rangle$ is a unit-norm eigenvector with J_z-eigenvalue m, then $J_\pm|\lambda, m\rangle$ is an eigenvector with J_z-eigenvalue $m + 1$:

$$J_\pm |\lambda, m\rangle = c_\pm(\lambda, m)|\lambda, m + 1\rangle.$$

From result 2, we know that $|m| \leq \lambda$. Now for m to possess a maximum value – which we call j – this procedure for stepping between consecutive subspaces must fail eventually:

$$J_+ |\lambda, j\rangle = 0,$$

which also implies that

$$J_- J_+ |\lambda, j\rangle = 0. \tag{17.3}$$

But

$$J_\pm J_+ = J^2 - J_z^2 - \hbar J_z.$$

Hence, Equation (17.3) reduces to

$$\hbar^2 \lambda - \hbar^2 j^2 - \hbar^2 j = 0,$$

or

$$\lambda = j(j + 1),$$

which is the form of the J^2-eigenvalues. By a similar argument, the minimum eigenvalue of J_z is $-j$, and

$$J_-|\lambda, -j\rangle = 0.$$

We operate repeatedly on the minimum state $|\lambda, -j\rangle$ with J_-. This produces states proportional to

$$|\lambda, -j\rangle, |\lambda, -j+1\rangle, \cdots, |\lambda, j-1\rangle, |\lambda, j\rangle.$$

This sequence has $2j+1$ elements – **hence, j must be an integer or a half-integer**.

We show that this list is exhaustive and includes all possible J_z-eigenvalues. We use a proof by contradiction: assume that there is a J_z-eigenvalue β, with

$$\beta \notin \{-j, -j+1, \cdots, j-1, j\}.$$

We act repeatedly on $|j, \beta\rangle$ to obtain a further J_z-eigenvalue α, such that

$$-j < \alpha < -j+1,$$

where the inequalities are strict. We now consider $J_-|j, \alpha\rangle$. Two possibilities arise:

(a) $J_-|j, \alpha\rangle \neq 0$. In this case, $J_-|j, \alpha\rangle$ is an eigenvector of J_z with eigenvalue $\alpha - 1$. However, this contradicts the fact that $-j$ is the minimum eigenvalue of J_z. This case cannot therefore occur.

(b) $J_-|j, \alpha\rangle = 0$. Then, $J_+J_-|j, \alpha\rangle = 0$ also, or

$$\left(J^2 - J_z^2 + \hbar J_z\right)|j, \alpha\rangle = 0.$$

But $|j, \alpha\rangle \neq 0$, hence

$$j(j+1) - \alpha(\alpha - 1) = 0,$$

with solution $\alpha = -j$. This contradicts the strictness of the inequalities $-j < \alpha < -j+1$. This case cannot therefore occur.

Indeed, the two cases are ruled out, which implies a contradiction. This means that $\beta \notin \{-j, -j+1, \cdots, j-1, j\}$ does not exist, which further implies that the set $\hbar\{-j, -j+1, \cdots, j-1, j\}$ is the complete set of J_z-eigenvalues.

Finally, for completeness, we derive the form of the constants of proportionality $c_\pm(j, m)$ (we replace the label λ with the label j from now on).

We have

$$\begin{aligned} J_+|j,m\rangle &= c_+(j,m)|j,m+1\rangle, \\ \langle j,m|J_P^\dagger &= c_+(j,m)^*\langle j,m+1|, \\ \langle j,m|J_- &= c_+(j,m)^*\langle j,m+1| \ \ldots \ J_- = J_+^\dagger. \end{aligned}$$

Hence,

$$\langle j,m|J_-J_+|h,m\rangle = |c_+(j,m)|^2\langle j,m+1|j,m+1\rangle = |c_+(j,m)|^2.$$

But $J_\pm J_+ = J^2 - J_z^2 - \hbar J_z$, hence

$$\begin{aligned} \langle j,m|J_-J_+|h,m\rangle &= \langle j,m|J^2 - J_z^2 - \hbar J_z|j,m\rangle, \\ &= \hbar^2 j(j+1) - \hbar^2 m - \hbar m, \\ &= \hbar^2\left[j(j+1) - m(m+1)\right], \\ &= |c_+(j,m)|^2. \end{aligned}$$

Taking c_+ to be real, we have

$$c_\pm(j,m) = \hbar\sqrt{j(j+1) - m(m\pm 1)}.$$

17.3 Representations

The following are matrix representations for the abstract algebra just defined:

- $j = 1/2$ **representation:** Consider again the **Pauli matrices**

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

  Form the angular-momentum operators

$$J_x = \frac{\hbar}{2}\sigma_x, \qquad J_y = \frac{\hbar}{2}\sigma_y, \qquad J_z = \frac{\hbar}{2}\sigma_z.$$

  We know that

$$\sigma^2 := \sigma_x^2 + \sigma_y^2 + \sigma_z^2 = 3\mathbb{I}, \qquad \sigma_z\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

  hence

$$J^2 = J_x^2 + J_y^2 + J_z^2 = \frac{3\hbar^2}{4} = \hbar^2\tfrac{1}{2}\left(\tfrac{1}{2}+1\right),$$

and

$$J_z \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \left(\tfrac{1}{2}\hbar\right) \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Hence, the Pauli matrices satisfy the commutation relation for the algebra of angular momentum for $j = 1/2$. The set $\{J_z, J^2\}$ is maximally commuting.

- $j = 1$ **representation:**

$$J_x = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad J_y = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \qquad J_z = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Form the Casimir operator

$$\begin{aligned} J^2 &= J_z^2 + J_y^2 + J_x^2, \\ &= \hbar^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \tfrac{1}{2}\hbar^2 \begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 \end{pmatrix} + \tfrac{1}{2}\hbar^2 \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \\ &= 2\hbar^2 \mathbb{I}. \end{aligned}$$

Identify $j(j+1) = 2$, hence $j = 1$. The matrices $\{J_x, J_y, J_z\}$ satisfy the CCR, and the set $\{J^2, J_z\}$ is maximally commuting.

A simultaneous basis for $\{J^2, J_z\}$ is the usual one:

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad |0\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad |-1\rangle = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix},$$

with

$$J_z|1\rangle = \hbar(+1)|1\rangle, \qquad J_z|0\rangle = \hbar(0)|0\rangle, \qquad J_z|-1\rangle = \hbar(-1)|0\rangle.$$

Chapter 18

Intrinsic angular momentum

Reading material for this chapter: Mandl, Chapters 2, 4, and 5

18.1 Overview

In the last chapter, we saw how to construct a matrix representation for angular-momentum quantum numbers $j = 1/2$ and $j = 1$:

- $j = 1/2$ **representation:**

$$\widehat{\mathrm{J}}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \widehat{\mathrm{J}}_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \widehat{\mathrm{J}}_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

  with

$$\widehat{\mathbf{J}}^2 = \tfrac{3}{2}\hbar^2\mathbb{I},$$

  and

$$\widehat{\mathrm{J}}_z|\pm\rangle = \hbar(\pm 1)|\pm\rangle.$$

- $j = 1$ **representation:**

$$\widehat{\mathrm{J}}_x = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \widehat{\mathrm{J}}_y = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \qquad \widehat{\mathrm{J}}_z = \hbar\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

  with

$$\widehat{\mathbf{J}}^2 \;=\; 2\hbar^2\mathbb{I},$$

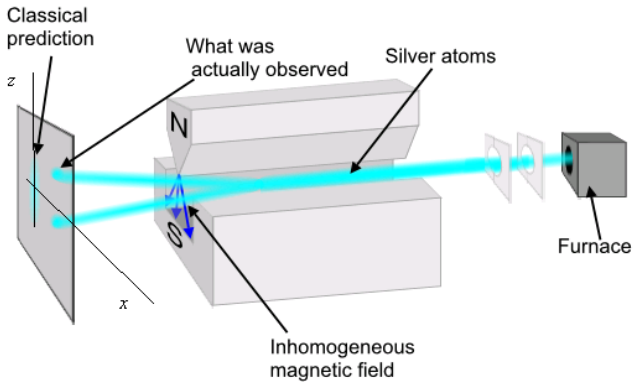


Figure 18.1: The Stern–Gerlach experiment

and

$$\widehat{\mathrm{J}}_z|1\rangle = \hbar(+1)|1\rangle, \qquad \widehat{\mathrm{J}}_z|0\rangle = \hbar(0)|0\rangle, \qquad \widehat{\mathrm{J}}_z|-1\rangle = \hbar(-1)|0\rangle.$$

It is tempting to ask, 'why bother'. Previously, we have found that the electron bound to the hydrogen atom possesses angular momentum $\hbar\sqrt{\ell(\ell+1)}$ because it 'orbits' a force centre – just as the earth orbiting the sun has angular momentum. In this chapter, we are going to find out that the electron also has **intrinsic angular momentum** that it possesses whether it is free or bound. As a loose analogy, compare this intrinsic angular momentum to the angular momentum of the earth due to its spinning on its axis[1]. Continuing with this analogy, we call intrinsic angular momentum **spin**.

18.2 Stern–Gerlach experiment

Consider the experimental setup shown in Fig. 18.1.

- A beam of (electrically neutral) atomic particles is passed through an inhomogeneous magnetic field $\boldsymbol{B}(\boldsymbol{x})$. The particles have a magnetic dipole moment $\boldsymbol{\mu}$, and the force experienced by the particles in the field is therefore

$$\boldsymbol{F} = -\nabla\mathcal{U} \qquad \mathcal{U} = -\boldsymbol{\mu}\cdot\boldsymbol{B},$$

or

$$F = \mu_z \frac{dB}{dz},$$

assuming the inhomogeneous direction coincides with the z-axis. The particles emerge from the device and are incident on an observation screen. Classically, one would expect to see a

[1] A very loose analogy, since the electron has no spatial extent, and thus, the formula $J = mrv$ ought to yield zero.

continuous pattern, spread along the z-axis, and symmetric about $z = 0$. If the particles have zero magnetic moment, one would expect to see a diffuse spot at the centre of the screen.

- Instead, one finds patterns of $1, 2, 3, \cdots$ discrete spots about the undeflected direction $z = 0$. The original beam splits into several beams. The number of spots depends on the **intrinsic angular momentum** of the atoms in the beam.

- Classically, it can be shown that the **magnetic dipole moment** of a particle with angular momentum is

$$\boldsymbol{\mu} = \frac{Q}{2m}\boldsymbol{J}, \qquad (*)$$

where $\boldsymbol{J}$ is the angular momentum and Q is the charge of the particle (but see the last point for the quantum-mechanical correction to this formula). If the angular momentum is quantised, with principal quantum number j, then there are only $2j+1$ possible values for the projection of the angular momentum on to the z-axis, and hence, only $2j + 1$ possible values for the magnetic force. Thus, the beam splits into $2j + 1$ sub-beams, each observed on the screen.

- This experiment can be repeated with electrons. The experiment is adjusted to take account of the Lorentz force – the fact that charged particles experience a force perpendicular to the motion, to prevent the electrons from being deflected. One finds only **two** spots on the observation screen – suggesting that the electrons have an intrinsic angular momentum with principal quantum number $2j + 1 = 2 \implies j = 1/2$.

- For electrons, the intrinsic magnetic moment associated with the spin $\boldsymbol{S}$ is given by

$$\boldsymbol{\mu} = \frac{-ge}{2m_{\mathrm{e}}}\boldsymbol{S},$$

where g is Dirac's g-factor, and $g \approx 2$. This is a consequence of solving the relativistic wave equation for the electron, a problem you might encounter in later modules.

18.3 Identical particles

Consider two particles with identical mass, charge, spin, etc. Classically, we can identify the particles by their position, and they can therefore be distinguished one from another. However, the uncertainty principle means that we cannot do this under quantum mechanics. Thus, the particles are identical. This has implications for the form of the wavefunction of the two-particle system.

Let $\Psi(1, A; 2, B)$ be the wavefunction of the two-particle system. This means that particle 1 occupies state A and particle 2 occupies state B. Consider the operation

$$\Psi(1, A; 2, B) \to E\Psi(1, A; 2, B) = \Psi(2, A; 1, B),$$

which means that particle 2 is now in state A, and particle 1 is in state B. Mathematically, we have

$$E^2\Psi(1, A; 2, B) = \Psi(1, A; 2, B). \tag{18.1}$$

Now, the fact that the particles are identical means that the wavefunction should be an eigenfunction of the exchange operator, such that

$$\begin{aligned} E\Psi(1, A; 2, B) &= \Psi(2, A; 1, B), \\ &= \lambda\Psi(1, A; 2, B); \end{aligned}$$

Equation (**??**) demonstrates that $\lambda^2 = 1$, hence

$$\lambda = \pm 1.$$

Thus, the state vector of a pair of identical particles is either **symmetric** under exchange:

$$\Psi(2, A; 1, B) = +\Psi(1, A; 2, B)$$

OR **antisymmetric** under exchange:

$$\Psi(2, A; 1, B) = -\Psi(1, A; 2, B)$$

This provides a useful classification of types of particles:

- Particles whose wavefunctions are **symmetric** under exchange are called **Bosons**;
- Particles whose wavefunctions are **antisymmetric** under exchange are called **Fermions**.

It also happens that

- Bosons have integer spin;
- Fermions have half-integer spin.

This is called the **spin-statistics theorem**, and is a consequence of quantum field theory.

Examples:

- Bosons: pions (spin zero), photons (spin 1).
- Fermions: protons, neutrons, electrons (all spin $1/2$).

18.4 Pauli's exclusion principle

Consider two identical, non-interacting Fermions. By the second postulate, the state of the composite system is formed by tensor products, such as

$$\Psi(1, A; 2, B) = \psi(A, 1)\psi(B; 2).$$

which means that particle 1 is in single-particle state A, and particle 2 is in single-particle state B. However, the correct wavefunction is antysymmetric:

$$\Psi(1, A; 2, B) = \psi(A, 1)\psi(B; 2) - \psi(B, 1)\psi(A, 2),$$

which is a state wherein particle 1 is in state A and particle 2 is in state B, OR, particle 2 is in state A and particle in is in state B (we cannot tell these situations apart). Now let's compute the wavefunction corresponding to both particles occupying the single-particle state A, it is

$$\Psi(1, A; 2, A) = \psi(A, 1)\psi(A, 2) - \psi(A, 1)\psi(A, 2) = 0,$$

and **the probability that both fermions occupy the same single-particle state A is identically zero**. This is Pauli's exclusion principle.

Example: Write down the ground-state wavefunction for a composite system comprising two non-interacting, spin-$1/2$ Fermions.

The answer involves two parts: A spatial wavefunction, for the spatial degrees of freedom, and a spin state. Because the spin- and spatial-degrees of freedom do not interact, the total wavefunction is a product:

$$\Psi(1, 2) = \psi(\boldsymbol{r}_1, \boldsymbol{r}_2)S(1, 2),$$

where $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ is the spatial wavefunction and $S(1, 2)$ is a spin state.

The total wavefunction $\Psi(1, 2)$ is antisymmetric under exchange. This means:

- The spatial wavefunction is symmetric and the spin state is antisymmetric, OR
- The spatial wavefunction is antisymmetric and the spin state is symmetric.

We consider **case 1** first. Let's look at the spin states. Particle 1 can be in a state $|+, 1\rangle$ and particle 2 can be in a state $|-, 1\rangle$, or vice versa, but it is not possible for both states to be in the same spin state. Thus

$$|1, 2\rangle = |+, 1\rangle|-, 2\rangle - |-, 1\rangle|+, 2\rangle,$$

or

$$|1,2\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \otimes \begin{pmatrix} 0 \\ -1 \end{pmatrix}_2 - \begin{pmatrix} 0 \\ -1 \end{pmatrix}_1 \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2 .$$

Hence,

$$\Psi(1,2) = \frac{1}{\sqrt{2}} \psi_{\text{symmetric}}(\boldsymbol{r}_1, \boldsymbol{r}_2) \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \otimes \begin{pmatrix} 0 \\ -1 \end{pmatrix}_2 - \begin{pmatrix} 0 \\ -1 \end{pmatrix}_1 \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2 \right].$$

Now let us examine **case 2**. We can form three symmetric spin states from the single-particle spin states:

$$\begin{aligned} \text{Both particles spin up} &= |1,+\rangle|2,+\rangle, \\ \text{Both particles spin down} &= |1,-\rangle|2,-\rangle, \\ \text{Spin up, down, symmetric combo.} &= |1,+\rangle|2,-\rangle + |1,-\rangle|2,\rangle, \end{aligned}$$

or

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2, \qquad \begin{pmatrix} 0 \\ -1 \end{pmatrix}_1 \otimes \begin{pmatrix} 0 \\ -1 \end{pmatrix}_2,$$
$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \otimes \begin{pmatrix} 0 \\ -1 \end{pmatrix}_2 + \begin{pmatrix} 0 \\ -1 \end{pmatrix}_1 \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2,$$

respectively. Thus, the other possible form for the composite wavefunction is

$$\begin{aligned} \Psi(1,2) &= \psi_{\text{antisymmetric}}(\boldsymbol{r}_1, \boldsymbol{r}_2) \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2 \right], \\ &\text{or } \psi_{\text{antisymmetric}}(\boldsymbol{r}_1, \boldsymbol{r}_2) \left[0 \begin{pmatrix} 0 \\ -1 \end{pmatrix}_1 \otimes \begin{pmatrix} 0 \\ -1 \end{pmatrix}_2 \right], \\ &\text{or } \frac{1}{\sqrt{2}} \psi_{\text{antisymmetric}}(\boldsymbol{r}_1, \boldsymbol{r}_2) \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \otimes \begin{pmatrix} 0 \\ -1 \end{pmatrix}_2 + \begin{pmatrix} 0 \\ -1 \end{pmatrix}_1 \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2 \right]. \end{aligned}$$

Now we focus on the spatial part of the wavefunction; in particular, we examine the ground state. We cannot make progress without specifying the details of the potential field. We focus for simplicity

on the one-dimensional case, and we assume

$$\begin{aligned}\hat{H} &= \hat{H}_1 + \hat{H}_2,\\ \hat{H}_1 &= -\frac{\hbar^2}{2m}\partial_{x_1}^2 + \mathcal{U}(x_1),\\ \hat{H}_2 &= -\frac{\hbar^2}{2m}\partial_{x_2}^2 + \mathcal{U}(x_2).\end{aligned}$$

The eigenvalue problem $\hat{H}\psi(x_1, x_2) = E\psi(x_1, x_2)$ therefore separates, and the ground-state of the system is a product of two single-particle ground states:

$$\psi_{\rm gs}(x_1, x_2) = \psi_{0\rm gs}(x_1)\psi_{0\rm gs}(x_2),$$

up to exchange symmetry, with eigenvalue

$$E_{\rm gs} = 2E_{0\rm gs}.$$

However, no antisymmetric ground-state wavefunction exists, since by definition of the ground state, both particles occupy the minimum energy level. Thus, the spatial wavefunction in the ground state is necessarily symmetric, and the total wavefunction is

$$\Psi_{\rm gs}(1,2) = \frac{1}{\sqrt{2}}\psi_{0\rm gs}(x_1)\psi_{0\rm gs}(x_2)\left[\begin{pmatrix}1\\0\end{pmatrix}_1 \otimes \begin{pmatrix}0\\-1\end{pmatrix}_2 + \begin{pmatrix}0\\-1\end{pmatrix}_1 \otimes \begin{pmatrix}1\\0\end{pmatrix}_2\right].$$

Thus, the ground state consists of a spin-up, spin-down pair of fermions. Such a pair is called a **singlet state**.

18.5 The periodic table

We discuss many-electrons atoms now, and outline how the periodic table follows from consideration of the hydrogen atom, and Pauli's exclusion principle.

- Consider first the dynamics of a single electron in the system. In a crude approximation, one ignores the repulsive interactions between the electron and its neighbours, and treats the electron as though it experiences a Coulomb force arising from the positive nucleus, of strength Ze (Z is the number of protons in the nucleus).

- At the next level of approximation, one takes account of the repulsive force between the electrons by supposing that the single electron of interest does not experience the 'bare' Coulombic force, but rather one diminished or 'screened' by the fact that a cloud of negative

electrons surrounds the positive core. This approximation can be described by a central potential.

- Thus, we are reduced again to the problem of motion in a central potential, using a screened potential.
- We compute the energy levels of this potential using the theory derived in Ch. 15. The potential is not the simple Coulombic one, and the energy levels are non-dengenerate with respect to the angular-momentum quantum number, $E_n \to E_{n\ell}$.
- It is possible, in this description, to write down the energy levels from the lowest to the highest. These are:

  $$1s,\ 2s,\ 2p,\ 3s,\ 3p,\ [4s,\ 3d],\ 4p,\ [5s,\ 4d],\ 5p, \cdots$$

  (the bracketed terms have very similar energies). For a given n, the energy $E_{n\ell}$ increases as ℓ increases: the large values of angular momentum create a 'centrifugal barrier' which prevents the electron from entering into the core region. When small-ℓ electrons enter this region, they sample the unscreened nuclear charge, which leads to a higher energy binding the electron to the nucleus.
- We fill each single-particle state or **orbital** starting with the ground state.
- From the previous section, we can fit two electrons into the ground state – one is spin up, and the other is spin down.
- The next state is also an s state, with $\ell = 0$, so it too can hold two electrons.
- The next state is a p state, with $\ell = 1$. This is $2\ell + 1 = 3$-fold degenerate, and can hold six $(= 3 \times 2)$ electrons.
- The most stable and non-reacting elements have **closed shells**, where **each energy level** is filled with electrons. The energy gap between filled shells and the next available 'slot' is large.
- For example, $Z = 2$ has $2(1s)$ electrons, and forms a closed shell. This is Helium.
- The next such atom has $Z = 10$, with the first shell closed $2(1s)$, as well as the second ($2(2s)$ and $6(2p)$) states. This is Neon.

Chapter 19

Addition of angular momenta

Reading material for this chapter: Mandl, Chapter 5

In this chapter we find the resultant angular momentum of a composite system where each component of the system has its own angular momentum.

Let $\widehat{\mathbf{J}}_1$, $\widehat{\mathbf{J}}_2$ be two independent angular momenta that satisfy the canonical commutation relation (CCR):

$$[\widehat{\mathrm{J}}_{nx}, \widehat{\mathrm{J}}_{ny}] = \mathrm{i}\hbar\widehat{\mathrm{J}}_{nz} \ + \text{CPs}, \qquad [\widehat{\mathrm{J}}_{nx}, \widehat{\mathrm{J}}_n^2] = 0, \ \text{\&c.},$$

where $n = 1, 2$ labels the subsystems. Because the subsystems are independent, we have

$$[\widehat{\mathrm{J}}_{1x}, \widehat{\mathrm{J}}_{2,y}] = 0, \ \text{\&c.}.$$

This is true for the spins of two particles. It is also true for a spin and an orbital angular momentum, since these operators act on different degrees of freedom.

We define a total angular momentum

$$\widehat{\mathbf{J}} := \widehat{\mathbf{J}}_1 \otimes \mathbb{I}_2 + \mathbb{I}_1 \otimes \widehat{\mathbf{J}}_2,$$

which acts on tensor products of the composite system.

Example: Consider again a composite system formed from two uncharged spin-$1/2$ Fermions. Recall the **singlet state**, which has one spin-up component and one spin-down component, in an antisymmetric form:

$$|\text{Singlet}\rangle = |+,1\rangle|-,2\rangle - |-,1\rangle|+,2\rangle,$$

or, in matrix representation,

$$|1,2\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \otimes \begin{pmatrix} 0 \\ -1 \end{pmatrix}_2 - \begin{pmatrix} 0 \\ -1 \end{pmatrix}_1 \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2.$$

This is an eigenstate of angular momentum: The angular momentum along the z-direction is given by the operator

$$\widehat{\mathrm{J}}_z = \widehat{\mathrm{J}}_{1z} \otimes \mathbb{I} + \mathbb{I} \otimes \widehat{\mathrm{J}}_{2z},$$

such that

$$\begin{aligned}
\widehat{\mathrm{J}}_z|1,2\rangle &= \left(\widehat{\mathrm{J}}_{1z} \otimes \mathbb{I} + \mathbb{I} \otimes \widehat{\mathrm{J}}_{2z}\right) \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \otimes \begin{pmatrix} 0 \\ -1 \end{pmatrix}_2 - \begin{pmatrix} 0 \\ -1 \end{pmatrix}_1 \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2 \right], \\
&= \left(\widehat{\mathrm{J}}_{1z} \otimes \mathbb{I}\right) \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \otimes \begin{pmatrix} 0 \\ -1 \end{pmatrix}_2 - \begin{pmatrix} 0 \\ -1 \end{pmatrix}_1 \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2 \right] \\
&\quad \left(\mathbb{I} \otimes \widehat{\mathrm{J}}_{2z}\right) \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \otimes \begin{pmatrix} 0 \\ -1 \end{pmatrix}_2 - \begin{pmatrix} 0 \\ -1 \end{pmatrix}_1 \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2 \right] \\
&= \left[\widehat{\mathrm{J}}_{1z} \begin{pmatrix} 1 \\ 0 \end{pmatrix}_1\right] \otimes \begin{pmatrix} 0 \\ -1 \end{pmatrix}_2 - \left[\widehat{\mathrm{J}}_{1z} \begin{pmatrix} 0 \\ -1 \end{pmatrix}_1\right] \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2, \\
&\quad \begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \otimes \left[\widehat{\mathrm{J}}_{2z} \otimes \begin{pmatrix} 0 \\ -1 \end{pmatrix}_2\right] - \begin{pmatrix} 0 \\ -1 \end{pmatrix}_1 \otimes \left[\widehat{\mathrm{J}}_{2z} \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2\right], \\
&= \frac{\hbar}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \otimes \begin{pmatrix} 0 \\ -1 \end{pmatrix}_2 + \frac{\hbar}{2} \begin{pmatrix} 0 \\ -1 \end{pmatrix}_1 \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2 \\
&\quad - \frac{\hbar}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \otimes \begin{pmatrix} 0 \\ -1 \end{pmatrix}_2 - \frac{\hbar}{2} \begin{pmatrix} 0 \\ -1 \end{pmatrix}_1 \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2, \\
&= 0, \\
&= 0|1,2\rangle.
\end{aligned}$$

Thus, the composite state has zero total angular momentum along the z-direction. Note also, that using formal tensor-product notation is very cumbersome, so instead, we will be more informal, and use notation such as

$$\widehat{\mathbf{J}} := \widehat{\mathbf{J}}_1 + \widehat{\mathbf{J}}_2,$$

for the tensor-product operator.

Having defined the addition of angular-momentum operators, $\widehat{\mathbf{J}} := \widehat{\mathbf{J}}_1 + \widehat{\mathbf{J}}_2$, note that $\{\widehat{\mathrm{J}}_x, \widehat{\mathrm{J}}_y, \widehat{\mathrm{J}}_z\}$ satisfy the CCR:

$$[\widehat{\mathrm{J}}_x, \widehat{\mathrm{J}}_y] = \mathrm{i}\hbar\widehat{\mathrm{J}}_z \ + \text{CPs.}$$

Note also the existence of a square operator:

$$\widehat{\mathrm{J}}^2 = \widehat{\mathrm{J}}_x^2 + \widehat{\mathrm{J}}_y^2 + \widehat{\mathrm{J}}_z^2,$$

that satisfies

$$[\widehat{\mathrm{J}}^2, \widehat{\mathrm{J}}_z] = 0, \text{ \&c.}$$

We also have

$$[\widehat{\mathrm{J}}^2, \widehat{\mathrm{J}}_1^2] = [\widehat{\mathrm{J}}^2, \widehat{\mathrm{J}}_2^2].$$

As a consequence of these commutation relations, the results of Ch. 17 apply: $\{\widehat{\mathrm{J}}_z, \widehat{\mathrm{J}}^2\}$ are simultaneously diagonalisable, $\widehat{\mathrm{J}}^2$ has eigenvalues $J(J+1)$, where J is integral or half-integral, and $\widehat{\mathrm{J}}_z$ has eigenvalues M, where

$$M = -J, -J+1, \cdots, J-1, J.$$

We have not yet been able to determine what are the allowed values of J. We do this now. The result is called the **angular momentum addition theorem**.

Theorem 19.1 *Let $\widehat{\mathbf{J}}_1$, $\widehat{\mathbf{J}}_2$ be two independent angular momenta that satisfy the CCR, and let j_1 and j_2 be the eigenvalues of $\widehat{\mathbf{J}}_1^2$ and $\widehat{\mathbf{J}}_2^2$, respectively. Form the sum*

$$\widehat{\mathbf{J}} = \widehat{\mathbf{J}}_1 + \widehat{\mathbf{J}}_2.$$

Then the eigenvalues of $\widehat{\mathbf{J}}^2$ take only discrete values $\hbar^2 J(J+1)$,

$$J = |j_1 - j_2|, \, |j_1 - j_2| + 1, \ldots \, j_1 + j_2.$$

Proof: Observe that the composite system possesses two complete sets of commutating observables:

1. The set $\{\widehat{\mathrm{J}}_1^2, \widehat{\mathrm{J}}_2^2, \widehat{\mathrm{J}}_{1z}, \widehat{\mathrm{J}}_{2z}\}$ is a maximally commuting set with a simultaneous basis of eigenvectors given by tensor products

   $$\{|j_1, m_1\rangle |j_2, m_2\rangle\}_{(j_1, j_2, m_1, m_2)},$$

   where $m_1 = -j_1, \cdots, j_1$ &c. We will also denote this basis by

   $$|j_1, m_1\rangle |j_2, m_2\rangle \equiv |j_1, m_1; j_2, m_2\rangle.$$

2. The set $\{\widehat{\mathrm{J}}_z, \widehat{\mathrm{J}}^2, \widehat{\mathrm{J}}_1^2, \widehat{\mathrm{J}}_2^2\}$ is also a maximally commuting set, with a simultaneous basis of eigenvectors

   $$\{|j_1, j_2, J, M\rangle\}_{(j_1, j_2, J, M)}.$$

Note that the two maximally commuting sets are incompatible. Basis (1) implies the completeness

relation

$$\sum_{m_1}\sum_{m_2}|j_1,m_1;j_2,m_2\rangle\langle j_1,m_1;j_2,m_2|=\mathbb{I},$$

while basis (2) implies the relation

$$\sum_{J}\sum_{M}|j_1,j_2,J,M\rangle\langle j_1,j_2,J,M|=\mathbb{I}.$$

Thus,

$$|j_1,j_2,J,M\rangle=\sum_{m_1}\sum_{m_2}\langle j_1,m_1;j_2,m_2|j_1,j_2,J,M\rangle|j_1,m_1;j_2,m_2\rangle.$$

We call

$$C(j_1,j_2,m_1,m_2;J,M):=\langle j_1,m_1;j_2,m_2|j_1,j_2,J,M\rangle$$

the **Clebsh–Gordon coefficient**.

Now $C(j_1,j_2,m_1,m_2;J,M)=0$ unless $M=m_1+m_2$. For, by definition, $\widehat{\mathrm{J}}_z=\widehat{\mathrm{J}}_{1z}+\widehat{\mathrm{J}}_{2z}$, hence

$$\left(\widehat{\mathrm{J}}_z-\widehat{\mathrm{J}}_{1z}-\widehat{\mathrm{J}}_{2z}\right)|j_1,j_2,J,M\rangle=0.$$

We pair this with the $\langle j_1,m_1;j_2,m_2|$:

$$\begin{aligned}0&=\langle j_1,m_1;j_2,m_2|\left(\widehat{\mathrm{J}}_z-\widehat{\mathrm{J}}_{1z}-\widehat{\mathrm{J}}_{2z}\right)|j_1,j_2,J,M\rangle,\\&=\hbar(M-m_1-m_2)\langle j_1,m_1;j_2,m_2|j_1,j_2,J,M\rangle,\\&=\hbar(M-m_1-m_2)C(j_1,j_2,m_1,m_2;J,M),\end{aligned}$$

which forces $C=0$ unless $M=m_1+m_2$. Thus, the maximum possible value of M is j_1+j_2, which coincides with the maximum possible value of J,

$$\max(J)=j_1+j_2.$$

Note that there are $(2j_1+1)$ possible m_1-values, and $(2j_2+1)$ possible m_2-values. Thus, the dimension of the vector space is $(2j_1+1)(2j_2+1)$, and there are $(2j_1+1)(2j_2+1)$ possible kets:

$$|j_1,m_1;j_2,m_2\rangle$$

(these are the type-1 kets from the first CSCO). However,

$$|j_1,j_2;J,M\rangle_{(j_1,j_2,J,M)}$$

is an equally good basis, so there must be $(2j_1+1)(2j_2+1)$ kets of this form, too (type 2). For

each J value there are $2J+1$ type-2 kets. Thus, there are

$$\sum_{J \text{ allowed}} 2J+1$$

type-2 kets, or

$$\sum_{J_{\min}}^{j_1+j_2} 2J+1$$

type-2 kets. But the number of type-1 and type-2 kets is the same:

$$\sum_{J_{\min}}^{j_1+j_2} 2J+1 = (2j_1+1)(2j_2+1).$$

The only solution to this equation is

$$J_{\min} = |j_1 - j_2|.$$

This concludes the proof.

Note: There is a general method for computing the Clesbsh–Gordon coefficients, but it is very unwieldy. For small-angular-momentum quantum numbers, the addition can be done in an intuitive fashion. We have already seen how to combine orbitals to obtain the angular momentum states of a pair of spin-$1/2$ fermions. We now look at a similar example.

Example: An electron bound to a nucleus is in an p state. Compute its total angular momentum.

Solution: The p state corresponds to $\ell = 1$. Thus, we must add the orbital angular momentum $\widehat{\mathbf{L}}$ with the spin angular momentum $\widehat{\mathbf{S}}$ and obtain a total angular momentum $\widehat{\mathbf{J}}$.

Two possibilities for J: $J = 3/2$ OR $J = 1/2$.

Consider case 1 first, $J = 3/2$. Then, the **top state** had $M = m_1 + m_2 = 3/2$, which implies $s_z = 1/2$ and $\ell_z = 1$. The only way for this to happen is for the electron to be spin up along the z-axis, and for the projection of orbital angular momentum along the z-axis to be positive. Thus, this state has the ket

$$|J = 3/2, M = 3/2\rangle = |+\rangle Y_{1,1}(\theta, \varphi).$$

To obtain lower states, act on this with the ladder operator

$$J_- = S_- \otimes \mathbb{I}_L + \mathbb{I}_s \otimes L_-,$$

where

$$S_-|+\rangle = |-\rangle, \qquad S_-|-\rangle = 0,$$

and where

$$\begin{aligned} L_-Y_{1,1} &= \sqrt{2}Y_{1,0}, \\ L_-Y_{1,0} &= \sqrt{2}Y_{1,-1}, \\ L_-Y_{1,-1} &= 0. \end{aligned}$$

Thus,

$$\begin{aligned} |3/2, M = 1/2\rangle &\propto J_- \left[|+\rangle Y_{1,1}\right], \\ &= |-\rangle Y_{1,1} + \sqrt{2}|+\rangle Y_{1,0}. \end{aligned}$$

Normalise:

$$|3/2, M = 1/2\rangle = \tfrac{1}{\sqrt{3}}\left[|-\rangle Y_{1,1} + \sqrt{2}|+\rangle Y_{1,0}\right].$$

Act again on this state with the lowering operator:

$$\begin{aligned} |3/2, M = -1/2\rangle &\propto |-\rangle L_-Y_{1,1} + \sqrt{2}|-\rangle Y_{1,0} + \sqrt{2}|+\rangle L_-Y_{1,0}, \\ &= \sqrt{2}|-\rangle Y_{1,0} + \sqrt{2}|-\rangle Y_{1,0} + 2|+\rangle Y_{1,-1}, \\ &= 2\sqrt{2}|-\rangle Y_{1,0} + 2|+\rangle Y_{1,-1}. \end{aligned}$$

Normalise:

$$|3/2, M = -1/2\rangle = \tfrac{1}{\sqrt{3}}\left[\sqrt{2}|-\rangle Y_{1,0} + |+\rangle Y_{1,-1}\right].$$

To find the **bottom state** we can act with the lowering operator again, or note simply that in this state, $m_s = -1/2$, and $m_\ell = -1$, hence,

$$|3/2, M = -3/2\rangle = |-\rangle Y_{1,-1}.$$

To find the $J = 1/2$ eigenstates, we start with the top state, with $M = 1/2$. In this state, $m_s = 1/2$ and $m_\ell = 0$ OR $m_s = -1/2$ and $m_\ell = 1$. Thus,

$$\langle 1/2, M = 1/2| = \rangle \alpha|+\rangle Y_{1,0} + \beta|-\rangle Y_{1,1},$$

and this state has no overlap with the other state built out of these product vectors:

$$\langle 3/2, 1/2|3/2, 1/2\rangle = 0,$$

or

$$\left[\sqrt{2}\langle +|Y_{1,0} + \langle -|Y_{1,1}\right] [\alpha|+\rangle Y_{1,0} + \beta|-\rangle Y_{1,1}] = 0,$$

hence

$$\beta = -\alpha\sqrt{2},$$

and

$$|1/2, 1/2\rangle = \tfrac{1}{\sqrt{3}} \left[|+\rangle Y_{1,0} - \sqrt{2}|-\rangle Y_{1,1}\right],$$

which is normalised.

Finally, the bottom state can be found using the lowering operator:

$$\begin{aligned} |1/2, -1/2\rangle &\propto |-\rangle Y_{1,0} + |+\rangle L_- Y_{1,0} - \sqrt{2}|-\rangle L_- Y_{1,1}, \\ &= |-\rangle Y_{1,0} + \sqrt{2}|+\rangle Y_{1,-1} - \sqrt{2}\sqrt{2}|-\rangle Y_{1,0}, \\ &= -|-\rangle Y_{1,0} + \sqrt{2}|+\rangle Y_{1,-1}, \end{aligned}$$

such that

$$|1/2, -1/2\rangle = \tfrac{1}{\sqrt{3}} \left[\sqrt{2}|+\rangle Y_{1,-1} - |-\rangle Y_{1,0}\right],$$

which is normalised and orthogonal to $|3/2, -1/2\rangle$:

$$\langle 3/2, -1/2|1/2, -1/2\rangle = 0.$$

In this course, you will only encounter simple problems like this one, where the angular momentum states can be worked out intuitively.

Chapter 20

Time-independent perturbation theory: non-degenerate case

Reading material for this chapter: Mandl, Chapter 7

20.1 The idea

In this section we focus again on solving the Schrödinger equation for time-independent systems. Recall that such systems reduce to an eigenvalue problem for the energy:

$$\hat{H}|\psi\rangle = E|\psi\rangle. \tag{20.1}$$

There are not many such problems that are exactly solvable. In fact, in this course we have considered most of them. Thus, it is helpful to consider approximate methods for general problems of the type (20.1)

The first such method we consider is **time-independent perturbation theory**, in the non-degenerate setting. Suppose that the problem

$$\hat{H}_0|\phi\rangle = E^{(0)}|\phi\rangle$$

is exactly solvable, with a discrete spectrum

$$E = E_1^{(0)}, E_2^{(0)}, \cdots,$$

that is non-degenerate:

$$E_i^{(0)} = E_j^{(0)} \implies i = j;$$

in other words,

$$E_i^{(0)} = E_j^{(0)} \implies |E_i^{(0)}\rangle = |E_j^{(0)}\rangle.$$

For definiteness, assume that the eigenvectors are the kets $\{|\phi_n\rangle\}_{n=1}^{\infty}$. We now focus on solving the **perturbed problem**

$$\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle, \qquad \hat{H} = \hat{H}_0 + \lambda V,$$

where λ is a **small dimensionless parameter**. If the perturbation is 'nice', and we assume it is, then the Hilbert space for the unperturbed and perturbed problems is the same:

$$\mathcal{H}(\hat{H}_0) = \mathcal{H}(\hat{H}).$$

Thus, by the completeness property of the basis $\{|\phi_n\rangle\}_{n=1}^{\infty}$, we may expand the solution $|\psi_n\rangle$ of the perturbed problem in terms of a known set of states. This is the subject of this chapter.

20.2 The method

We are to solve

$$\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle, \qquad \hat{H} = \hat{H}_0 + \lambda V,$$

given that λ is a small parameter. We pose the series solution

$$\begin{aligned} E_n &= \sum_{p=0}^{\infty} \lambda^p e_{np}, \\ |\psi_n\rangle &= \sum_{p=0}^{\infty} \lambda^p |\phi_{np}\rangle. \end{aligned}$$

We assume wlog that the states $|\phi_{n1}\rangle, |\phi_{n2}\rangle, \cdots$ are orthogonal to the n^{th} eigenstate of the unperturbed problem:

$$\langle\phi_n|\phi_{n1}\rangle = \langle\phi_n|\phi_{n2}\rangle = \cdots = 0.$$

Next, we substitute our trial solution into the eigenvalue problem:

$$\begin{aligned} (\hat{H}_0 + \lambda V)|\psi_n\rangle &= (\hat{H}_0 + \lambda V)\sum_{p=0}^{\infty} \lambda^p |\phi_{np}\rangle, \\ &= E_n|\psi_n\rangle, \\ &= \left(\sum_{p=0}^{\infty} \lambda^p e_{np}\right)\left(\sum_{p=0}^{\infty} \lambda^p |\phi_{np}\rangle\right). \end{aligned}$$

In other words,

$$(\hat{H}_0 + \lambda V) \sum_{p=0}^{\infty} \lambda^p |\phi_{np}\rangle = \left(\sum_{p=0}^{\infty} \lambda^p e_{np} \right) \left(\sum_{p=0}^{\infty} \lambda^p |\phi_{np}\rangle \right).$$

This is a power-series identity in λ; the identity must hold term-by-term.

We examine the **zeroth-order term**:

$$\hat{H}_0 |\phi_{n0}\rangle = e_{n0} |\phi_{n0}\rangle.$$

Thus, $|\phi_{n0}\rangle = |\phi_n\rangle$ and $e_{n0} = E_n^{(0)}$, and the zeroth-order problem is exactly the same as the unperturbed problem.

Next, we examine the **first-order term**:

$$\lambda \hat{H}_0 |\phi_{n1}\rangle + \lambda V |\phi_{n0}\rangle = \lambda e_{n0} |\phi_{n1}\rangle + \lambda e_{n1} |\phi_{n0}\rangle.$$

Dividing out by λ and using the information about the zeroth-order terms, this becomes

$$\hat{H}_0 |\phi_{n1}\rangle + V |\phi_n\rangle = E_n^{(0)} |\phi_{n1}\rangle + e_{n1} |\phi_n\rangle.$$

Re-arrange:

$$\left(\hat{H}_0 - E_n^{(0)} \right) |\phi_{n1}\rangle = (-V + e_{n1}) |\phi_n\rangle.$$

We take the scalar product of this identity with $|\phi_n\rangle$:

$$\langle \phi_n | \left(\hat{H}_0 - E_n^{(0)} \right) |\phi_{n1}\rangle = \langle \phi_n | (-V + e_{n1}) |\phi_n\rangle.$$

Consider the LHS. The operator $\hat{H}_0$ is Hermitian, so we can choose for it to operate on the bra instead of the ket. But the bra is an eigenstate of $\hat{H}_0$, so the identity becomes

$$\langle \phi_n | \left(E_n^{(0)} - E_n^{(0)} \right) |\phi_n\rangle = 0 = \langle \phi_n | (-V + e_{n1}) |\phi_n\rangle.$$

In other words,

$$e_{n1} = \langle \phi_n | V | \phi_n \rangle,$$

and

$$E_n = E_n^{(0)} + \lambda \langle \phi_n | V | \phi_n \rangle + O(\lambda^2).$$

To compute the corrected state vector, we consider the identity

$$\left(\hat{H}_0 - E_n^{(0)}\right)|\phi_{n1}\rangle = \left(-V + e_{n1}\right)|\phi_n\rangle$$

again. Take its scalar product with $\langle\phi_p|$, with $p \neq n$. Thus,

$$\langle\phi_p|\left(\hat{H}_0 - E_n^{(0)}\right)|\phi_{n1}\rangle = \langle\phi_p|\left(-V + e_{n1}\right)|\phi_n\rangle$$

We expand the correction $|\phi_{n1}\rangle$ in terms of the basis elements $|\phi_n\rangle$:

$$|\phi_{n1}\rangle = \sum_{n\neq p} A_{np}|\phi_p\rangle.$$

Combining the last two equations, we have

$$\begin{aligned}
\langle\phi_p|\left(\hat{H}_0 - E_n^{(0)}\right)\left(\sum_{n\neq q} A_{nq}|\phi_q\rangle\right) &= \langle\phi_p|\left(-V + e_{n1}\right)|\phi_n\rangle, \\
\left(E_p^{(0)} - E_n^{(0)}\right)\sum_{n\neq q} A_{nq}\langle\phi_p|\phi_q\rangle &= -\langle\phi_p|V|\phi_n\rangle, \\
A_{np}\left(E_p^{(0)} - E_n^{(0)}\right) &= -\langle\phi_p|V|\phi_n\rangle, \\
A_{np} &= -\frac{\langle\phi_p|V|\phi_n\rangle}{E_p^{(0)} - E_n^{(0)}}, \qquad p \neq n.
\end{aligned}$$

Thus,

$$\begin{aligned}
|\phi_{n1}\rangle &= \sum_{n\neq p} A_{np}|\phi_p\rangle, \\
&= \sum_{n\neq p} \frac{\langle\phi_p|V|\phi_n\rangle}{E_n^{(0)} - E_p^{(0)}}|\phi_p\rangle,
\end{aligned}$$

$$|\psi_n\rangle = |\phi_n\rangle + \lambda|\phi_{n1}\rangle + O(\lambda^2),$$

and

$$|\psi_n\rangle = |\phi_n\rangle + \lambda\sum_{n\neq p} \frac{\langle\phi_p|V|\phi_n\rangle}{E_n^{(0)} - E_p^{(0)}}|\phi_p\rangle + O(\lambda^2).$$

We pass on to **second-order perturbation theory**, and derive the corrected energy only. At second order, the power series in λ yields

$$\lambda^2\hat{H}_0|\phi_{n2}\rangle + \lambda^2 V|\phi_{n1}\rangle = \lambda^2 e_{n2}|\phi_{n0}\rangle + \lambda e_{n1}|\phi_{n1}\rangle + \lambda^2 e_{n0}|\phi_{n2}\rangle.$$

We divide out by λ^2 and use the information supplied by the zeroth-order theory. The result is

$$\hat{H}_0|\phi_{n2}\rangle + V|\phi_{n1}\rangle = e_{n2}|\phi_n\rangle + e_{n1}|\phi_{n1}\rangle + E_n^{(0)}|\phi_{n2}\rangle.$$

Re-arrange:

$$\left(\hat{H}_0 - E_n^{(0)}\right)|\phi_{n2}\rangle = e_{n2}|\phi_n\rangle + (e_{n1} - V)\,|\phi_{n1}\rangle.$$

Take the scalar product with $|\phi_n\rangle$. The result is

$$0 = e_{n2} + \langle\phi_n|\,(e_{n1} - V)\,|\phi_{n1}\rangle. \qquad (*)$$

Now use the information from the first-order theory. For example,

$$\begin{aligned}
\langle\phi_n|e_{n1}|\phi_{n1}\rangle &= e_{n1}\langle\phi_n|\phi_{n1}\rangle, \\
&= e_{n1}\langle\phi_n|\left(\sum_{n\neq p}\frac{\langle\phi_p|V|\phi_n\rangle}{E_n^{(0)} - E_p^{(0)}}|\phi_p\rangle\right), \\
&= 0.
\end{aligned}$$

Hence, Eq. (*) becomes

$$e_{n2} = \langle\phi_n|V|\phi_{n1}\rangle.$$

We use the first-order theory again:

$$\begin{aligned}
e_{n2} &= \langle\phi_n|V\left(\sum_{n\neq p}\frac{\langle\phi_p|V|\phi_n\rangle}{E_n^{(0)} - E_p^{(0)}}|\phi_p\rangle\right), \\
&= \sum_{n\neq p}\frac{\langle\phi_p|V|\phi_n\rangle}{E_n^{(0)} - E_p^{(0)}}\langle\phi_n|V|\phi_p\rangle, \\
&= \sum_{n\neq p}\frac{|\langle\phi_p|V|\phi_n\rangle|^2}{E_n^{(0)} - E_p^{(0)}}.
\end{aligned}$$

Thus, to **second order** in perturbation theory,

$$E_n = E_n^{(0)} + \lambda\langle\phi_n|V|\phi_n\rangle + \lambda^2\sum_{n\neq p}\frac{|\langle\phi_p|V|\phi_n\rangle|^2}{E_n^{(0)} - E_p^{(0)}} + O(\lambda^3).$$

Note: We require the corrections to the energy to be small, since λ is a small parameter. Typically, the first-order correction to the energy is small. For the first-order correction to the wavefunction

to be small, we require that

$$|\lambda\langle\phi_p|V|\phi_n\rangle| \ll |E_n^{(0)} - E_p^{(0)}|, \qquad \text{for all } n \neq p.$$

If this is not the case, then the perturbation theory breaks down.

20.3 Example of nondegenerate perturbation theory

The Hamiltonian for a harmonic oscillator at frequency ω is the following:

$$\hat{H}_0 = -\frac{\hbar^2}{2m}\partial_x^2 = \tfrac{1}{2}m\omega^2 x^2.$$

Consider instead a particle that experiences an anharmonic potential, such that its Hamiltonian is shifted to a new form:

$$\hat{H} = \hat{H}_0 + qx^4.$$

Identify a dimensionless parameter λ for the problem and compute the anharmonic correction to the ground-state energy assuming this parameter is small.

Solution: We have the eigenvalue problem

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + \tfrac{1}{2}m\omega^2 x^2\psi + qx^4\psi = E\psi.$$

Multiply up by $2m/\hbar^2$:

$$-\frac{\partial^2\psi}{\partial x^2} + \frac{m^2\omega^2}{\hbar^2}x^2 + \frac{2mq}{\hbar^2}x^4\psi = \left(2mE/\hbar^2\right)\psi.$$

Each term now has dimensions of $[\text{Length}]^{-2}\,[\psi]$. Focus on the second term. We have,

$$\frac{1}{[\text{Length}]^2} = \left[\frac{m^2\omega^2}{\hbar^2}\right][\text{Length}]^2.$$

We identify a length scale a:

$$\frac{1}{a^2} = \frac{m^2\omega^2}{\hbar^2}a^2,$$

or

$$a = \sqrt{\frac{\hbar}{m\omega}}.$$

We re-write the oscillator equation as

$$-\frac{\partial^2\psi}{\partial x^2} + \frac{x^2}{a^4}\psi + \frac{2mq}{\hbar^2}x^4\psi = \left(2mE/\hbar^2\right)\psi,$$

or

$$-\frac{\partial^2\psi}{\partial x^2}+\frac{x^2}{a^4}\psi+\frac{2mqa^6}{\hbar^2}\frac{x^4}{a^6}\psi=\left(2mE/\hbar^2\right)\psi,$$

Hence, we identify

$$\lambda=\frac{2mqa^6}{\hbar^2}=\frac{2mq}{\hbar^2}\frac{\hbar^3}{m^3\omega^3}=\frac{2q\hbar}{m^2\omega^3}.$$

In other words,

$$q=\lambda\left(\frac{m^2\omega^3}{2\hbar}\right),$$

and the perturbed problem to solve is

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2}+\tfrac{1}{2}m\omega^2x^2\psi+\lambda\left(\frac{m^2\omega^3}{2\hbar}\right)x^4\psi=E\psi.$$

We therefore identify

$$V:=\left(\frac{m^2\omega^3}{2\hbar}\right)x^4.$$

It is easy to check that this has dimensions of energy:

$$\left[\frac{m^2\omega^3}{2\hbar}x^4\right]=\frac{M^2T^{-3}L^4}{ML^2T^{-1}}=ML^2T^{-2}=[\text{Energy}].$$

Thus, perturbation theory is valid provided the parameter λ is small:

$$\lambda:=\frac{2q\hbar}{m^2\omega^3}\ll 1.$$

The lowest-order correction to the ground-state energy of the oscillator is given by

$$E_0=\tfrac{1}{2}\hbar\omega+\lambda\Delta E,$$

where

$$\Delta E=\langle\psi_0|V|\psi_0\rangle,$$

and where $|\psi_0\rangle$ is the ground state of the associated **harmonic oscillator**. In the position representation,

$$\psi_0(x)=\sqrt{\frac{1}{\sqrt{\pi}a}}e^{-x^2/2a^2},\qquad a=\sqrt{\frac{\hbar}{m\omega}}.$$

Hence,

$$\begin{aligned}\lambda\Delta E &= \lambda\frac{m^2\omega^3}{2\hbar}\frac{1}{\sqrt{\pi}a}\int_{-\infty}^{\infty}x^4e^{-x^2/a^2}\mathrm{d}x,\\ &= \frac{m^2\omega^3}{2\hbar}\frac{1}{\sqrt{\pi}}a^4\int_{-\infty}^{\infty}s^4e^{-s^2}\mathrm{d}s,\\ &= \frac{2q\hbar}{m^2\omega^3}\frac{m^2\omega^3}{2\hbar}a^4\left(\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}s^4e^{-s^2}\mathrm{d}s\right),\\ &= qa^4I,\end{aligned}$$

where I is just a pure number which we determine now. Consider

$$J(\gamma)=\int_{-\infty}^{\infty}e^{-\gamma s^2}\mathrm{d}s=\sqrt{\pi/\gamma}.$$

Hence,

$$\frac{dJ}{d\gamma}=-\tfrac{1}{2}\sqrt{\pi}\gamma^{-3/2}=-\int_{-\infty}^{\infty}s^2e^{-\gamma s^2}\mathrm{d}s.$$

Similarly,

$$\frac{d^2J}{d\gamma^2}=\tfrac{3}{4}\sqrt{\pi}\gamma^{-5/2}=\int_{-\infty}^{\infty}s^4e^{-\gamma s^2}\mathrm{d}s.$$

Setting $\gamma=1$ here gives

$$\tfrac{3}{4}\sqrt{\pi}=\int_{-\infty}^{\infty}s^4e^{-\gamma s^2}\mathrm{d}s,$$

hence, the integral I has the value $3/4$, and

$$E_0=\tfrac{1}{2}\hbar\omega+\tfrac{3}{4}qa^4+O(\lambda^2).$$

Note: We have been very careful here in specifying a dimensionless parameter λ and in constraining it to be small before doing any calculations. Technically, this is essential. However, in practical applications, we simply go to the last step; then, we would solve this problem simply by writing down the relation

$$E_0=\hbar\omega+\langle\psi_0|qx^4|\psi_0\rangle+\cdots.$$

This is what we will do from now on.

The second order

For mischief[1], we go to second order in the perturbation theory, wherein the next correction to the ground-state energy is the following:

$$E_0^{(2)} = q^2 \sum_{p=1}^{\infty} \frac{|\langle \phi_p | x^4 | \phi_0 \rangle|^2}{-\hbar\omega p}.$$

We consider the following integral:

$$\begin{aligned} I_p &= \int_{-\infty}^{\infty} \phi_p x^4 \phi_0 \, \mathrm{d}x, \qquad p \neq 0, \\ &= N_p N_0 \int_{-\infty}^{\infty} H_p(x/a) \mathrm{e}^{-x^2/a^2} x^4 \, \mathrm{d}x, \\ &= N_p N_0 \int_{-\infty}^{\infty} H_p(s) \mathrm{e}^{-s^2} s^4 \, \mathrm{d}s. \end{aligned}$$

But consider

$$\begin{aligned} H_4(s) &= 16s^4 - 48s^2 + 12, \\ H_2(s) &= 4s^2 - 2, \\ H_0(s) &= 1. \end{aligned}$$

Thus,

$$s^4 = \tfrac{1}{16} H_4(s) + \tfrac{3}{4} H_2(s) + \tfrac{3}{4} H_0(s).$$

Thus,

$$\begin{aligned} I_p &= N_p N_0 a^5 \int_{-\infty}^{\infty} H_p(s) \left[\tfrac{1}{16} H_4(s) + \tfrac{3}{4} H_s(s) + \tfrac{3}{4} H_0(s) \right] \mathrm{d}s, \\ &= N_p N_0 a^5 \left(\tfrac{1}{16} 4! 2^4 \sqrt{\pi} \delta_{p,4} + \tfrac{3}{4} 2! 2^2 \sqrt{\pi} \delta_{p,2} + 0 \right). \end{aligned}$$

[1] Going to high order in perturbation theory is sometimes fruitless as well as mischievous. The reason is because it is not known *a priori* what is the radius of convergence of the power-series expansions. A strange heuristic is the following: given a complex-valued function $f(z)$ analytic on a disc D of radius R, it is sometimes possible to approximate $f(z)$ by a truncated Taylor series *even outside of the disc* D. The approximation becomes poorer as more terms are added to the (divergent) series. Therefore, outside the radius of convergence of the perturbation theory, a low-order expansion can give some information about the energy spectrum, while a higher-order expansion gives less information.

Now,

$$\begin{aligned} E_0^{(2)} &= q^2 \sum_{p=1}^{\infty} \frac{I_p^2}{-\hbar\omega p}, \\ &= q^2 \sum_{p=1}^{\infty} (-\hbar\omega p)^{-1} \left[N_p N_0 a^5 \left(\tfrac{1}{16} 4! 2^4 \sqrt{\pi} \delta_{p,4} + \tfrac{3}{4} 2! 2^2 \sqrt{\pi} \delta_{p,2} \right) \right]^2, \end{aligned}$$

and the only terms that survive in the sum are at $p = 4$ and $p = 2$, for which we have

$$\begin{aligned} N_4 N_0 a^5 \left(\tfrac{1}{16} 4! 2^4 \sqrt{\pi} \right) &= a^5 \left(\frac{1}{\sqrt{4! 2^4}} \frac{1}{\pi^{1/4}} \frac{1}{a^{1/2}} \right) \left(\frac{1}{\pi^{1/4}} \frac{1}{a^{1/2}} \right) \left(\tfrac{1}{16} 4! 2^4 \sqrt{\pi} \right), \\ &= \sqrt{\frac{4!}{2^4}} a^4, \end{aligned}$$

and

$$\begin{aligned} N_2 N_0 a^5 \left(\tfrac{3}{4} 2! 2^2 \sqrt{\pi} \right) &= a^5 \left(\frac{1}{\sqrt{2! 2^2}} \frac{1}{\pi^{1/4}} \frac{1}{a^{1/2}} \right) \left(\frac{1}{\pi^{1/4}} \frac{1}{a^{1/2}} \right) \left(\tfrac{3}{2} 2! 2^2 \sqrt{\pi} \right), \\ &= 3 \sqrt{\frac{2!}{2^2}} a^4. \end{aligned}$$

Combine:

$$\begin{aligned} \frac{1}{2\hbar\omega} I_2^2 + \frac{1}{4\hbar\omega} I_4^2 &= \frac{1}{2\hbar\omega} \left(\frac{4!}{2^4} a^8 \right) + \frac{1}{4\hbar\omega} \left(\frac{9 \times 2!}{2^2} a^8 \right), \\ &= -\tfrac{21}{8} \frac{a^8}{\hbar\omega}. \end{aligned}$$

Hence,

$$E_0 = \hbar\omega + \tfrac{3}{4} q a^4 - \tfrac{21}{8} \frac{a^8}{\hbar\omega} + O(q^3).$$

Chapter 21

Time-independent perturbation theory: degenerate case

Reading material for this chapter: Mandl, Chapter 7

21.1 Overview

In this chapter we continue with the time-independent perturbation theory, this time for cases where the eigenvalues of energy are degenerate. The problem to solve is therefore modified from that in Ch. 20: We start with the exactly-solvable problem

$$\hat{H}_0|\phi\rangle = E^{(0)}|\phi\rangle$$

with a discrete spectrum

$$E = E_1^{(0)}, E_2^{(0)}, \cdots,$$

The n^{th} energy level is assumed to be s-**fold degenerate**, with eigenvectors

$$|u_{n1}\rangle, \cdots, |u_{ns}\rangle,$$

such that

$$\langle u_{n\alpha}|u_{n\beta}\rangle = \delta_{\alpha\beta}, \qquad \alpha, \beta = 1, \ldots, s.$$

It is required to compute the changes to the n^{th} energy level due to the presence of a perturbation,

$$\hat{H}_0 \to \hat{H} := \hat{H}_0 + \lambda V,$$

where λ is a small dimensionless parameter.

It is not obvious *a priori* that the degeneracy will remain in place once the perturbation is added to the problem. Thus, we assume that the energy level $E_n^{(0)}$ splits into s new levels:

$$E_{ni} = E_n^{(0)} + \lambda E_{ni}^{(1)} + \lambda^2 E_{ni}^{(0)} + \cdots, \qquad i = 1, \cdots, s.$$

Associated with each new energy level, there is an eigenvector:

$$|\psi_{ni}\rangle = |\phi_{ni}\rangle + \lambda|\phi_{ni}^{(1)}\rangle + \lambda^2|\phi_{ni}^{(2)}\rangle + \cdots,$$

where the states on the right-hand side are to be determined.

21.2 The solution

We focus on finding two quantities:

- The first-order correction to the energy, $E_{ni} = E_n^{(0)} + \lambda E_{ni}^{(1)}$;
- The zeroth-order perturbed state: $|\psi_{ni}\rangle = |\phi_{ni}\rangle + O(\lambda)$.

As before, we focus first of all on the zeroth-order expansion in the problem

$$\left(\hat{H}_0 + \lambda\right)\left[|\phi_{ni}\rangle + \lambda|\phi_{ni}^{(1)}\rangle\right] = \left(E_n^{(0)} + \lambda E_{ni}^{(1)}\right)\left[|\phi_{ni}\rangle + \lambda|\phi_{ni}^{(1)}\rangle\right],$$

or

$$\hat{H}_0|\phi_{ni}\rangle = E_n^{(0)}|\phi_{ni}\rangle.$$

Now we go over to the first-order term:

$$\hat{H}_0|\phi_{ni}^{(1)}\rangle + V|\phi_{ni}\rangle = E_n^{(0)}|\phi_{ni}^{(1)}\rangle + E_{ni}^{(1)}|\phi_{ni}\rangle.$$

Re-arranging gives

$$\left(\hat{H}_0 - E_n^{(0)}\right)|\phi_{ni}^{(1)}\rangle = \left(E_{ni}^{(1)} - V\right)|\phi_{ni}\rangle.$$

This is similar to the result in the non-degenerate case. However, one key difference is that now we do NOT know what the state $|\phi_{ni}\rangle$ is. We now determine it, and hence determine the first-order corrections to the energy. Certainly, the state $|\phi_{ni}\rangle$ is a mixture of the eigenstates of the unperturbed problem:

$$|\phi_{ni}\rangle = \sum_{\alpha=1}^{s} C_{i\alpha}|u_{n\alpha}\rangle.$$

Thus, it suffices to determine the $C_{i\alpha}$'s. We go back to the first-order equation:

$$\left(\hat{H}_0 - E_n^{(0)}\right)|\phi_{ni}^{(1)}\rangle = \left(E_{ni}^{(1)} - V\right)|\phi_{ni}\rangle,$$

or

$$\left(\hat{H}_0 - E_n^{(0)}\right)|\phi_{ni}^{(1)}\rangle = \left(E_{ni}^{(1)} - V\right)\left(\sum_{\alpha=1}^{s} C_{i\alpha}|u_{n\alpha}\rangle\right).$$

We take the scalar product of both sides with the bra $\langle u_{n\beta}|$. Certainly $\langle u_{n\beta}|\left(\hat{H}_0 - E_n^{(0)}\right) = 0$, hence

$$\begin{aligned}
0 &= \langle u_{n\beta}|\left(E_{ni}^{(1)} - V\right)\left(\sum_{\alpha=1}^{s} C_{i\alpha}|u_{n\alpha}\rangle\right), \\
&= \left(\sum_{\alpha=1}^{s} C_{i\alpha}\langle u_{n\beta}|\right)\left(E_{ni}^{(1)} - V\right)|u_{n\alpha}\rangle, \\
&= \sum_{\alpha=1}^{s} C_{i\alpha}\left[E_{ni}^{(1)}\delta_{\alpha\beta} - \langle u_{n\beta}|V|u_{n\alpha}\rangle\right], \\
&= \sum_{\alpha=1}^{s} C_{i\alpha}\left[E_{ni}^{(1)}\delta_{\alpha\beta} - V_{\beta\alpha}\right]
\end{aligned}$$

Thus, we have a set of s homogeneous equations:

$$\begin{aligned}
\sum_{\alpha=1}^{s} C_{1\alpha}\left[E_{n1}^{(1)}\delta_{\alpha\beta} - V_{\beta\alpha}\right] &= 0, \\
\vdots &\quad \vdots \\
\sum_{\alpha=1}^{s} C_{s\alpha}\left[E_{ns}^{(1)}\delta_{\alpha\beta} - V_{\beta\alpha}\right] &= 0.
\end{aligned}$$

However, this is identical to s copies of the problem

$$\left[E_{ni}^{(1)}\mathbb{I}_{s\times s} - \boldsymbol{V}\right]\begin{pmatrix} C_{i1} \\ \dots \\ C_{is} \end{pmatrix} = 0,$$

which is an eigenvalue problem in the eigenvalue E_{ni}^{i}. In conclusion,

- The perturbed level-n energies are computed as

$$E_{ni} = E_n^{(0)} + \lambda\Delta E_{ni}, \qquad i = 1, \cdots, s$$

where ΔE_{ni} are the s eigenvalues of the problem

$$[\Delta E_{ni}\mathbb{I}_{s\times s} - \boldsymbol{V}]\,\boldsymbol{C}_i = 0.$$

- The perturbed level-n states are computed as

$$|\psi_{ni}\rangle = |\phi_{ni}\rangle + O(\lambda),$$

where the states $|\phi_{ni}\rangle$ are determined from eigenvectors of the problem:

$$|\phi_{ni}\rangle = \sum_{\alpha=1}^{n} C_{i\alpha}|u_{n\alpha}\rangle.$$

Before looking at an example, consider again the result just derived, namely that the perturbations to the energy levels are eigenvalues of the problem

$$|\boldsymbol{V} - \Delta E_{ni}\mathbb{I}| = 0, \qquad V_{\alpha\beta} = \langle u_{n\alpha}|V|u_{n\beta}\rangle. \qquad (*)$$

Suppose we can find a clever basis $\{|u_{n\alpha}\rangle\}_{\alpha=1}^{s}$ for the Hamiltonian $\hat{H}_0$ that is simultaneously a set of eigenvectors for V. Then the eigenvalue problem (*) is diagonal, with eigenvalues

$$\Delta E_{ni} = \langle u_{ni}|V|u_{ni}\rangle, \qquad i = 1, \cdots, s.$$

This is guaranteed if $(\hat{H}_0, V)$ are compatible:

$$\left[\hat{H}_0, V\right] = 0.$$

For large problems ($s \gg 1$), it is a good idea to find such a clever basis before solving the determinant problem. It will be a good idea to keep this approach in mind when we consider spin-orbit coupling in Ch. 22.

Example: Consider a basic system

$$\hat{H}_0 = \begin{pmatrix} E_0 & 0 \\ 0 & E_0 \end{pmatrix},$$

to which is added a perturbation

$$\hat{H}_0 \to \hat{H}_0 + \lambda \begin{pmatrix} V_0 & V_0 \\ V_0 & 0 \end{pmatrix}.$$

Show (i) that the basic system is degenerate; (ii) that the perturbation brakes the degeneracy. Hence, compute the lowest-order correction to the energy, and write down the perturbed eigenstates.

The basic system is degenerate:

$$\begin{pmatrix} E_0 & 0 \\ 0 & E_0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = E_0 \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$
$$\begin{pmatrix} E_0 & 0 \\ 0 & E_0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = E_0 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Hence, the states

$$|u_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |u_2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

both have the same energy.

From the theory, the corrections ΔE to this energy are determined by the eigenvalue problem

$$|\Delta E \mathbb{I} - \boldsymbol{V}| = 0,$$

where

$$\begin{aligned} V_{11} &= \langle u_1|V|u_1\rangle = V_0, \\ V_{12} &= \langle u_1|V|u_2\rangle = V_0, \\ V_{21} &= \langle u_2|V|u_1\rangle = V_0, \\ V_{22} &= \langle u_2|V|u_2\rangle = 0. \end{aligned}$$

Thus, we solve

$$\begin{vmatrix} V_0 - \Delta E & V_0 \\ V_0 & -\Delta E \end{vmatrix} = 0. \qquad (*)$$

Hence,

$$\Delta E = V_0 \varphi_\pm, \qquad \varphi_\pm = \frac{1 \pm \sqrt{5}}{2}.$$

The perturbation therefore breaks the degeneracy and introduces new energy levels:

$$E_{01} = E_0 + \lambda V_0 \varphi_+,$$
$$E_{02} = E_0 + \lambda V_0 \varphi_-.$$

The corresponding new energy states are given by the eigenvectors of the problem (*). Up to normalisation, these are

$$C_1 = (1, -\varphi_-), \qquad C_2 = (\varphi_-, 1).$$

Thus, the perturbed upper state E_{01} has eigenvector

$$|\psi_{01}\rangle = |u_1\rangle - \varphi_-|u_2\rangle,$$

up to normalisation, whle the perturbed lower state E_{02} has eigenvector

$$|\psi_{02}\rangle = \varphi_-|u_1\rangle + |u_2\rangle,$$

up to normalisation.

Of course, this is a very silly example, because the perturbed system can be solved exactly. It is readily seen that the **exact** solution to the perturbed problem is

$$E_{01} = E_0 + \lambda V_0 \varphi_+,$$
$$E_{02} = E_0 + \lambda V_0 \varphi_-,$$

with eigenvalues

$$|\psi_{01}\rangle = \begin{pmatrix} 1 \\ -\varphi_- \end{pmatrix},$$

and

$$|\psi_{02}\rangle = \begin{pmatrix} \varphi_- \\ 1 \end{pmatrix}$$

(up to normalisation). But these are precisely the lowest-order solutions of the perturbed problem. Thus, we conclude that we have been very lucky, and that the lowest-order degenerate perturbation theory agrees with the exact solution. It is very rare for this to happen.

Chapter 22

The fine structure of hydrogen

Reading material for this chapter: Mandl, Chapter 7; Young and Freedman, Chapters 28–29

22.1 Classical magnetic moments

Consider a particle of charge Q and mass m doing circular motion of radius r (Fig. 22.1). To an observer in the lab frame, the particle carries a current, since

$$\text{Current} = I = \frac{\text{Charge in motion}}{\text{Time}}.$$

The appropriate value of time here is the period of the circular motion:

$$T = \frac{2\pi}{\omega} = \frac{2\pi}{v/r} = \frac{2\pi r}{v}.$$

Thus,

$$I = \frac{Qv}{2\pi r}.$$

We **define** the **magnetic moment** as

$$\mu = \text{magnetic moment} := \text{Current} \times \text{Area},$$

hence

$$\mu = IA = \left(\frac{Qv}{2\pi r}\right)\pi r^2 = \frac{Qvr}{2}.$$

Note, however, that the particle's angular momentum is

$$L = mvr.$$

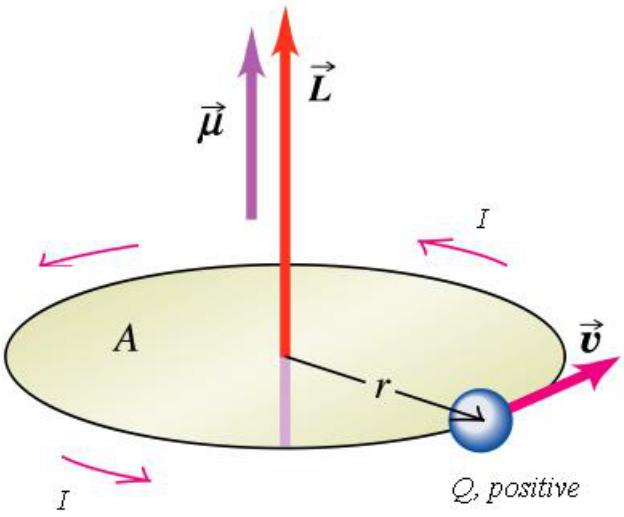


Figure 22.1: Classical magnetic-moment vector of a current loop

Thus, the magnetic moment and the angular momentum are proportional:

$$\mu = \frac{Q}{2m} L.$$

Because angular momentum is a vector, we promote the magnetic moment to vector status:

$$\boldsymbol{\mu} = \frac{Q}{2m} \boldsymbol{L},$$

which is perpendicular to the plane of the motion.

Next, we place the current loop in a uniform magnetic field $\boldsymbol{B}$. In this exercise, we consider instead a square current loop, although the principles are the same. The system is shown schematically in Fig. 22.2. We focus on the highlighted point, and carry out a cross-section in the $z - x$ plane (Fig. 22.3). Here, the current loop consists of a charged particle (charge $\mathrm{d}Q$) moving at velocity v in the $-y$-direction. The particle experiences the Lorentz force $\mathrm{d}\boldsymbol{F} = \mathrm{d}Q\boldsymbol{v} \times \boldsymbol{B}$, which is in the positive x-direction:

$$\mathrm{d}F = \mathrm{d}QvB, \qquad \text{in the positive } x \text{ direction.}$$

Thus, a torque is exerted on the loop, that causes it to rotate. The torque is

$$\mathrm{d}\tau = \mathrm{d}\hat{F} r,$$

where $\mathrm{d}\hat{F}$ is the projection of the force on to a direction perpendicular to the loop axis. In other words,

$$\mathrm{d}\tau = \mathrm{d}F r \cos\alpha,$$

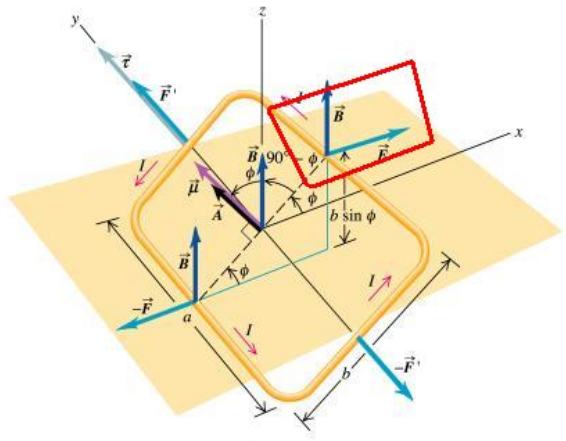


Figure 22.2: Current loop in a magnetic field

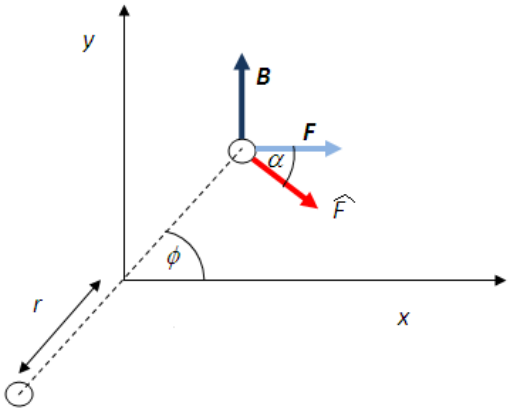


Figure 22.3: Current loop in a magnetic field (zoom in on a point of interest)

or

$$\mathrm{d}\tau = \mathrm{d}Fr\sin\phi.$$

Restoring $\mathrm{d}Q$, this is

$$\mathrm{d}\tau = \mathrm{d}QvBr\sin\phi = \frac{\mathrm{d}Qv}{2}Bb\sin\phi.$$

Recall the definition of current:

$$I = \frac{\mathrm{d}Q}{\mathrm{d}t},$$

hence

$$\mathrm{d}Q = I\mathrm{d}t = \frac{I}{v}\mathrm{d}x.$$

Hence, the increment of torque along the top part of the current loop is

$$\mathrm{d}\tau = \frac{\mathrm{d}Qv}{2}Bb\sin\phi = \tfrac{1}{2}IB\sin\phi\,\mathrm{d}x.$$

Integrating along the top segment of the loop gives $\mathrm{d}x \to a$. However, there is an identical contribution to the **total** torque on the loop coming from the opposite wide. Thus, the total torque on the loop is

$$IBab\sin\phi.$$

But $A = ab$, hence

$$\tau = IAB\sin\phi,$$

or

$$\tau = \mu B\sin\phi.$$

Next, we compute the work done by the magnetic force in rotating the loop through an angular

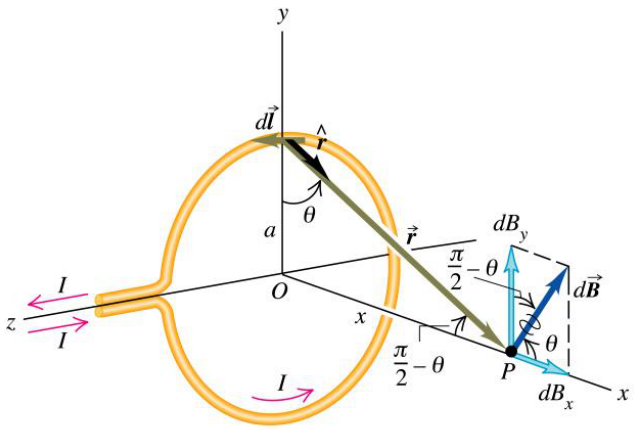


Figure 22.4: As a consequence of Ampère's Law (Maxwell's equations), a current loop generates a magnetic field.

increment $\mathrm{d}\phi$. This is

$$\mathrm{d}W = \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{x} = 2\hat{F} r \mathrm{d}\theta.$$

The factor of 2 comes from the fact that work is done by the force along both lengths of the loop. Moreover, $r = b/2$. Hence,

$$\mathrm{d}W = \tau \mathrm{d}\phi = \mu B \sin\phi \mathrm{d}\phi.$$

Integrating gives

$$W(\phi_2) - W(\phi_1) = -\mu B \cos\phi \Big|_{\phi_1}^{\phi_2} = -\boldsymbol{\mu} \cdot \boldsymbol{B} \Big|_{\phi_1}^{\phi_2},$$

which implies the existence of a **magnetic potential energy**

$$\mathcal{U} = -\boldsymbol{\mu} \cdot \boldsymbol{B}.$$

22.2 Biot–Savart Law

We state without proof the following result: A current loop creates a magnetic field whose sense is given by the right-hand rule; the magnitude of the field at the centre of the loop is

$$B = \frac{\mu_0 I}{2r}$$

where μ_0 is the magnetic constant. This is a simple application of the **Biot–Savart Law**, which in turn is a simple consequence of Maxwell's equations in the static case (See Fig. 22.4).

Consider now a small, charged, 'spinning particle' with finite magnetic moment $\boldsymbol{\mu}$ that sits at the

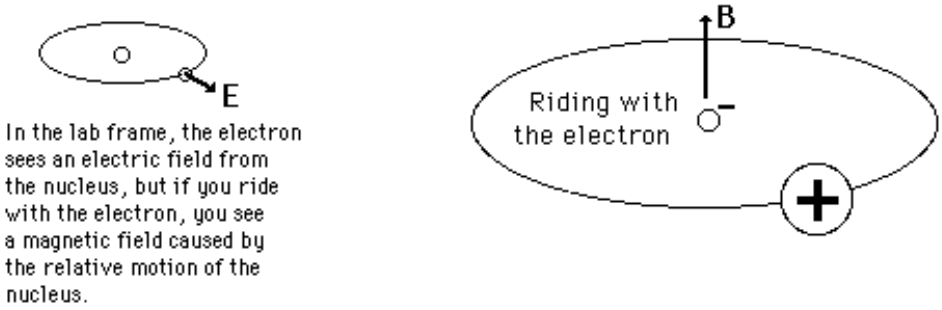


Figure 22.5: The electron bound to a hydrogen atom, viewed in two different frames of reference. Left: lab frame; right: electron's rest frame.

centre of a current loop. The particle sees a magnetic field

$$\boldsymbol{B} = \frac{\mu_0 I}{2r}\hat{\boldsymbol{z}},$$

where $\hat{\boldsymbol{z}}$ is a unit vector perpendicular to the plane of the loop. The particle therefore experiences a potential

$$\mathcal{U}_{SO} = -\boldsymbol{B}\cdot\boldsymbol{\mu}.$$

We apply these ideas to the electron in a hydrogen atom.

22.3 Spin-orbit coupling in the hydrogen atom

In the lab frame, the electron 'sees' an electric field from the nucleus. However, if we go over to the electron's rest frame, it sees a current loop formed by the now-orbiting positive nucleus (Fig. 22.5). Thus, in the frame of reference of the electron, there is a magnetic field

$$B = \frac{\mu_0 I}{2r}, \qquad I = \frac{e}{T} = \frac{e}{2\pi}\omega = \frac{e}{2\pi r}v,$$

hence

$$B = \frac{\mu_0}{4\pi}\frac{ev}{r^2}.$$

The sense of this field is given by

$$\boldsymbol{B} = \frac{\mu_0}{4\pi}\frac{e}{r^3}\boldsymbol{r}\times\boldsymbol{v},$$

$$\boldsymbol{B} = \frac{\mu_0}{4\pi m_{\mathrm{e}}}\frac{e}{r^3}\boldsymbol{L},$$

where $\boldsymbol{L}$ is the angular momentum of the electron **as measured in the lab frame**. Staying in the electron's rest frame, we remind ourselves that it has a finite magnetic moment:

$$\boldsymbol{\mu} = -\frac{g}{2m_{\mathrm{e}}}\boldsymbol{S}, \qquad g \approx 2,$$

and thus, there is a **spin–orbit** interaction potential:

$$\mathcal{U}_{SO} = -\boldsymbol{\mu}\cdot\boldsymbol{B} = +\frac{\mu_0 e^2}{4\pi m_{\mathrm{e}}^2}\frac{1}{r^3}\boldsymbol{L}\cdot\boldsymbol{S}.$$

But

$$\mu_0\epsilon_0 = c^{-2},$$

and this expression can be tidied up:

$$\mathcal{U}_{SO} = -\boldsymbol{\mu}\cdot\boldsymbol{B} = +\frac{e^2}{4\pi\epsilon_0 r^2}\frac{1}{m_{\mathrm{e}}c^2 r}\boldsymbol{L}\cdot\boldsymbol{S},$$

or or

$$\mathcal{U}_{SO} = \frac{1}{m_{\mathrm{e}}c^2}\frac{1}{r}\frac{d\mathcal{U}}{dr}\boldsymbol{L}\cdot\boldsymbol{S}.$$

Unfortunately, this is wrong by a factor of two. If we carry out the calculation in a relativistically correct fashion, we obtain the result

$$\boxed{\mathcal{U}_{SO} = \frac{1}{2m_{\mathrm{e}}c^2}\frac{1}{r}\frac{d\mathcal{U}}{dr}\boldsymbol{L}\cdot\boldsymbol{S}.}$$

Note that this result, derived in the electron's frame of reference, is exactly the same in the laboratory frame. We return to this frame and compute the effects of this **spin-orbit coupling** on the energy levels of hydrogen.

We consider the following perturbed Hamiltonian for the hydrogen atom:

$$\hat{H} = \underbrace{\left(-\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\epsilon_0 r}\right)}_{=\hat{H}_0} + = \frac{1}{2m_{\mathrm{e}}c^2}\frac{1}{r}\frac{d\mathcal{U}}{dr}\boldsymbol{L}\cdot\boldsymbol{S}$$

(we suppress the hats on the angular momentum operators). The eigenvalues $E_{n\ell}^{(0)}$ of $\hat{H}_0$ are $2(2\ell+1)$-fold degenerate with respect to the orbital angular momentum (quantum number ℓ). Treating the spin-orbit interaction as small, we must use degenerate perturbation theory. Note that the unperturbed eigenfunctions

$$R_n(r)Y_{\ell,m_\ell}(\theta,\varphi)|\pm\rangle, \qquad (*)$$

do not diagonalise the perturbation $\mathcal{U}_{SO}$ because this contains a mixture of angular momentum projections along various axes. However, if we re-write the perturbation as

$$\mathcal{U}_{SO} = \frac{1}{4m_{\rm e}c^2}\frac{1}{r}\frac{d\mathcal{U}}{dr}\left(\boldsymbol{J}^2 - \boldsymbol{L}^2 - \boldsymbol{S}^2\right),$$

where $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$ is the addition of the spin and orbital angular momenta, then the functions

$$|n, \ell, s, J, M\rangle$$

do diagonalise the perturbation. Here $|n, \ell, s, J, M\rangle$ is an eigenstate of the CSCO $\{\boldsymbol{L}^2, \boldsymbol{S}^2, \boldsymbol{J}^2, J_z\}$ got by a linear combination of the functions (*), and by the angular-momentum addition theorem.

Thus, the 'clever eigenstates' $|u_{ni}\rangle$ in the theoretical presentation of degenerate perturbation theory that diagonalise both $\hat{H}_0$ and V are in fact the states $|n, \ell, s, J, M\rangle$. The theoretical formula for the corrections to the energy levels was

$$\Delta E^{(1)}_{ni} = \lambda\langle u_{ni}|V|u_{ni}\rangle.$$

Letting $|u_{ni}\rangle \to |n, \ell, s, J, M\rangle$, this is

$$\Delta E(n, \ell, J) = \langle n, \ell, s, J, M|\mathcal{U}_{SO}|n, \ell, s, J, M\rangle,$$

or

$$\begin{aligned}\Delta E(n, \ell, J) &= \frac{1}{4m_{\rm e}^2c^2}\langle \ell, s, J, M|\left(\boldsymbol{J}^2 - \boldsymbol{L}^2 - \boldsymbol{S}^2\right)|\ell, s, J, M\rangle\left\langle\frac{1}{r}\frac{dU}{dr}\right\rangle_{n\ell}, \\ &= \frac{\hbar^2}{4m^2c^2}\left[j(j+1) - \ell(\ell+1) - \tfrac{3}{4}\right]\left\langle\frac{1}{r}\frac{dU}{dr}\right\rangle_{n\ell},\end{aligned}$$

where

$$\left\langle\frac{1}{r}\frac{dU}{dr}\right\rangle_{n\ell}$$

denotes the expectation value of $r^{-1}U'(r)$ with respect to the function $R_{n\ell}Y_{\ell,m_\ell}(\theta, \varphi)$. This value is independent of m_ℓ because the operator $r^{-1}U'(r)$ is independent of φ. Carrying out this integral (homework), we have

$$\left\langle\frac{1}{r}\frac{dU}{dr}\right\rangle_{n\ell} = \frac{1}{a_0^3n^3\ell(\ell+1)(\ell+\frac{1}{2})},$$

hence

$$\Delta E(n\ell j) = \frac{1}{4m_{\rm e}^2c^2}\frac{\hbar^2e^2}{4\pi\epsilon_0a^3n^3}\frac{j(j+1) - \ell(\ell+1) - \frac{3}{4}}{\ell(\ell+1)(\ell+\frac{1}{2})},$$

or

$$\begin{aligned}\Delta E(n\ell j) &= \frac{|E_n|}{n}\alpha^2\frac{j(j+1)-\ell(\ell+1)-\frac{3}{4}}{\ell(\ell+1)(\ell+\frac{1}{2})},\\ \alpha^2 &= \frac{e^2}{4\pi\epsilon_0\hbar c},\\ E_n &= -\tfrac{1}{2}\frac{e^2}{4\pi\epsilon_0 n^2}.\end{aligned}$$

Note that there is no correction to the energy of s-states ($\ell = 0$), since then $j = s = 1/2$, and the numerator is identically zero. In reality, there is a second $O(\alpha^2)$ effect, due to relativistic effects, wherein the dependence of the electron mass on its velocity is considered. This consideration leads to a shift in the s-states.

That the energy levels are shifted is called **splitting**. The splitting is very difficult to see without precise equipment. Thus, the spin-orbit features of the hydrogen atom are called **fine structure** (Fig. 22.6). The small perturbation parameter α is called the **fine-structure constant**.

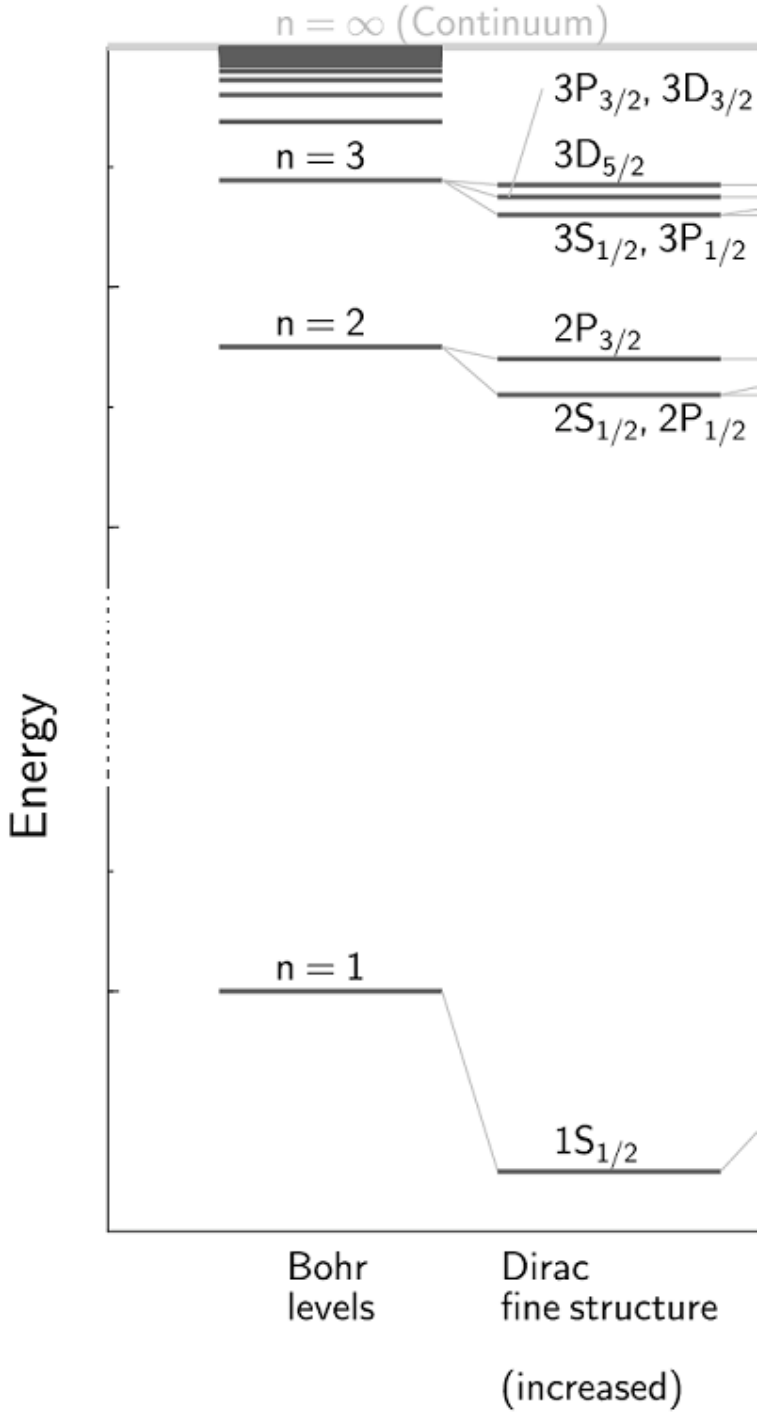


Figure 22.6: Hydrogenic fine structure. Schematic shows effects of spin-orbit coupling and relativistic-mass effect. Diagram uses **spectroscopic notation** – the letter is for the total orbital angular momentum and the letter with the subscript is for the total (spin+orbital) angular momentum.

Chapter 23

Variational methods

Reading material for this chapter: Mandl, Chapter 8

23.1 Estimating the ground state of an arbitrary system

In this chapter we develop a neat trick to estimate the ground state of a fairly general system. It is based on simple integrations and avoids the messy sums involved in perturbation theory.

23.2 The idea

Consider a system described by a Hamiltonian $\hat{H}$, which possesses the complete set of orthonormal eigenstates $|u_1\rangle, |u_2\rangle, \cdots$, which are unknown. We write down the corresponding energy levels in an ordered sequence:

$$E_1 \leq E_2 \leq \cdots .$$

Any state $|\psi\rangle$ of the system can be expanded in terms of a sum of these eigenvectors:

$$|\psi\rangle = \sum_{n=1}^{\infty} c_n |u_n\rangle.$$

Hence,

$$\frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{\sum_{n=1}^{\infty} |c_n|^2 E_n}{\sum_{n=1}^{\infty} |c_n|^2}.$$

Now the sequence of energy levels is ordered, hence

$$\frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{\sum_{n=1}^{\infty} |c_n|^2 E_n}{\sum_{n=1}^{\infty} |c_n|^2} \geq \frac{\sum_{n=1}^{\infty} |c_n|^2 E_1}{\sum_{n=1}^{\infty} |c_n|^2} = E_1,$$

hence

$$E_1 \leq \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle},$$

for **any** state $|\psi\rangle$ in the Hilbert space of solutions.

Thus, to estimate the ground state energy of the system, we write down a wavefunction $\psi(\alpha_1, \cdots, \alpha_s)$, which possesses the qualitative features of the correct but unknown ground-state energy, and which contains several free parameters $\alpha_1, \cdots, \alpha_s$. Then,

$$E_1 \leq E(\alpha_1, \cdots, \alpha_s) := \frac{\langle\psi(\alpha_1, \cdots, \alpha_s)|\hat{H}|\psi(\alpha_1, \cdots, \alpha_s)\rangle}{\langle\psi(\alpha_1, \cdots, \alpha_s)|\psi(\alpha_1, \cdots, \alpha_s)\rangle},$$

for all values of the parameters $\alpha_1, \cdots, \alpha_s$. By minimising over the parameters $\alpha_1, \cdots, \alpha_s$, the **upper bound** for the ground-state energy can be sharpened:

$$E_1 \leq \min_{\alpha_1, \cdots, \alpha_s} E(\alpha_1, \cdots, \alpha_s).$$

This procedure is called the **variational technique**.

23.3 The Yukawa potential

Estimate the ground-state energy of a particle experiencing the attractive central potential

$$\mathcal{U}(r) = -g^2 \, \frac{e^{-Mr}}{r},$$

where g^2 and M are positive numbers.

Let's write down the eigenvalue problem for the potential:

$$-\frac{\hbar^2}{2m}\nabla^2\psi - \frac{g^2}{r}e^{-Mr}\psi = E\psi.$$

Multiply up by $2m/\hbar^2$:

$$-\nabla^2\psi - \frac{2mg^2}{\hbar^2}\frac{e^{-Mr}}{r}\psi = (2mE/\hbar^2)\psi = -k^2\psi.$$

Identify

$$a := \hbar^2/mg^2.$$

Thus, the eigenvalue problem is

$$-\nabla^2\psi - \frac{2}{ar}e^{-Mr}\psi = -k^2\psi.$$

Formally, if $M = 0$, we recover the hydrogen atom. This suggests that the trial function should look like the hydrogenic ground state, which is

$$\frac{1}{\sqrt{\pi}a^{3/2}}e^{-r/a}.$$

However, we still need to take care of the exponential term. This term damps the potential to zero very rapidly, suggesting wavefunctions that are localised very close to the force centre. Thus, we propose a trial wavefunction that decays to zero more rapidly than the hydrogenic one:

$$\psi(\boldsymbol{r}) = \frac{\alpha^{3/2}}{\sqrt{\pi}a^{3/2}}e^{-\alpha r/a}.$$

We now compute the expectation values, in detail.

First, the **kinetic energy**. We have

$$\begin{aligned}
\nabla^2 e^{-\alpha r/a} &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}e^{-\alpha r/a}\right), \\
&= \frac{\partial^2}{\partial r^2}e^{-\alpha r/a} + \frac{2}{r}\frac{\partial}{\partial r}e^{-\alpha r/a}, \\
&= \left(\frac{\alpha^2}{a^2} - \frac{2\alpha}{ra}\right)e^{-\alpha r/a}.
\end{aligned}$$

In addition,

$$\begin{aligned}
\int d^3 r e^{-2\alpha r/a} &= \int_0^\infty r^2 \mathrm{d}r \int_\Omega \mathrm{d}\Omega\, e^{-2\alpha r/a}, \\
&= 4\pi\int_0^\infty r^2 e^{-2\alpha r/a}\mathrm{d}r, \\
&= 4\pi\left(\frac{a}{2\alpha}\right)^3\int_0^\infty u^2 e^{-u}\mathrm{d}\alpha, \\
&= 4\pi\left(\frac{a}{2\alpha}\right)^3 2!, \\
&= \pi\frac{a^3}{\alpha^3},
\end{aligned}$$

as well as

$$\begin{aligned}\int d^3r \frac{1}{r} e^{-2\alpha r/a} &= \int_0^\infty r\mathrm{d}r \int_\Omega \mathrm{d}\Omega\, e^{-2\alpha r/a}, \\ &= 4\pi \int_0^\infty r e^{-2\alpha r/a}\mathrm{d}r, \\ &= 4\pi \left(\frac{a}{2\alpha}\right)^2 \int_0^\infty u^2 e^{-u}\mathrm{d}\alpha, \\ &= 4\pi \left(\frac{a}{2\alpha}\right)^2 1!, \\ &= \pi \frac{a^2}{\alpha^2}.\end{aligned}$$

Hence,

$$\begin{aligned}\int \mathrm{d}^3r\, e^{-\alpha r/a}\nabla^2 e^{-\alpha r/a} &= \int \mathrm{d}^3r\, e^{-\alpha r/a}\left(\frac{\alpha^2}{a^2} - \frac{2\alpha}{ra}\right) e^{-\alpha r/a}, \\ &= \frac{\alpha^2}{a^2}\int \mathrm{d}^3r\, e^{-2\alpha r/a} - \frac{2\alpha}{a}\int \mathrm{d}^3r\, \frac{1}{r} e^{-2\alpha r/a}, \\ &= \frac{\alpha^2}{a^2}\left(\pi \frac{a^3}{\alpha^3}\right) - \frac{2\alpha}{a}\left(\pi \frac{a^2}{\alpha^2}\right), \\ &= -\frac{\pi a}{\alpha}.\end{aligned}$$

Thus, the expected value of the kinetic energy in this state is

$$\int \mathrm{d}d^3r\, \frac{\alpha^{3/2}}{\sqrt{\pi}a^{3/2}} e^{-r/a}\left(-\frac{\hbar^2}{2m}\nabla^2\right)\frac{\alpha^{3/2}}{\sqrt{\pi}a^{3/2}} e^{-r/a} = -\frac{\hbar^2}{2m}\frac{\alpha^3}{\pi a^3}\int \mathrm{d}^3r\, e^{-\alpha r/a}\nabla^2 e^{-\alpha r/a} = \frac{\hbar^2}{2m}\frac{\alpha^3}{\pi a^3}\frac{\pi a}{\alpha} = \frac{\hbar^2\alpha^2}{2ma^2}.$$

Next, we compute

$$\begin{aligned}\int \mathrm{d}^3 e^{-\alpha r/a}\frac{e^{-Mr}}{r} e^{-\alpha r/a} &= \int \mathrm{d}^3 \frac{e^{-Mr}}{r} e^{-2\alpha r/a}, \\ &= \int_0^\infty r^2\mathrm{d}r \int_\Omega \mathrm{d}\Omega\, \frac{e^{-r(M+2\alpha/a)}}{r}, \\ &= 4\pi \int_0^\infty \mathrm{d}r\, r e^{-r(M+2\alpha/a)}, \\ &= \frac{4\pi}{(M+2\alpha/a)^2}\int_0^\infty \mathrm{d}u\, u e^{-u}, \\ &= \frac{4\pi}{(M+2\alpha/a)^2}.\end{aligned}$$

and the expected value of the potential energy is therefore

$$\int \mathrm{d}^3 r \, \frac{\alpha^{3/2}}{\sqrt{\pi} a^{3/2}} e^{-r/a} \left(-g^2 \frac{e^{-Mr}}{r} \right) \frac{\alpha^{3/2}}{\sqrt{\pi} a^{3/2}} e^{-r/a} = \\ - g^2 \frac{\alpha^3}{\pi a^3} \int \mathrm{d}^3 r \, e^{-2r/a} \frac{e^{-Mr}}{r} = -g^2 \frac{\alpha^3}{\pi a^3} \frac{4\pi}{(M + 2\alpha/a)^2}.$$

Putting it all together, we have

$$\begin{aligned} E(\alpha) &= \frac{\hbar^2 \alpha^2}{2ma^2} - \frac{4\alpha^3}{a^3} \frac{g^2}{(M + 2\alpha/a)^2}, \\ &= \frac{\hbar^2 \alpha^2}{2ma^2} - \frac{4\alpha^3}{a^3} \frac{g^2}{\frac{4\alpha^2}{a^2} \left(1 + \frac{Ma}{2\alpha}\right)^2}, \\ &= \frac{\hbar^2 \alpha^2}{2ma^2} - \frac{\alpha}{a} \frac{g^2}{\left(1 + \frac{Ma}{2\alpha}\right)^2}, \\ &= \frac{\hbar^2 \alpha^2}{2ma^2} - \frac{\alpha}{a} \frac{\hbar^2}{ma} \left(1 + \frac{Ma}{2\alpha}\right)^{-2}, \qquad g^2 = \hbar^2/ma, \\ &= \frac{\hbar^2 \alpha^2}{2ma^2} - \frac{\hbar^2 \alpha}{ma^2} \left(1 + \frac{Ma}{2\alpha}\right)^{-2}. \end{aligned}$$

Before continuing, it is salutary to check that by setting $M = 0$ and $\alpha = 1$, we recover the functional form for the state of hydrogen:

$$E(\alpha = 1, M = 0) = \frac{\hbar^2}{2ma^2} - \frac{\hbar^2}{ma^2} = -\frac{\hbar^2}{2ma^2}.$$

This is indeed the case, since the ground state of hydrogen is

$$1\mathrm{Ry} = 13.6\mathrm{eV} = -\frac{\hbar^2}{2m_\mathrm{e} a_0^2}, \qquad a_0 := \text{Bohr radius.}$$

Thus, our estimate for the ground state of the Yukawa potential is

$$E(\alpha) = \frac{\hbar^2 \alpha^2}{2ma^2} - \frac{\hbar^2 \alpha}{ma^2} \left(1 + \frac{Ma}{2\alpha}\right)^{-2}.$$

Next, we minimise $E(\alpha)$ as a function of α. This is just ordinary calculus, but it is tricky. Therefore, we minimise the function graphically: We introduce an auxiliary function

$$D(\alpha; \mu) = \alpha^2 - \frac{2\alpha}{\left(1 + \frac{\mu}{\alpha}\right)^2}, \qquad \mu = Ma/2,$$

plot the function $D(\alpha)$ for different values of the parameter μ, and obtain the minimum that way. Fig. 23.1 shows the curve $D(\alpha; \mu = 0)$ and $D(\alpha; \mu = 0.4)$. Both curves possess minima. The

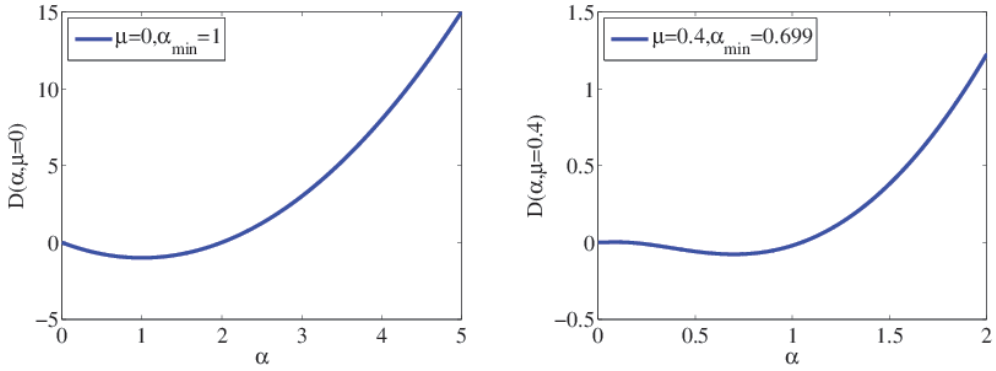


Figure 23.1: Minimisation procedure for computing the ground-state energy of the Yukawa potential.

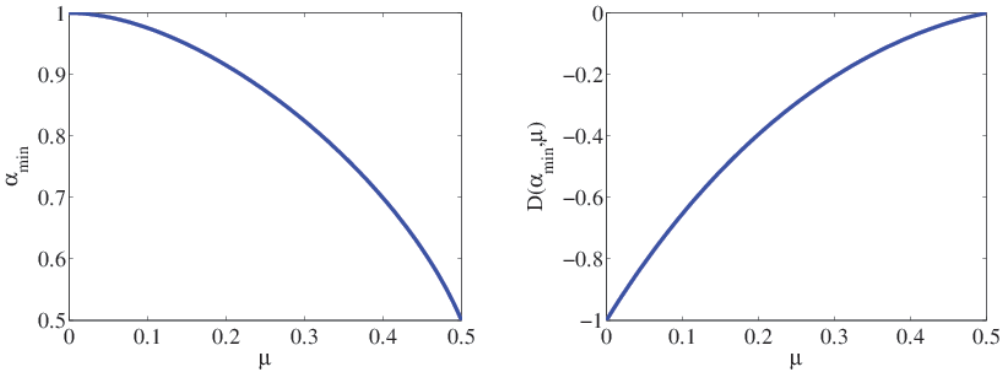


Figure 23.2: Minimisation procedure for computing the ground-state energy of the Yukawa potential (continued).

minimum at $\mu = 0$ is at exactly $\alpha_{\min} = 1$ – precisely the value for a hydrogenic system. The minimum value of α clearly decreases as μ is increased. The minimum disappears completely at $\mu = 0.5$. Thus, only for cases below this value can the hydrogenic model be used to construct bound states. Finally, the parametric dependence of $\alpha_{\min}$ and $D_{\min}$ on μ is shown in Fig. 23.2. In conclusion, our estimate for the ground-state energy of the Yukawa system is

$$E_0 \approx \frac{\hbar^2}{2ma^2} D(\alpha_{\min}; \mu),$$

where $\mu = Ma/2$, and where the non-dimensional function $D(\alpha_{\min}; \mu)$ is obtained graphically. Note that $D(\alpha_{\min}; \mu) < 0$ for $\mu < 1/2$: the model ground state is indeed a bound state, provided M is not too large.

23.4 Ground state of helium

Helium contains two electrons orbiting a nucleus of two protons and two neutrons. Thus, the nucleus has charge $+2e$. We know from the discussion in Ch. 18 that the electrons in a multi-electron atom occupy single-particle-like states. Thus, we fill the single-particle states of the helium atom with electrons. The ground state has no orbital angular momentum, and the only electron quantum number that can vary is therefore the spin. Energy minimisation dictates that both electrons occupy the ground state, thus they must have opposite spin. Thus, the spin component of the wavefunction is antisymmetric, and the spatial part is symmetric. The spatial part can therefore be written as

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \psi_{\rm gs}(\boldsymbol{r}_1)\psi_{\rm gs}(\boldsymbol{r}_2).$$

We have absolutely no idea what $\psi_{\rm gs}(\cdot)$ is. In a naive picture, we might assume that the electron ignores its neighbour altogether, and experiences the bare nuclear charge. Then, the ground-state wavefunction would he hydrogenic:

$$\psi_{\rm gs}(\boldsymbol{r}) = \frac{Z^{3/2}}{\sqrt{\pi}a_0^{3/2}}e^{-Z|\boldsymbol{r}|/a_0}, \qquad a_0 = \frac{4\pi\epsilon_0\hbar^2}{me^2}, \qquad Z = 2,$$

where $Z = 2$ is the number of positive charges in the nucleus. In this picture, the total ground-state wavefunction would be

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \frac{Z^{3/2}}{\sqrt{\pi}a_0^{3/2}}e^{-Z|\boldsymbol{r}_1|/a_0}\frac{Z^{3/2}}{\sqrt{\pi}a_0^{3/2}}e^{-Z|\boldsymbol{r}_2|/a_0} = \frac{Z^3}{\pi a_0^3}e^{-Z(|\boldsymbol{r}_1|+|\boldsymbol{r}_2|)/a_0}.$$

However, a more sophisticated picture involves taking account of the effect of one electron on the other. Thus, we imagine that electron B 'gets in the way' of electron A, and effectively reduces the amount of positive charge electron A experiences from interacting with the nucleus. Therefore, instead of a nuclear charge of Ze, electron A experiences a 'screened' nuclear charge αe, where α is some unknown number between 0 and 2. Thus, our more sophisticated estimate for the ground-state wavefunction is simply

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2; \alpha) = \frac{\alpha^3}{\pi a_0^3}e^{-\alpha(r_1+r_1)/a_0};$$

note that this state is normalised. We compute the expectation value of the Hamiltonian

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla_1^2 - \frac{\hbar^2}{2m}\nabla_2^2 - \frac{Ze^2}{4\pi\epsilon_0 r_1} - \frac{Ze^2}{4\pi\epsilon_0 r_2} + \frac{e^2}{4\pi\epsilon_0|\boldsymbol{r}_1 - \boldsymbol{r}_2|}$$

in the state $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2; \alpha)$ (note the inclusion of the electron-electron interaction term). We have,

- Kinetic energy term:
$$\langle\psi|\left(-\frac{\hbar^2}{2m}\nabla_1^2-\frac{\hbar^2}{2m}\nabla_2^2\right)|\psi\rangle=\frac{\hbar^2\alpha^2}{ma_0^2};$$

- Potential term (interactions with nucleus):
$$\langle\psi|\left(-\frac{Ze^2}{4\pi\epsilon_0 r_1}-\frac{Ze^2}{4\pi\epsilon_0 r_2}\right)|\psi\rangle=-\frac{2Ze^2}{4\pi\epsilon_0 a_0}\alpha;$$

- Electron-electron interaction term:
$$\langle\psi|\frac{e^2}{4\pi\epsilon_0|\boldsymbol{r}_1-\boldsymbol{r}_2|}|\psi\rangle=\frac{e^2}{4\pi\epsilon_0}\left(\frac{Z^3}{\pi a_0^3}\right)^2\int\int \mathrm{d}^3r_1\,\mathrm{d}^3r_2\frac{1}{|\boldsymbol{r}_1-\boldsymbol{r}_2|}e^{-2Z(|\boldsymbol{r}_1|+|\boldsymbol{r}_2|)/a_0}.$$

  The integral is tricky but it can be done analytically. We are left with
$$\langle\psi|\frac{e^2}{4\pi\epsilon_0|\boldsymbol{r}_1-\boldsymbol{r}_2|}|\psi\rangle=\tfrac{5}{4}\alpha\mathrm{Ry},$$

  where
$$1\,\mathrm{Ry}=\frac{\hbar^2}{2ma_0^2}=13.6\mathrm{eV}.$$

Putting it all together, we have

$$E(\alpha)=\langle\psi(\boldsymbol{r}_1,\boldsymbol{r}_2;\alpha)|\hat{H}|\psi(\boldsymbol{r}_1,\boldsymbol{r}_2;\alpha)\rangle=\left[\alpha^2-\left(2Z-\tfrac{5}{8}\right)\right]\mathrm{Ry}.$$

Computing $dE/d\alpha=0$ gives

$$\alpha=Z-\tfrac{5}{16}=Z_{\mathrm{eff}}.$$

The corresponding energy is

$$E(\alpha=Z_{\mathrm{eff}})=-2\left(Z-\tfrac{5}{16}\right)^2\mathrm{Ry}.$$

For helium, we have $Z=2$, hence

$$E(\alpha=Z_{\mathrm{eff}})=-2\left(\tfrac{27}{16}\right)^2\mathrm{Ry}\approx-5.7\mathrm{Ry}.$$

The true (measured) ground-state energy is

$$E_0=-5.81\mathrm{Ry}.$$

Our estimate is true to within 2% – a remarkable agreement! This reinforces the claim made in Ch. 18 that the electrons in an atom live in states that resemble single-particle states.

Chapter 24

Numerical methods

In this section we develop a numerical method to solve the one-dimensional eigenvalue problem

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \mathcal{U}(x)\psi = E\psi$$

using the **Chebyshev collocation method**. Before doing this, we outline the method for a simpler problem, for which analytical solutions are known.

The following books might help in understanding this last chapter:

- *Chebyshev and Fourier spectral methods*, J. P. Boyd, Dover Publications (2000). Boyd himself has put a copy of this on his website and is therefore available for free in pdf form.
- *Spectral methods in Matlab*, L. N. Trefethen, SIAM Publications (2001).

You will see that this section of the course is more contemporary than others!

24.1 A simpler problem

Consider the equation[1]

$$\frac{d^2f}{dy^2} = -\lambda f, \qquad y \in [-L/2, L/2]\,,$$

which is to be solved with **vanishing boundary conditions**

$$f(-L/2) = f(L/2) = 0.$$

[1] Matlab code: `simple.m`

This is an **eigenvalue problem** in the eigenvalue λ. However, we already know the solution: it is

$$f(y) = f_n(y) = \sin(\sqrt{\lambda_n}y), \qquad \lambda_n = \frac{4\pi^2}{L^2}n^2, \qquad n = 1, 2, \cdots$$

or

$$f(y) = f_n(y) = \cos(\sqrt{\lambda_n}y), \qquad \lambda_n = \frac{4\pi^2}{L^2}\left(n + \tfrac{1}{2}\right)^2, \qquad n = 0, 1, \cdots$$

where the apparently free parameter λ is now forced to take discrete values, $\lambda = \lambda_n$.

We are now going to 'shoot a pigeon with a cannon', and solve this problem numerically. We are going to expand the solution in terms of a set of **basis functions**,

$$f(y) = \sum_{n=0}^{\infty} a_n T_n(x), \qquad x = \frac{2}{L}y,$$

where $\{T_n(x)\}_{n=0}^{\infty}$ are a complete set of basis functions on the interval $[-1, 1]$ called the Chebyshev polynomials:

$$T_n(x) = \cos(n \arccos(x)).$$

Although this does not really look like a polynomial in x, it is!. The first few are shown here:

$$\begin{aligned} T_0(x) &= 1, \\ T_1(x) &= x, \\ T_2(x) &= 2x^2 - 1, \\ T_3(x) &= 4x^3 - 3x, \\ T_4(x) &= 8x^4 - 8x^2 + 1. \end{aligned}$$

For more information on the properties of these functions, you may, in this instance, check out the Wikipedia article. I can personally vouch for this article since I have contributed to it myself!

Just as

$$\left\{1, \sin\left(\frac{2n\pi}{L}x\right), \cos\left(\frac{2n\pi}{L}x\right)\right\}_{n=1}^{\infty}$$

are a good set of basis functions for **periodic functions** on an interval $[-L/2, L/2]$, so too are the Chebyshev polynomials for **arbitrary functions** on the same interval. Thus, we in expanding the solution in terms of these exotic functions, instead of familiar sines and cosines, we are taking into account the fact that the solution is not necessarily periodic. Of course, we must truncate the expansion in a numerical framework, so we work with the approximate solution

$$f_N(y) = \sum_{n=0}^{N} a_n T_n(x).$$

There are $N+1$ undetermined coefficients and two boundary conditions. That leaves $N-1$ conditions to obtain. We therefore evaluate the ODE at $N-1$ interior points to give $N+1$ constraints on the coefficients:

$$\begin{aligned} f_N(-L/2) &= 0, \\ \left.\frac{d^2 f_N}{dy^2}\right|_{y_1} &= -\lambda f_N(y_1), \\ \vdots & \quad \vdots \\ \left.\frac{d^2 f_N}{dy^2}\right|_{y_{N-1}} &= -\lambda f_N(y_{N-1}), \\ f_N(+L/2) &= 0, \end{aligned}$$

or

$$\begin{aligned} \sum_{n=0}^{N} a_n T_n(-1) &= 0, \\ \sum_{n=0}^{N} a_n \left(\frac{2}{L}\right)^2 T_n''(x_1) &= -\lambda \sum_{n=0}^{N} a_n T_n(x_1), \\ \vdots & \quad \vdots \\ \sum_{n=0}^{N} a_n \left(\frac{2}{L}\right)^2 T_n''(x_{N-1}) &= -\lambda \sum_{n=0}^{N} a_n T_n(x_{N-1}), \\ \sum_{n=0}^{N} a_n T_n(+1) &= 0. \end{aligned}$$

The interior points are NOT arbitrary: we evaluate at the $N-1$ points

$$x_1,\, x_2, \cdots, x_{N-1} = \cos\left(\frac{\pi}{N}\right),\, \cos\left(2\frac{\pi}{N}\right), \cdots, \cos\left((N-1)\frac{\pi}{N}\right);$$

these are the **collocation points**.

But now we have a **generalised eigenvalue problem**:

$$L\boldsymbol{a} = \lambda M \boldsymbol{a},$$

where

$$L = \begin{pmatrix} T_0(-1) & \cdots & T_N(-1) \\ (2/L)^2 T_0''(x_1) & \cdots & (2/L)^2 T_N''(x_1) \\ \vdots & & \vdots \\ (2/L)^2 T_0''(x_{N-1}) & \cdots & (2/L)^2 T_N''(x_{N-1}) \\ T_0(+1) & \cdots & T_N(+1) \end{pmatrix},$$

$$M = -\begin{pmatrix} 0 & \cdots & 0 \\ T_0(x_1) & \cdots & T_N(x_1) \\ \vdots & & \vdots \\ T_0(x_{N-1}) & \cdots & T_N(x_{N-1}) \\ 0 & \cdots & 0 \end{pmatrix},$$

and

$$\boldsymbol{a} = (a_0, \cdots, a_n)^T.$$

This is a standard problem, and can be solved using a numerical package, such as 'eig' in Matlab.

- Typing

  ```
  d=eig(L,M);
  ```

  in Matlab yields the first $N+1$ eigenvalues.

- We must then check that the eigenvalues are real (a check for bugs in the code):

  ```
  plot(imag(d),'o')
  ```

- Having done that, we sort the eigenvalues in increasing order:

  ```
  d=sort(d);
  ```

- Then, we plot the results.

  ```
  plot(d,'o')
  ```

- Typically, the solver yields an accurate answer only for the first few eigenvalues. Suppose we want to find the first two eigenvalues accurately. We fix N and compute the first two eigenvalues. We then increase N and compute the eigenvalues again. We continue increasing N until the first two eigenvalues do not change upon varying N. The solver is then said to have **converged**.

Happily, these solvers such as 'eig' tell us the eigenvectors as well as the eigenvalues. Typing

```
[V,D]=eig(L,M);
```

gives two $(N+1)\times(N+1)$ matrices. The matrix D is diagonal and corresponds to the eigenvalues,

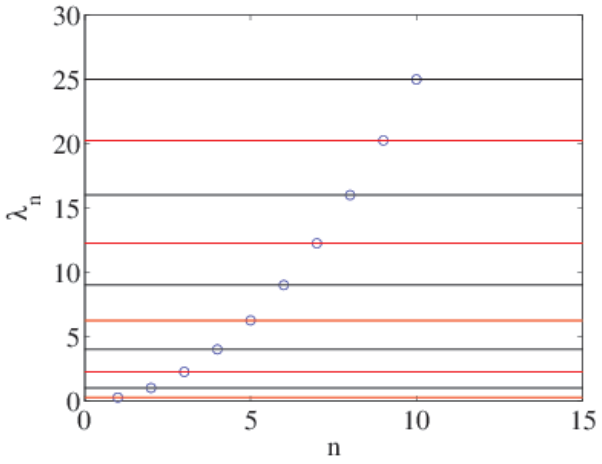


Figure 24.1: The spectrum of the problem $f''(y) = -\lambda f(y)$: comparison between numerical method and theory. Here $N = 100$ and $L = 2\pi$.

```
for i=1:(N+1)
  d(i)=D(i,i);
end
```

while the matrix V corresponds to the eigenvectors. Suppose we want to find the leading eigenvector. We would pick out the leading eigenvalue:

```
[maxd,imax]=max(d);
```

(do NOT sort them!). The corresponding eigenvector is

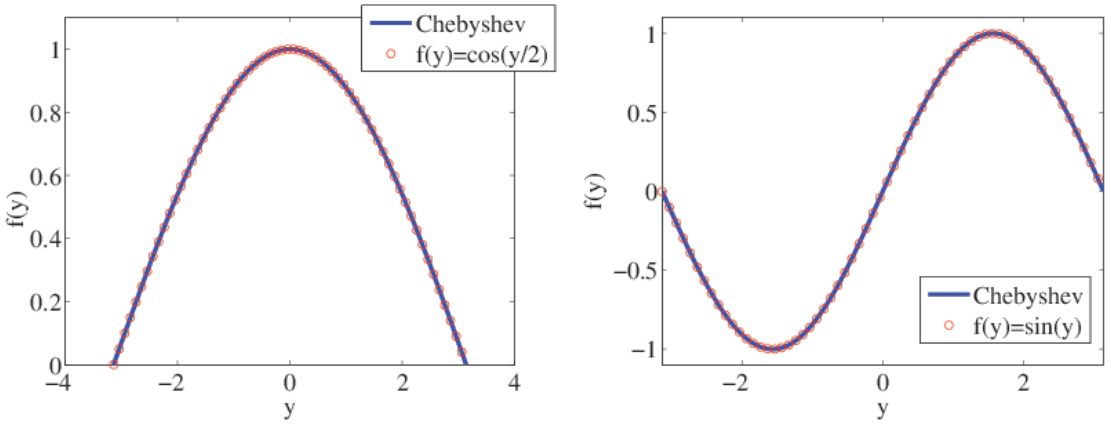
```
a=V(:,imax),
```

i.e. the $imax^{\text{th}}$ column of the matrix V. Finally then, our guess for the leading vector is

$$f_N(y) = \sum_{n=0}^{n} a_n T_n(x), \qquad x = \frac{2}{L} y.$$

The results of implementing this algorithm, with $N = 100$, are shown in Fig. 24.1. The first ten numerically-generated modes are shown in the figure (dots), along with the analytical modes: red lines for $\lambda = (n + 1/2)^2$, and black lines for $\lambda = n^2$ (Here $L = 2\pi$). The two calculations agree exactly. I have also picked out the first two modes and computed the corresponding eigenfunctions (Fig. 24.2)[2]. These eigenfunctions are $\psi = \cos(y/2)$ (lowest), and $\psi = \sin(y)$ (second lowest). Again, the exact calculation and the numerical calculation agree very well. In the next section, we answer the question, 'how well?'

[2]Matlab code: `make_eigenfunction_simple.m`



(a) (b)

Figure 24.2: The first two eigenfunctions of the problem $f''(y) = -\lambda f(y)$. Here $N = 100$ and $L = 2\pi$.

24.2 Exponential convergence

In this section we examine some numerical issues surrounding the Chebyshev collocation method. So far we have been quite casual in our use of nomenclature. For definiteness, we work on the interval $[-1, 1]$. We start with the operator problem

$$\mathcal{L}f = \lambda \mathcal{M} f,$$

and construct the approximate solution

$$f(y) \approx f_N(y) = \sum_{n=0}^{N} a_n T_n(x), \qquad x \in [-1, 1]\,.$$

Until now, we have called this a **truncation**, although really it is an **interpolation**. Let's see why the latter label is more appropriate.

First, recall the following result, due to Lagrange:

Theorem 24.1 *Let $f(x)$ be some function whose value is known at the discete points $x_0, x_1, \cdots, x_N$. Then there exist polynomials $C_0(x)$, $C_1(x), \cdots, C_N(x)$ such that the function*

$$P_N(x) = \sum_{i=0}^{N} f(x_i) C_i(x)$$

agrees with $f(x)$ at the points $x_0, x_1, \cdots, x_N$:

$$P_N(x_i) = f(x_i), \qquad i = 0, 1, \cdots, N.$$

Proof: Take

$$C_i(x) = \prod_{j=0, j\neq i}^{N} \frac{x - x_j}{x_i - x_j}.$$

Noting that

$$C_i(x_k) = \delta_{ik},$$

the result follows. This result establishes the existence of interpolating polynomials, but does not tell us which ones are best. It turns out that the Chebyshev polynomials are among the better polynomials, and that the non-uniform Chebyshev grid is best. In what follows, we explain why.

For illustration purposes, consider the problem $\mathcal{L}f = \lambda \mathcal{M} f$ where boundary conditions are not important. We pose the interpolation approximation

$$f_N(x) = \tfrac{1}{2} b_0 T_0(x) + \sum_{n=1}^{N-1} b_n T_n(x) + \tfrac{1}{2} b_N T_N(x)$$

We impose the condition that $f_N(x)$ and $f(x)$ agree exactly at the points $x_0, x_1, \cdots, x_N$. We do not know the value of $f(x)$, but we do know the differential equation it solves. Thus, we have

$$\mathcal{L} f_N(x_k) = \lambda \mathcal{M} f_N(x_k), \qquad k = 0, 1, \cdots N.$$

Then the following theorem holds:

Theorem 24.2 *Let the interpolation grid be given by*

$$x_k = \cos(k\pi/N), \qquad k = 0, 1, \cdots N.$$

Let $f_N(x)$ be the interpolating polynomial of degree N which interpolates to $f(x)$ on this grid:

$$f_N(x) = \tfrac{1}{2} b_0 T_0(x) + \sum_{n=1}^{N-1} b_n T_n(x) + \tfrac{1}{2} b_N T_N(x).$$

Finally, let $\{\alpha_n\}_n$ be the coefficients of the exact expansion of $f(x)$ in Chebyshev polynomials:

$$f(x) = \tfrac{1}{2} \alpha_0 T_0(x) + \sum_{n=1}^{\infty} \alpha_n T_n(x)$$

Then,

$$b_n = \frac{2}{N} \left[\tfrac{1}{2} f(x_0) T_n(x_0) + \sum_{k=0}^{N-1} f(x_k) T_n(x_k) + \tfrac{1}{2} f(x_N) T_n(x_N) \right],$$

which leads to the following bound:

$$|f(x) - f_N(x)| \leq 2 \sum_{n=N+1}^{\infty} |\alpha_n|.$$

Unfortunately, the proof of this theorem is beyond the scope of this course. Happily, however, we can prove the following corollary:

Theorem 24.3 *If the problem $\mathcal{L}f = \lambda \mathcal{M} f$ is analytic, then the convergence of the interpolation approximation in Theorem 24.2 is exponential.*

Proof: If there are no singularities in the problem $\mathcal{L}f = \lambda \mathcal{M} f$, then a power-series solution is possible, with finite radius of convergence. Continuing the power series into the complex plane gives a solution that has derivatives of all order. Thus, we may assume that

$$|f^{(p)}(x)| \leq M_p,$$

where the bound is independent of $x \in [-1, 1]$.

Next, we note that a Chebyshev series is but a Fourier series in disguise! For, let $\theta = \arccos(x)$. Then,

$$f(x) = \tfrac{1}{2}\alpha_0 + \sum_{n=1}^{\infty} \alpha_n T_n(x) = \tfrac{1}{2}\alpha_0 + \sum_{n=1}^{\infty} \alpha_n \cos(n\theta).$$

Differentiating both sides p times w.r.t. θ gives

$$\sum_{n=1}^{\infty} \alpha_n n^p \Re\left(\mathrm{i}^p \mathrm{e}^{\mathrm{i}n\theta}\right) = \frac{d^p f}{d\theta^p}.$$

But note:

$$\begin{aligned} \frac{df}{d\theta} &= \frac{dx}{d\theta}\frac{df}{dx} = -\sin\theta \frac{df}{dx}, \\ \frac{d^2 f}{d\theta} &= \sin^2\theta \frac{d^2 f}{dx^2} - \cos\theta \frac{df}{dx}, \end{aligned}$$

and so on, implying that $|d^p f / d\theta^p| \leq \widetilde{M}_p$, where the bound is independent of θ or x. Hence,

$$\left| \sum_{n=0}^{\infty} \alpha_n n^p \Re\left(\mathrm{i}^p \mathrm{e}^{\mathrm{i}n\theta}\right) \right| \leq \tilde{M}_p,$$

and this is a convergent series. It follows that the general term tends to zero:

$$\lim_{n\to\infty} |\alpha_n| n^p = 0.$$

At worst,

$$|\alpha_n| \leq Ae^{-\gamma n^\delta}, \qquad n \to \infty,$$

for some positive parameters A and γ, and δ that are independent of n. Hence, there exists $N_0 \in \mathbb{N}$ such that

$$|\alpha_n| < Ae^{-\gamma n^\delta}, \qquad \text{for all } n > N_0.$$

Returning to the bound in the Theorem 24.2, we have

$$\begin{aligned} |f(x) - f_{N_0}(x)| &\leq 2 \sum_{n=N_0+1}^{\infty} |\alpha_n|, \\ &\leq 2A \sum_{n=N_0+1}^{\infty} e^{-\gamma n^\delta}, \\ &\leq 2Ae^{-\gamma(N_0+1)^\delta} \sum_{r=0}^{\infty} e^{-\gamma r^\delta}, \\ &\leq Be^{-\gamma N_0^\delta} \end{aligned}$$

The error is thus proportional to $e^{-\gamma N_0^\delta}$ and we therefore say that the Chebyshev collocation method **converges exponentially**. Typically, this result generalises to situations where the boundary conditions are built in to the interpolation coefficients.

24.3 The Schrödinger equation

We return to the Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \mathcal{U}(x)\psi = E\psi.$$

As usual, we multiply up by $2m/\hbar^2$:

$$-\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}\mathcal{U}(x)\psi = \left(2mE/\hbar^2\right)\psi.$$

We are going to assume that there is **typical value** of the potential energy, such that

$$\mathcal{U}(x) = \mathcal{U}_0 v(x),$$

such that $v(x)$ is a dimensionless shape function. Thus,

$$-\frac{d^2\psi}{dx^2} + \frac{2m\mathcal{U}_0}{\hbar^2}v(x)\psi = \left(2mE/\hbar^2\right)\psi.$$

Now

$$\frac{1}{a^2} = +\frac{2m\mathcal{U}_0}{\hbar^2}$$

defines a typical lengthscale a, and we are left with

$$-\frac{d^2\psi}{dx^2} + \frac{1}{a^2}\upsilon(x)\psi = \left(2mE/\hbar^2\right)\psi.$$

However, we are going to define a dimensionless distance variable,

$$y := x/a,$$

hence

$$-\frac{1}{a^2}\frac{d^2\psi}{ds^2} + +\frac{1}{a^2}\upsilon(y)\psi = \frac{1}{a^2}\left(2mE/\hbar^2\right)\psi.$$

Calling

$$\lambda := \frac{1}{a^2}\left(2mE/\hbar^2\right),$$

we are left with the following eigenvalue problem:

$$\frac{d^2\psi}{dy^2} - \upsilon(y)\psi = -\lambda\psi, \qquad (*)$$

which we call the non-dimensional Schrödinger equation (NDSE). We now solve this equation numerically.

We would like to expand the solution in terms of Chebyshev polynomials. However, the interval of these polynomials is $[-1, 1]$, while the non-dimensional Schrödinger equation (*) is defined on the whole line. Therefore, we introduce a coordinate transformation,

$$y = \frac{\alpha x}{\sqrt{1-x^2}}, \qquad x = \frac{y}{\sqrt{\alpha^2+x^2}},$$

where α is a positive real parameter that can be varied. Letting $x \in [-1, 1]$ gives a y-variable that ranges over the whole real line; the points $x = \pm 1$ correspond to $y = \pm\infty$. Thus, we propose an approximate solution

$$\psi_N(y) = \sum_{n=0}^{N} a_n T_n(x).$$

The second derivative of the approximate solution is

$$\frac{d^2\psi_N}{dy^2} = \sum_{n=0}^{N} a_n \left[\left(\frac{dx}{dy}\right)^2 T_n''(x) + \frac{d^2x}{dy^2}T_n'(x)\right],$$

where

$$\frac{dx}{dy} = \frac{\alpha^2}{(\alpha^2+y^2)^{3/2}}, \qquad \frac{d^2x}{dy^2} = \frac{-3y\alpha^2}{(\alpha^2+y^2)^{5/2}}.$$

We now proceed to set up the collocation matrices.

The boundary conditions require that ψ should vanish at $|y| = \infty$:

$$\psi(y = \pm\infty) = \psi(x = \pm 1) = 0,$$

hence

$$\sum_{n=0}^{N} a_n T_n(\pm 1) = 0.$$

This gives two conditions on $N+1$ unknowns. To obtain the $N-1$ other conditions, we evaluate the trial solution against the differential equation at $N-1$ interior points,

$$x_k = \cos\left(k\frac{\pi}{N}\right), \qquad k = 1, 2, \cdots, N-1,$$

where

$$y_k = \frac{\alpha x_i}{\sqrt{1-x_k^2}}.$$

This gives

$$\sum_{n=0}^{N} a_n \left[\left(\frac{dx}{dy}\right)^2_{x_k} T_n''(x_k) + \left(\frac{d^2x}{dy^2}\right)_{x_k} T_n'(x_k)\right] - \upsilon(y_k)\sum_{n=0}^{N} a_n T_n''(x_k) = \sum_{n=0}^{N} a_n T_n''(x_k).$$

We therefore form the following matrices:

$$L = \begin{pmatrix} T_0(-1) & \cdots & T_N(-1) \\ \left(\frac{dx}{dy}\right)^2_{x_1} T_0''(x_1) + \left(\frac{d^2x}{dy^2}\right)_{x_1} T_0'(x_1) - \upsilon(y_1) & \cdots & \left(\frac{dx}{dy}\right)^2_{x_1} T_N''(x_1) + \left(\frac{d^2x}{dy^2}\right)_{x_1} T_N'(x_1) - \upsilon(y_1) \\ \vdots & & \vdots \\ \left(\frac{dx}{dy}\right)^2_{x_{N-1}} T_0''(x_{N-1}) + \left(\frac{d^2x}{dy^2}\right)_{x_{N-1}} T_0'(x_{N-1}) - \upsilon(y_{N-1}) & \cdots & \left(\frac{dx}{dy}\right)^2_{x_{N-1}} T_0''(x_{N-1}) + \left(\frac{d^2x}{dy^2}\right)_{x_{N-1}} T_N'(x_{N-1}) - \upsilon(y_{N-1}) \\ T_0(+1) & \cdots & T_N(+1) \end{pmatrix}$$

$$M = -\begin{pmatrix} 0 & \cdots & 0 \\ T_0(x_1) & \cdots & T_N(x_1) \\ \vdots & & \vdots \\ T_0(x_{N-1}) & \cdots & T_N(x_{N-1}) \\ 0 & \cdots & 0 \end{pmatrix},$$

and

$$\boldsymbol{a} = (a_0, \cdots, a_n)^T,$$

to give an eigenvalue problem

$$L\boldsymbol{a} = \lambda M\boldsymbol{a}.$$

24.4 Harmonic oscillator revisited

In this section, we again use a cannon to shoot birds, and apply the Chebyshev collocation method to the harmonic oscillator[3]. To do this, we must first of all write down the NDSE.

Eigenvalue problem:

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial y^2} + \tfrac{1}{2}m\omega^2x^2\psi = E\psi;$$

multiply up by $2m/\hbar^2$:

$$-\frac{\partial^2\psi}{\partial x^2} + \frac{m^2\omega^2}{\hbar^2}x^2\psi = (2mE/\hbar^2)\psi.$$

Identify a standard unit of length:

$$\frac{1}{a^4} = \frac{m^2\omega^2}{\hbar^2}, \qquad a = \sqrt{\hbar/m\omega}.$$

Hence, the eigenvalue problem reads

$$-\frac{\partial^2\psi}{\partial x^2} + \frac{1}{a^4}x^2\psi = (2mE/\hbar^2)\psi.$$

Identify a non-dimensional variable of length:

$$y = x/a :$$

$$-\frac{1}{a^2}\frac{\partial^2\psi}{\partial s^2} + \frac{1}{a^2}s^2\psi = (2mE/\hbar^2)\psi.$$

Multiply up by a^2 and then by -1:

$$\frac{\partial^2\psi}{\partial s^2} - s^2\psi = -\lambda\psi, \qquad \lambda = 2mE(a^2/\hbar^2).$$

The NDSE for the harmonic oscillator is therefore

$$\frac{\partial^2\psi}{\partial s^2} - s^2\psi = -\lambda\psi, \qquad v(s) = s^2.$$

[3]Matlab code: `schrodinger1.m`, with `u=y*y` in lines 155-161

As in the previous section, we propose the solution

$$\psi_N(y) = \sum_{n=0}^{N} a_n T_n(x),$$

where

$$y = \frac{\alpha x}{\sqrt{1-x^2}}, \qquad x = \frac{y}{\sqrt{\alpha^2+x^2}}.$$

We introduce collocation points

$$x_k = \cos\left(k\frac{\pi}{N}\right), \qquad k = 1, 2, \cdots, N-1,$$

or

$$y_k = \frac{\alpha x_i}{\sqrt{1-x_k^2}}.$$

This gives

$$L = \begin{pmatrix} T_0(-1) & \cdots & T_N(-1) \\ \left(\frac{dx}{dy}\right)^2_{x_1} T_0''(x_1) + \left(\frac{d^2x}{dy^2}\right)_{x_1} T_0'(x_1) - y_1^2 & \cdots & \left(\frac{dx}{dy}\right)^2_{x_1} T_N''(x_1) + \left(\frac{d^2x}{dy^2}\right)_{x_1} T_N'(x_1) - y_1^2 \\ \vdots & & \vdots \\ \left(\frac{dx}{dy}\right)^2_{x_{N-1}} T_0''(x_{N-1}) + \left(\frac{d^2x}{dy^2}\right)_{x_{N-1}} T_0'(x_{N-1}) - y_{N-1}^2 & \cdots & \left(\frac{dx}{dy}\right)^2_{x_{N-1}} T_0''(x_{N-1}) + \left(\frac{d^2x}{dy^2}\right)_{x_{N-1}} T_N'(x_{N-1}) - y_{N-1}^2 \\ T_0(+1) & \cdots & T_N(+1) \end{pmatrix}$$

$$M = -\begin{pmatrix} 0 & \cdots & 0 \\ T_0(x_1) & \cdots & T_N(x_1) \\ \vdots & & \vdots \\ T_0(x_{N-1}) & \cdots & T_N(x_{N-1}) \\ 0 & \cdots & 0 \end{pmatrix},$$

and

$$\boldsymbol{a} = (a_0, \cdots, a_n)^T.$$

The eigenvalues are obtained by solving

$$L\boldsymbol{a} = \lambda M \boldsymbol{a}.$$

The spectrum is shown in Fig. 24.3 The first few numerically-generated modes are shown in the figure (dots), along with the analytical modes: the parameter $2n$ in the Hermite differential equation corresponds exactly to $\lambda - 1$, hence $\lambda = \lambda_n = 2n + 1$, where $n = 0, 1, \cdots$. The two calculations agree exactly for small n-values. The agreement is spoilt for higher n-values. However, increasing N beyond $N = 100$ yields better agreement for this portion of the spectrum.

I have also picked out the first mode and computed the corresponding eigenfunction (Fig. 24.4).

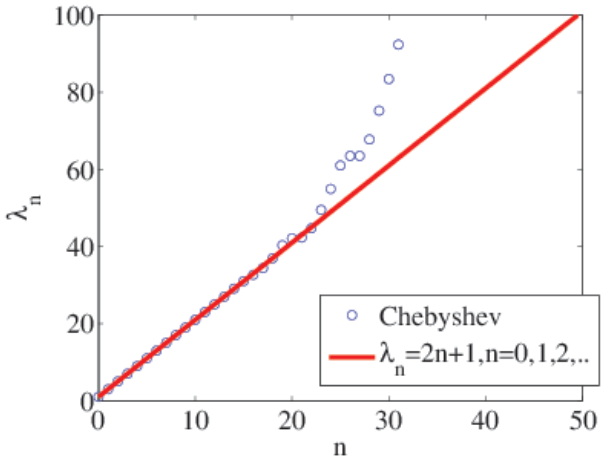


Figure 24.3: The spectrum of quantum harmonic oscillator: comparison between numerical method and theory. Here $N = 100$.

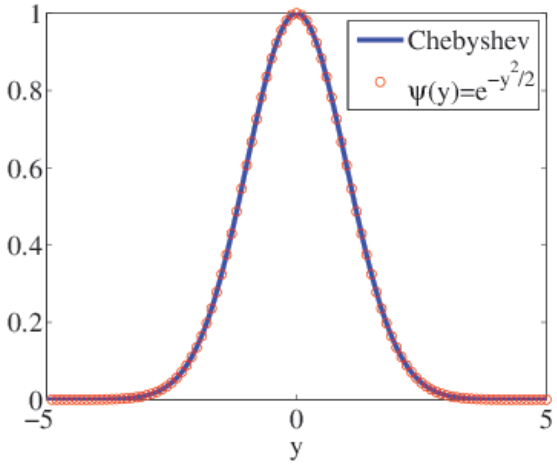


Figure 24.4: First eigenfunction of the quantum harmomic oscillator: comparison between numerical method and theory. Here $N = 100$.

This eigenfunction is $\psi = e^{-y^2}$. Again, excellent agreement is obtained.

24.5 Exotic potentials

Let's compute the spectrum of the anharmonic oscillator with corresponding eigenvalue problem

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + \left(\tfrac{1}{2}m\omega^2 x^2 + qx^4\right)\psi = E\psi.$$

We need to find the NDSE. However, let's take a shortcut. From Ch. 20, we know that a is the lengthscale, where

$$a = \sqrt{\hbar/m\omega},$$

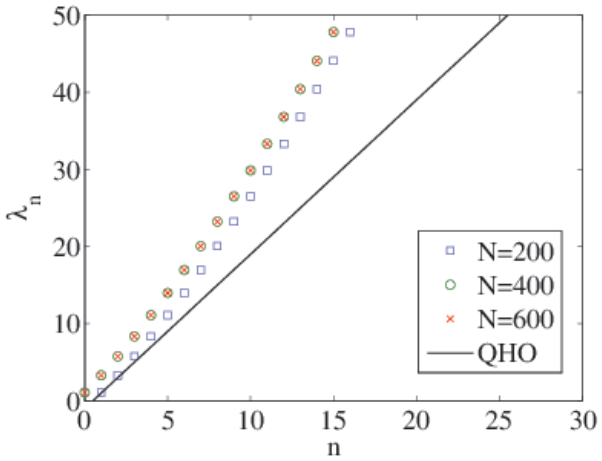


Figure 24.5: The spectrum of the anharmonic oscillator: convergence study. Here $\epsilon = 0.5$.

and that the Schrödinger equation can be re-written as

$$-\frac{\partial\psi}{\partial x^2} + \frac{x^2}{a^4}\psi + \left(\frac{2mqa^6}{\hbar^2}\right)\frac{x^4}{a^6}\psi = (2mE/\hbar^2)\psi.$$

Introducing a parameter $\epsilon := 2mqa^6/\hbar^2$, this is

$$-\frac{\partial\psi}{\partial x^2} + \frac{x^2}{a^4}\psi + \epsilon\frac{x^4}{a^6}\psi = (2mE/\hbar^2)\psi.$$

We identify the non-dimensional distance variable $y = x/a$, hence

$$\frac{\partial\psi}{\partial y^2} - \left(s^2 + \epsilon s^4\right)\psi = -\lambda\psi.$$

As a final check on the correctness of the method, we compute the spectrum of this system with $\epsilon = 0.01$. We expect the ground-state eigenvalue to be

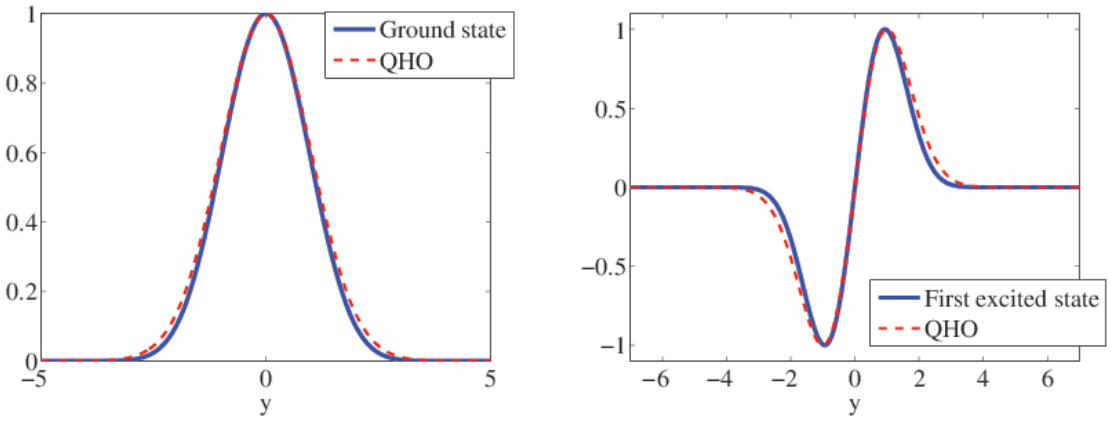
$$\lambda_{\rm gs} = 1 + \tfrac{3}{4}\epsilon = 1.0075.$$

The result with $N = 500$ or $N = 600$ is

$$\lambda_{\rm gs} = 1.007373672,$$

and the small discrepancy can be explained by $O(\epsilon^2)$ terms.

However, we can go beyond perturbation theory int the numerical setting. Thus, we compute the eigenvalues and eigenvectors for $\epsilon = 0.5$. Fig. 24.5 shows that the energy levels of the anharmonic



(a) (b)

Figure 24.6: The first two eigenfunctions anharmonic oscillator. Here $\epsilon = 0.5$ and $N = 400$.

oscillator are shifted above the harmonic analogues, and grow superlinearly ($\lambda_n \sim n^a$, with $a > 1$). Convergence is achieved for low n-values for $N = 400$. Fig. 24.6 shows the first two eigenfunctions[4]. They look very similar to the solution of the ordinary quantum harmonic oscillator! This also indicates why variational methods work so well: typically, the shape of the wavefunctions is determined by symmetry considerations (odd, even), and by the condition that they should vanish at $|x| = \infty$; these conditions place severe constraints on the shape, and thus systems that have the same kind of Hamiltonian will also have the same kind of eigenfunctions.

[4]Matlab code: `make_eigenfunction.m`

Chapter 25

Perspectives

Recall from the introduction that Quantum Mechanics was introduced to eliminate a divergence or an 'infinity' in the calculation for the spectral density of a blackbody. Recall also how relativistic quantum mechanics was needed to develop a correct theory (with the correct prefactors) of spin-orbit coupling in hydrogen.

If you take more advanced courses in quantum-field theory, you will find that this complete theory, which involves the coupling of electronic charge to light, gives rise to other infinities. Feynman managed to get rid of these infinities by introducing a **renormalisation** of the field theory, in which the 'bare' electronic mass and charge are replaced with effective values, thus leading to convergent probability amplitudes.

The Standard Model of particle physics contains only renormalisable operators. However, if you combine general relativity and quantum field theory to obtain a quantum gravity, the result does not appear to be renormalisable if the theory is constructed in a standard fashion. Thus, we are apparently left with infinities.

Personally, I am not holding my breath, waiting for a solution to this problem (I strayed back into classical mechanics, preferring its certainties). Indeed, I would much prefer to know if the **many-worlds** solution to the problem of measurement is valid or not. Again, however, I am not holding my breath. Answers on a postcard (or by other means) to Room 24, Science Building, UCD.

Appendix A

Matlab codes

A.1 Matlab code for generating spherical harmonics

```
function []=test_draw1(ell,m,flag)

dtheta=1;
dphi=1;
phi=0:dphi:360;
theta=0:dtheta:180;

[Phi Theta] = meshgrid(phi,theta);

Theta_rad=Theta*(pi/180);
Phi_rad=Phi*(pi/180);

if(ell==0)
    temp=1+0*Theta_rad;
else
    N=legendre(ell,cos(Theta_rad));
    temp=N(abs(m)+1,:,:);
    temp=((-1)^abs(m))*reshape(temp,length(theta),length(phi));
end

if(m>=0)
    Pellm=temp;
else
```

```
    val1=(-1)^abs(m);
    val2=factorial(ell-abs(m));
    val3=factorial(ell+abs(m));
    Pellm=val1*(val2/val3)*temp;
end

val0=(-1)^m;
val1=(2*ell+1)/(4*pi);
val2=factorial(ell-m);
val3=factorial(ell+m);

size(Pellm)

C=val0*sqrt(val1*val2/val3)*Pellm.*exp(sqrt(-1)*m*Phi_rad);

[x y z] = sph2cart(Phi_rad,(pi/2)-Theta_rad,1);

surf(x,y,z,real(C),'edgecolor','none')

drawnow
colorbar
set(gca,'fontsize',18,'fontname','times new roman')
xlabel('x')
ylabel('y')
lighting phong
axis equal
camlight('right')

end
```

Appendix B

The Hamiltonian Formulation of Classical Mechanics

B.1 Lagrangian mechanics

We start with the **Lagrangian** formulation of classical mechanics (CM), which we studied in ACM 20150: for generalized coordinates $\{q_i\}_{i=1}^N$ and generalized velocities $\{\dot{q}_i\}_{i=1}^N$, the **Lagrangian** encodes all the information about the mechanical system:

$$L = T(q_i, \dot{q}_i) - \mathcal{U}(q_i), \tag{B.1}$$

where T is the kinetic energy and $\mathcal{U}$ is the potential energy. The dynamics of the mechanical system are obtained by imposing stationarity of the action

$$S[q_i, \dot{q}_i] = \int_{t_1}^{t_2} L(q_i, \dot{q}_i)\mathrm{d}t, \tag{B.2}$$

which leads to the following **Euler–Lagrange** equations:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i}, \qquad i = 1, \cdots, N. \tag{B.3}$$

Example: Consider particle motion in one space dimension. Then, q is the position ($q \equiv x$) and $\dot{q}$ is the velocity. The Lagrangian is

$$L = \tfrac{1}{2}m\dot{q}^2 - \mathcal{U}(q), \tag{B.4}$$

and the Euler–Lagrange equations give

$$\frac{d}{dt}(m\dot{q}) = -\mathcal{U}'(q) \implies m\ddot{q} = -\mathcal{U}'(q). \tag{B.5}$$

Thus, the Lagrangian formulation of mechanics implies the Newtonian formulation.

B.2 Hamiltonian mechanics

We define the generalized momenta conjugate the generalized coordinates q_i:

$$p_i := \frac{\partial L}{\partial \dot{q}_i}. \tag{B.6}$$

We are going to regard p_i, q_i, and $\dot{q}_i$ as **independent symbols** and we are going to "get rid of" the $\dot{q}_i$'s from the description of the dynamics. To do this, we introduce a new function:

$$H(p_i, q_i, \dot{q}_i) := \sum_i p_i \dot{q}_i - L(q_i, \dot{q}_i). \tag{B.7}$$

The function H is called the **Legendre transformation** of L w.r.t. the pair $(\dot{q}_i, p_i)$. We have the following theorem:

Theorem: H in Eq. (B.7) is independent of $\dot{q}_i$.

To prove this, it suffices to differentiate H in Eq. (B.7) w.r.t. $\dot{q}_i$ and show that the result is zero:

$$\frac{\partial H}{\partial \dot{q}_i} = p_i - \frac{\partial L}{\partial \dot{q}_i},$$

which is zero by definition. Hence,

$$H = H(q_i, p_i) \tag{B.8}$$

only. The quantity H so constructed is the **Hamiltonian** of the system.

Let us form the differential of H. We do so in two ways, based on Eqs. (B.7) and (B.8) respectively. Consider the **first way**:

$$\begin{aligned}
\mathrm{d}H &= \dot{q}_i \mathrm{d}p_i + p_i \mathrm{d}\dot{q}_i - \frac{\partial L}{\partial q_i}\mathrm{d}q_i - \frac{\partial L}{\partial \dot{q}_i}\mathrm{d}\dot{q}_i, \\
&= \dot{q}_i \mathrm{d}p_i - \frac{\partial L}{\partial q_i}\mathrm{d}q_i + \left(p_i - \frac{\partial L}{\partial \dot{q}_i}\right)\mathrm{d}\dot{q}_i, \\
&= \dot{q}_i \mathrm{d}p_i - \frac{\partial L}{\partial q_i}\mathrm{d}q_i, \\
&\overset{\text{E.L.}}{=} \dot{q}_i \mathrm{d}p_i - \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right)\mathrm{d}q_i, \\
&= \dot{q}_i \mathrm{d}p_i - \left(\frac{dp_i}{dt}\right)\mathrm{d}q_i, \\
&= \dot{q}_i \mathrm{d}p_i - \dot{p}_i \mathrm{d}q_i.
\end{aligned}$$

Consider also the **second way**:

$$\mathrm{d}H = \frac{\partial H}{\partial p_i}\mathrm{d}p_i + \frac{\partial H}{\partial q_i}\mathrm{d}q_i.$$

However, these approaches are totally equivalent, so we have

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \qquad i = 1, \cdots, N. \tag{B.9}$$

These are **Hamilton's equations** of motion.

Example: We return to one-dimensional particle dynamics, where $L = (m\dot{q}^2/2) - \mathcal{U}(q)$. We identify the momentum p conjugate to the generalized coordinate q:

$$p = \frac{\partial L}{\partial \dot{q}} = m\dot{q} \implies \dot{q} = \frac{1}{m}p.$$

The Legendre transformation is

$$\begin{aligned} H &= \dot{q}p - L, \\ &= m\dot{q}^2 - \tfrac{1}{2}m\dot{q}^2 + \mathcal{U}(q), \\ &= \tfrac{1}{2}m\dot{q}^2 + \mathcal{U}(q), \\ &= \frac{1}{2m}p^2 + \mathcal{U}(q), \\ &= H(q,p). \end{aligned}$$

Thus, H is the system's energy!

Also,

$$\frac{\partial H}{\partial q} = \mathcal{U}'(q), \qquad \frac{\partial H}{\partial p} = \frac{1}{m}p,$$

Hence,

$$\dot{q} = \frac{\partial H}{\partial p} \implies \dot{q} = \frac{1}{m}p,$$

and

$$\dot{p} = -\frac{\partial H}{\partial q} \implies \dot{p} = -\mathcal{U}'(q).$$

B.3 Noether's Theorem

A **symmetry** of the mechanical system is some transformation that acts on the system, and leaves the mechanical properties of the system unchanged. Noether's theorem says that **any such sym-**

metry gives rise to a conserved quantity. In this section, we give some demonstrations of this theorem, although we do not prove it in a general setting.

Example: Consider N particles interacting, for which the Hamiltonian is

$$H = \sum_{i=1}^{N} \frac{1}{2m_i} p_i^2 + \mathcal{U}(q_1, \cdots, q_N).$$

Here t does not appear explicitly in H:

$$H = H(q_i, p_i) \implies H(t + \Delta t) = H(t).$$

Thus, there is a conserved quantity associated with time translation $t \to t + \Delta t$. We know automatically what this conserved quantity is: it is H itself. Now, we differentiate H w.r.t. time to check if it is conserved:

$$\frac{dH}{dt} = \frac{\partial H}{\partial q_i}\dot{q}_i + \frac{\partial H}{\partial p_i}\dot{p}_i + \frac{\partial H}{\partial t},$$

and $\partial H/\partial t = 0$ because t does not appear explicitly in H. Thus,

$$\begin{aligned} \frac{dH}{dt} &= \frac{\partial H}{\partial q_i}\dot{q}_i + \frac{\partial H}{\partial p_i}\dot{p}_i, \\ &= \frac{\partial H}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial p_i}\frac{\partial H}{\partial q_i}, \\ &= 0. \end{aligned}$$

Thus, the Noether quantity associated with time translation is H itself. **Energy conservation and the invariance of the system under time translation are intimately linked**. Note finally, this result relies on $\partial_t H = 0$ **and** Hamilton's equations.

We now consider one final example to demonstrate Noether's theorem. We consider the **Kepler problem** in the plane. We start with the associated Lagrangian problem in polar coordinates:

$$L = T - \mathcal{U} = \tfrac{1}{2} m \left(\frac{ds}{dt}\right)^2 - \mathcal{U}(r),$$

where the force is central, such that $\mathcal{U} = \mathcal{U}(r)$ only. Also, we have $\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\varphi^2$, such that

$$\left(\frac{ds}{dt}\right)^2 = \dot{r}^2 + r^2\dot{\varphi}^2,$$

and

$$L = \tfrac{1}{2} m \left(\dot{r}^2 + r^2\dot{\varphi}^2\right) - \mathcal{U}(r).$$

The **generalized coordinates** are (r, φ), and the conjugate momenta are thus

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r}, \qquad p_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = mr^2\dot{\varphi}.$$

We carry out the Legendre transformation:

$$\begin{aligned} H &= p_r\dot{r} + p_\varphi\dot{\varphi} - L, \\ &= m\dot{r}^2 + mr^2\dot{\varphi} - \tfrac{1}{2}m\dot{r}^2 - \tfrac{1}{2}mr^2\dot{\varphi} + \mathcal{U}(r), \\ &= \tfrac{1}{2}m\dot{r}^2 + \tfrac{1}{2}mr^2\dot{\varphi} + \mathcal{U}(r), \\ &= \frac{1}{2m}p_r^2 + \frac{1}{2mr^2}p_\varphi^2 + \mathcal{U}(r). \end{aligned}$$

Also,

$$\frac{\partial H}{\partial p_r} = \frac{1}{m}p_r, \qquad \frac{\partial H}{\partial p_\varphi} = \frac{1}{mr^2}p_\varphi,$$

and

$$\frac{\partial H}{\partial r} = -\frac{1}{m}\frac{p_\varphi^2}{r^3} + \mathcal{U}'(r), \qquad \underline{\frac{\partial H}{\partial \varphi} = 0.}$$

We assemble these partial derivatives into Hamilton's equations. Start with the radial direction:

$$\dot{r} = \frac{\partial H}{\partial p_r}, \qquad \dot{p}_r = -\frac{\partial H}{\partial r}.$$

This gives

$$\dot{r} = \frac{1}{m}p_r, \qquad \dot{p}_r = \frac{1}{m}\frac{p_\varphi^2}{r^3} - \mathcal{U}'(r). \tag{B.10}$$

We also have the tangential direction:

$$\dot{\varphi} = \frac{\partial H}{\partial p_\varphi}, \qquad \dot{p}_\varphi = -\frac{\partial H}{\partial \varphi}.$$

This gives

$$\dot{\varphi} = \frac{1}{mr^2}p_\varphi, \qquad \dot{p}_\varphi = 0.$$

Hence,

$$p_\varphi = \text{Const.} := J$$

This result is substituted into Eq. (B.10) to give

$$\dot{p}_r = \frac{1}{m}\frac{J^2}{r^3} - \mathcal{U}'(r), \qquad p_r = m\dot{r}$$

or

$$m\ddot{r} = \frac{1}{m}\frac{J^2}{r^3} - \mathcal{U}'(r),$$

or

$$m\ddot{r} = -\mathcal{U}'_{\text{eff}}(r), \qquad \mathcal{U}_{\text{eff}}(r) = \frac{J}{2mr^2} + \mathcal{U}(r).$$

Hamilton's formulation is so clever that we have almost missed the existence of a Noether conserved quantity! Let us go back and find it. Consider again the Kepler problem

$$H = \frac{1}{2m}p_r^2 + \frac{1}{2mr^2}p_\varphi^2 + \mathcal{U}(r).$$

Here, H is independent of φ: $H = H(r, p_r, p_\varphi)$ only, such that

$$H(\varphi + \Delta\varphi) = H(\varphi).$$

Thus, the mechanical system is invariant under **rotations** $\varphi \to \varphi + \Delta\varphi$. Since $H = H(r, p_r, p_\varphi)$, we have that

$$\frac{\partial H}{\partial \varphi} = 0.$$

But, by Hamilton's equations,

$$\dot{p}_\varphi = -\frac{\partial H}{\partial \varphi} = 0.$$

Thus, the rotational symmetry $\partial H/\partial\varphi = 0$ implies the conservation of angular momentum, $p_\varphi = \text{Const.}$. This connection between rotational symmetry and conservation of angular momentum is not so obvious in the other formulations of mechanics. That and the coordinate-free formulation in Hamilton's equations (i.e. not necessarily Cartesian) makes this particular formulation of CM very useful.