

P3 Quantum Mechanics

Dr Rosaria Lena

Invalid Date

Table of contents

Preface	6
1 Origins of QM	8
Learning objectives	8
1.1 A quick introduction to Quantum Mechanics	8
1.1.1 Quantum mechanics timeline	8
1.1.2 Two equivalent formulations of quantum mechanics	10
1.2 Blackbody radiation and the ultraviolet catastrophe	10
The black-body spectrum... observed? In the 1890s?!	12
1.2.1 Wien's law (1893)	12
1.2.2 Wien's displacement law	16
1.2.3 Rayleigh-Jeans model (1900)	17
1.2.4 Derivation of the Rayleigh-Jeans law	20
1.2.5 Planck's Law (1900)	22
1.2.6 Exercise	26
1.2.7 Exercise	29
1.2.8 Blackbody radiation and the cosmic microwave background	29
1.3 Photoelectric effect	32
1.4 The nuclear atom	34
1.4.1 The 'classical' problems with Rutherford's model	37
1.4.2 The Bohr model	37
1.4.3 Bohr's correspondence principle	37
Historical remark: Arthur Erich Haas	38
1.4.4 Compton's scattering	39
1.4.5 The problems with Bohr's model	42
1.5 De Broglie's hypothesis and matter waves	43
1.5.1 Interpretation of the Bohr's circular orbits	44
1.5.2 Exercise	44
1.5.3 Electron diffraction experiments	46
1.6 (1926) Schrödinger's atomic model: a <i>really</i> quantum model of the atom	47
1.6.1 Electron orbitals	49
1.7 Recap	49
2 Wavefunctions and the uncertainty principle	51
Learning objectives	51

2.1	Wave-particle duality	51
2.2	The double-slit experiment	51
2.2.1	Feynman's lectures	52
2.2.2	Experiment 1: An experiment with bullets	52
2.2.3	Experiment 2: An experiment with waves	53
2.2.4	Experiment 3: An experiment with electrons	55
2.2.5	So is the electron a wave, a particle or both?	57
2.2.6	Experiment 4: An experiment with electrons... but we monitor the electrons through the slits!	58
2.3	Wavefunctions and Born's probability interpretation	60
2.3.1	Exercise	62
2.4	The wave formalism for free particles	63
2.4.1	Free particle as plane wave?	63
2.4.2	Superposition of two plane waves	66
2.4.3	Adding more waves to form a wavepacket	68
2.4.4	Wavepacket dispersion	70
2.4.5	Conservation of probability	71
2.5	Reciprocity in wavepackets and the Heisenberg uncertainty principle	72
2.5.1	Gaussian wavepackets	72
2.5.2	Heisenberg's uncertainty principle - implications	75
2.6	Summary	78
2.7	References	78
3	The Schrödinger equation and some QM formalism	79
	Learning objectives	79
3.1	The importance of the superposition principle	79
3.2	The time-dependent Schrödinger equation	80
3.2.1	Operators in the Schrödinger equation	81
3.2.2	Exercise	82
3.2.3	Continuity condition	83
3.2.4	The time-dependent Schrödinger equation in 3D	83
3.3	Probability conservation	84
3.4	The time-independent Schrödinger equation	86
3.5	Operators, expectation values and eigenfunctions	88
3.5.1	Operators	88
3.5.2	Expectation values	89
3.5.3	Eigenfunctions	91
3.6	General solution of the TDSE - expansion in energy eigenfunctions	93
3.7	Commutators	96
3.7.1	Properties of commutators	96
3.7.2	The canonical commutation relation	97
3.7.3	The generalised uncertainty principle	97
3.7.4	Exercise	97

3.8	Time evolution of expectation values	98
3.8.1	The Ehrenfest theorem and the correspondence principle	99
3.9	Summary	100
4	Solutions of Schrödinger Equations in 1D - Bound states	101
	Learning objectives	101
4.1	Solving a general problem with the TDSE	101
4.2	Solving the TISE	102
4.2.1	Properties of the eigenfunctions	102
4.2.2	Properties of the eigenenergies - Bound states and scattering states	103
4.3	The infinite well potential (particle in a box)	106
4.3.1	Properties of infinite-square-well potential eigenfunctions	109
4.3.2	Parity	111
4.4	The delta-function potential	112
4.5	The finite square well	115
4.5.1	Bound states	116
4.6	The harmonic well potential (quantum harmonic oscillator)	120
4.6.1	An introduction to the Dirac notation	124
4.6.2	Solving the harmonic oscillator	127
4.7	Rules for wavefunctions in general potentials	129
4.7.1	Examples of general potentials	131
4.7.2	Periodic boundary conditions	131
4.8	More on the Dirac notation	136
4.8.1	Bra, ket and operators	136
4.8.2	Inner and outer product	137
4.8.3	Mapping to waves formalism	137
4.8.4	Exercise	140
4.9	Summary	141
5	Solutions of Schrödinger Equations in 1D - Scattering states	142
5.1	Scattering states	142
5.2	Scattering from a step potential	143
5.2.1	Scattering from “allowed” potential step ($E > V$)	144
5.2.2	Scattering from “forbidden” potential step ($E < V$)	146
5.3	Scattering from potential well (double step)	148
5.3.1	Resonant scattering from potential well - Ramsauer effect	149
5.4	Potential barriers and tunneling ($E < V$)	150
6	Angular momentum theory	154
6.1	Schrödinger’s model of the atom and electron orbitals	154
6.1.1	Electron orbitals	154
6.2	Orbital angular momentum (OAM)	155
6.2.1	Commutation relations for the orbital angular momentum	156

6.2.2	Angular momentum in spherical polar coordinates	157
6.2.3	Solution of the 3D TISE in spherical polar coordinates	158
6.2.4	Action of parity, L and L^2 on spherical harmonics	160
6.2.5	Orbital angular momentum and spherical harmonics in Dirac notation .	164
6.2.6	Angular momentum ladder operators	164
6.2.7	Computing angular momentum observables	167
6.3	Spin: intrinsic angular momentum	168
6.4	Total angular momentum	169
6.5	Summary	171
7	Summary of exercises	172
7.1	Chapter 1	172
7.2	Chapter 2	172
7.3	Chapter 3	172
7.4	Chapter 4	172
7.5	Chapter 5	172
8	FAQs	173

Preface

This is a set of notes for the P3 Quantum Mechanics course at the University of Glasgow (2024/2025). The notes were adapted from content previously developed and used by other lecturers who came before me (Prof. Andy Buckley was the last before me), so I do not take full credits for the content, which has not been subjected to major changes this year. This set of notes in this format has been however developed by me, so if you have any issues with it or any feedback about it, please let me know.

The material provided is mainly split into two sources of information:

- Slides: contain the essential and are used only as a guideline during lectures, so they will have to be complemented by notes that you take during lectures. Please *do take notes*, it will help you learn.
- Notes: you may know where to find them already if you are reading this. These are usually (hopefully!) quite comprehensive and include interactive material (at least those in html format) so they are more complete than the slides. You are expected to use these notes as the interactive elements can help you metabolise concepts. Also, it takes me a long time to write them, so it would be great if someone actually uses them (and finds them useful - if you don't, let me know). The notes can be accessed in different formats:
 - html: just follow the link on Moodle. These contain some extra things that I may mention during the lectures (*not all of them!*), worked examples, code scripts, etc...
 - pdf: see download icon at the top left corner in the html link, it is usually updated when I change something, but I'll make sure I update the pdf file on Moodle too. Please note that the pdf produced is not optimised in its structure and, most importantly, does *not* include any interactive elements, so I would encourage you to still read the html notes if possible, which are also more easily accessible and easier to navigate
 - epub: there should also be an epub version but I have not managed to make it work well (yet!)

Recommended textbooks for **this** course (with formality level in brackets):

- “*Introduction to Quantum Mechanics*” — *Griffiths & Schroeter*: Excellent presentation, nicely balanced between formality and accessibility. This course fits in the first 4 chapters, with next year’s QM course covered by the rest. (Beginner)

- “*Quantum Mechanics*” — *Jain*: A more basic introduction, but very well matched to this course. This course covers about 80% of the book. (Beginner)
- “*The Feynman Lectures on Physics, Volume III by Richard P. Feynman*” Feynman’s lectures are famous for their intuitive explanations and conceptual insight. It is a great supplemental resource to build intuition and understanding, though it doesn’t follow the standard textbook structure or problem-solving approach. Personal experience: good for some conceptual understanding of quantum mechanics, especially for the wave-particle duality and the double-slit experiment. This book is online and I have included some chapters in the notes so you can go through them.
- “*Quantum Mechanics*” - *Bransden & Joachain*: This will be often used during this course, still accessible but at times requiring more mathematical rigor. (Intermediate)
- “*Quantum Physics*” — *Gasiorowicz*: An older textbook, with some annoying use of CGS units, but good on many aspects. (Intermediate)
- “*Modern Quantum Mechanics*” — *Sakurai*: An excellent book, taking a modern “backwards” approach to QM that starts with angular momentum and the algebraic structure of the theory, more formal to start with and maybe not the best approach from the beginning. (Intermediate - Advanced)
- “*Lectures on Quantum Mechanics*” — *Weinberg*: An advanced text by a master of the subject. Mainly of interest here for the perspectives on historical development and the “meaning” of QM. (Advanced)

1 Origins of QM

Learning objectives

1. Understanding the key historical context that led to quantum mechanics
2. Identifying key problems in classical physics, such as the ultraviolet catastrophe, and understanding how quantum concepts provided solutions.
3. Explaining and summarising the significance of foundation experiments like black-body radiation, Compton scattering and the photoelectric effect, in the development of quantum mechanics.
4. Understanding the concept of wave-particle duality and its experimental confirmations
5. Tracing the evolution of atomic models from Rutherford to Bohr
6. Recognising the foundational role of early quantum mechanics in modern physics.

This lecture mostly covers Chapter 6 of [1].

1.1 A quick introduction to Quantum Mechanics

What do you know about quantum mechanics?

- Schrödinger's cat?
- Heisenberg's uncertainty principle?
- Double slit experiment?
- Ultraviolet catastrophe?

1.1.1 Quantum mechanics timeline

The [following timeline](#) is a bit more detailed... still very brief. Pay attention however to the timescale at which these discoveries happened!

Quantum Mechanics Timeline

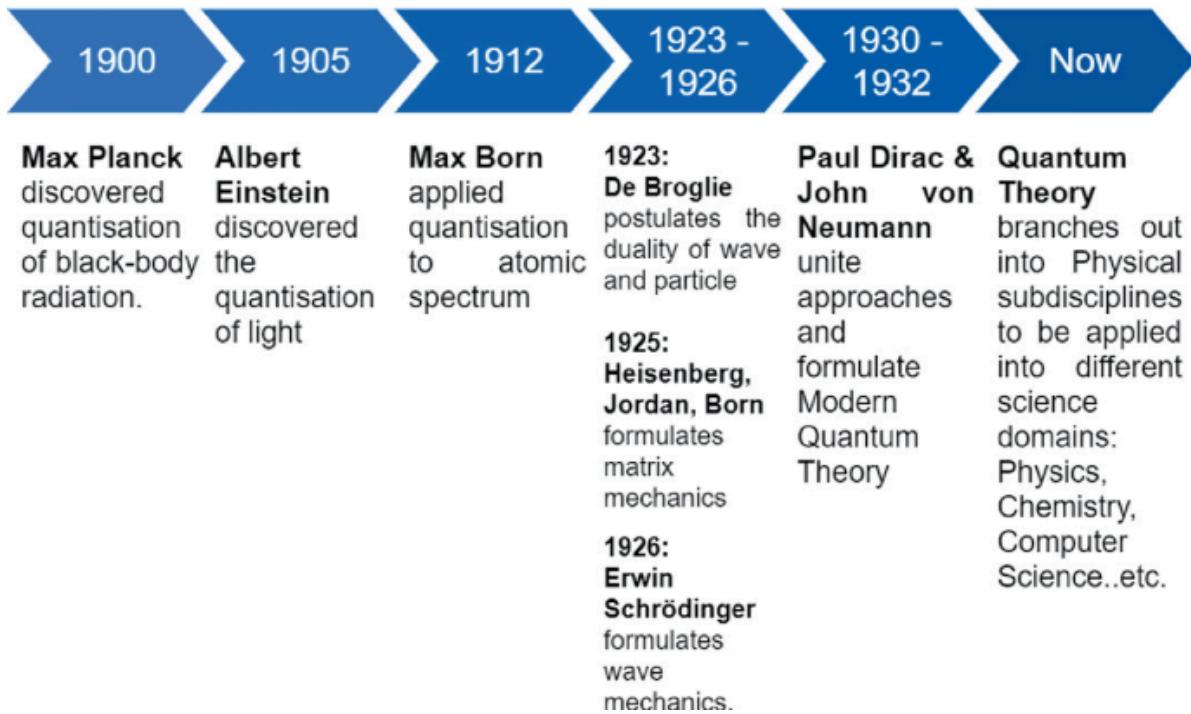


Figure 1.1: A brief timeline of Quantum Mechanics

1.1.2 Two equivalent formulations of quantum mechanics

Quantum mechanics was proposed at nearly the same time (see timeline above) using two different but equivalent formulations:

- **Wave mechanics:** proposed by Schrödinger in 1925, following the de Broglie idea about matter waves (we'll see this towards the end of this chapter). We will start with this description, that you are more familiar with, for the first part of the course.
- **Matrix mechanics:** developed by Heisenberg, Born and Jordan in 1925-1926, which uses non-commutative algebra and associates to each physical quantity and physical “observables” a matrix. This can be more convenient to describe two-level systems, the quantum harmonic oscillator or the angular momentum. We will start using this once we introduce the quantum harmonic oscillator.

These two approaches were then unified in a general formulation of quantum mechanics, developed by Dirac and Von Neumann in 1930-1932.

What led to the necessity of a quantum mechanics framework? Where and how did classical physics fail?

In the following we will discuss the historical motivations that led to the development of quantum mechanics.

1.2 Blackbody radiation and the ultraviolet catastrophe

- A body that totally absorbs all radiation that falls upon it is called a **blackbody**. You can think of it as a perfect absorber of radiation.
- A blackbody is in thermal equilibrium, so it has a temperature T .
- Since a blackbody absorbs energy and it is in thermal equilibrium, then it must also be a perfect emitter (i.e. emit across all wavelengths): a blackbody is a perfect absorber but it is also a perfect emitter.
- Emitted radiation depends only upon the radiator's temperature, and the *total emissive power*, or *total emittance*, or *spectral radiance*, of a black hole follows the **Stefan-Boltzmann law**:

$$R(T) = \sigma T^4, \quad (1.1)$$

where we defined the Stefan's constant

$$\sigma = 5.61 \times 10^{-8} \text{Wm}^{-2}\text{K}^{-4}.$$

A typical classic model for a perfect blackbody is a black cavity with a small hole (see Figure 1.2): all the light entering is reflected multiple times across the black walls and is (almost completely) absorbed. Since the cavity is in thermal equilibrium, the emitted radiation depends only on its temperature, so the cavity emits like a black body.

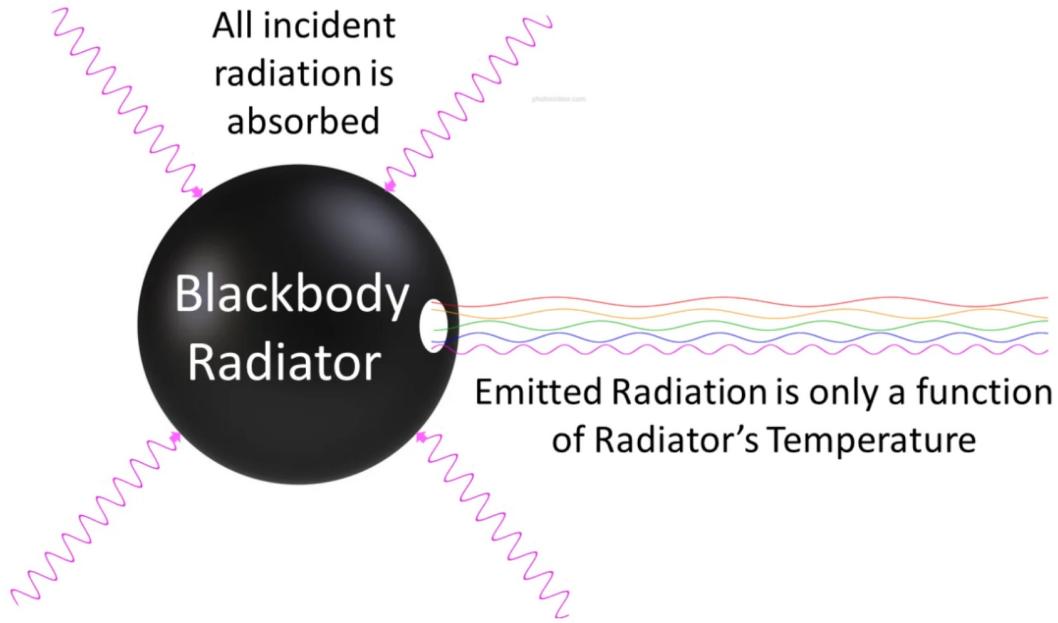


Figure 1.2: A perfect blackbody radiator.

Equation 1.1 describes the *total* emittance, i.e. the emission power per unit area at all wavelengths, but what is the distribution of radiation across wavelengths (or frequencies) that a black body emits? In mathematical terms, we previously found a form for $R(T)$, but want to know the *emittance at a given wavelength* $R(\lambda, T)$, and the total emittance is obviously obtained by integrating this over all the wavelengths:

$$R(T) = \int_0^{\infty} R(\lambda, T) d\lambda. \quad (1.2)$$

i Note

The emission spectra of black bodies (the energy emitted against the wavelength or frequency) depend only on their temperature T .

In the following we will also refer to the wavelength (or frequency) *monochromatic energy density* $u(\lambda, T)$ (or $u(\nu, T)$). The relation between the spectral emittance and the energy

density can be shown [5] to be

$$u(\lambda, T) = \frac{4}{c} R(\lambda, T), \quad (1.3)$$

so from the measurement of the emission spectrum one could determine the energy density.

So what is the form of the emission spectrum of a black body? Although the spectrum of black-body radiation was *known observationally* (how?!), its theoretical description has puzzled physicists at the end of the 19th Century and many attempts to describe this ran into trouble. It cannot be fully explained using classical physics, it requires quantum physics, as we will see in the following.

The black-body spectrum... observed? In the 1890s?!

You might be wondering how back in the days (1890-ish) physicists already knew the shape of the EM power spectrum... if you are, that's a good question and you're not alone. For this, let's thank Samuel Langley and his *bolometer*.

Ohhhh...
 Langley invented the bolometer,
 A very fine sort of thermometer.
 It can measure the heat,
 Of a polar bear's feet
 At the distance of half a kilometer.

[via J.M. Pasachoff, Contemporary Astronomy (Saunders College Publishers, Philadelphia, 1985)]

The rest of Langley's work/life story is also worth checking out, ranging from effectively inventing standard time to pioneering powered flight!

1.2.1 Wien's law (1893)

From general thermodynamical arguments, in 1893 Wien showed that the black-body energy density had to take the form

$$u_\lambda(\lambda, T) = \frac{dE}{d\lambda} = \lambda^{-5} f(\lambda T), \quad (1.4)$$

or in terms of the frequency:

$$u_\nu(\nu, T) = \frac{dE}{d\nu} = \nu^3 g(\nu/T), \quad (1.5)$$

where $f(\lambda T)$ and $g(\nu/T)$ are universal functions that depend solely upon λT and ν/T respectively and cannot be derived thermodynamically, so they have to be found empirically.

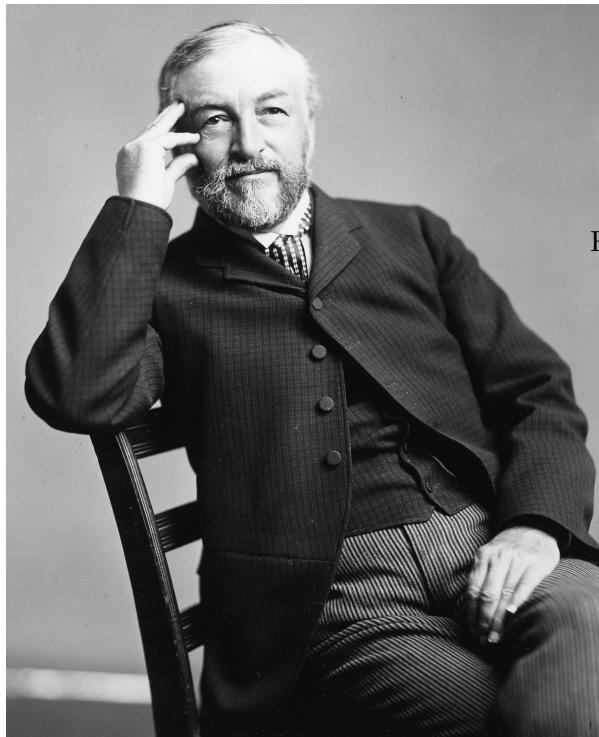


Figure 1.3: Samuel Langley, inventor of the bolometer.

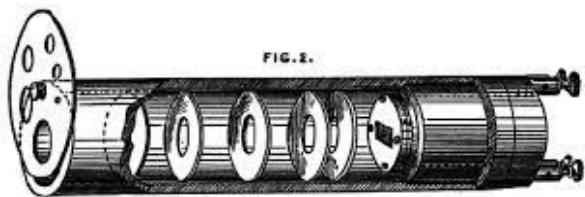


Figure 1.4: Scheme and illustration of the bolometer.

Wien's empirical guess was that $f(\lambda T)$ in Equation 1.5 had the form

$$f(\lambda T) = C \exp\left(-\frac{A}{\lambda T}\right).$$

Equation 1.5 is called Wien's law, also referred to as Wien's approximation or Wien's model (not to be confused with 'Wien's displacement law'!), and it is a good model for the observed spectrum at short wavelength (high frequencies) $hc \gg k_B T \lambda$, but it is not accurate at high wavelengths.

In Equation 1.3 we showed the relation between the spectral emittance and the energy density, from which Wien derived the explicit form of **Wien's approximation** for the spectral emittance (after attempting to derive the function $f(\lambda T)$). This is shown in Equation 1.6 and an example is plotted in Figure 1.5.

Note that Wien's law did not initially contain fundamental constants, which were later introduced by Max Planck.

Wien's approximation for the spectral emittance

$$R_\lambda^{\text{Wien}}(\lambda, T) = \frac{2hc^2}{\lambda^5} \exp\left(-\frac{hc}{\lambda k_B T}\right) \quad (1.6)$$

```
import warnings
warnings.filterwarnings('ignore')
import numpy as np
from astropy.constants import k_B, c, h
from matplotlib import pyplot as plt
from matplotlib import rc
rc('text', usetex=True)
rc('font', size=15)
rc('figure', figsize=(8,6), dpi=300)

l = np.linspace(0,3000,1000) # Wavelength, nm
T = 4000 # Temperature, Kelvin
T2 = 3000 # Temperature, Kelvin

B_wien=2 * h.value *c.value**2 * np.exp(-h.value*c.value/((l*1e-9)*k_B.value*T)) / (l*1e-9)
B_wien2 = 2 * h.value *c.value**2 * np.exp(-h.value*c.value/((l*1e-9)*k_B.value*T2)) / (l*1e-9)

plt.plot(l, B_wien, label="T=4000K")
plt.plot(l, B_wien2, label="T=3000K")
```

```

plt.xlabel(r'Wavelength $\lambda$ (nm)')
plt.ylabel(r'Spectral Radiance $R(\lambda, T)$ ($Wm^{-3}sr^{-1}$)')
plt.title('Spectral radiance - Wien model')
plt.ylim([0,5e12])
plt.grid()
plt.legend();

```

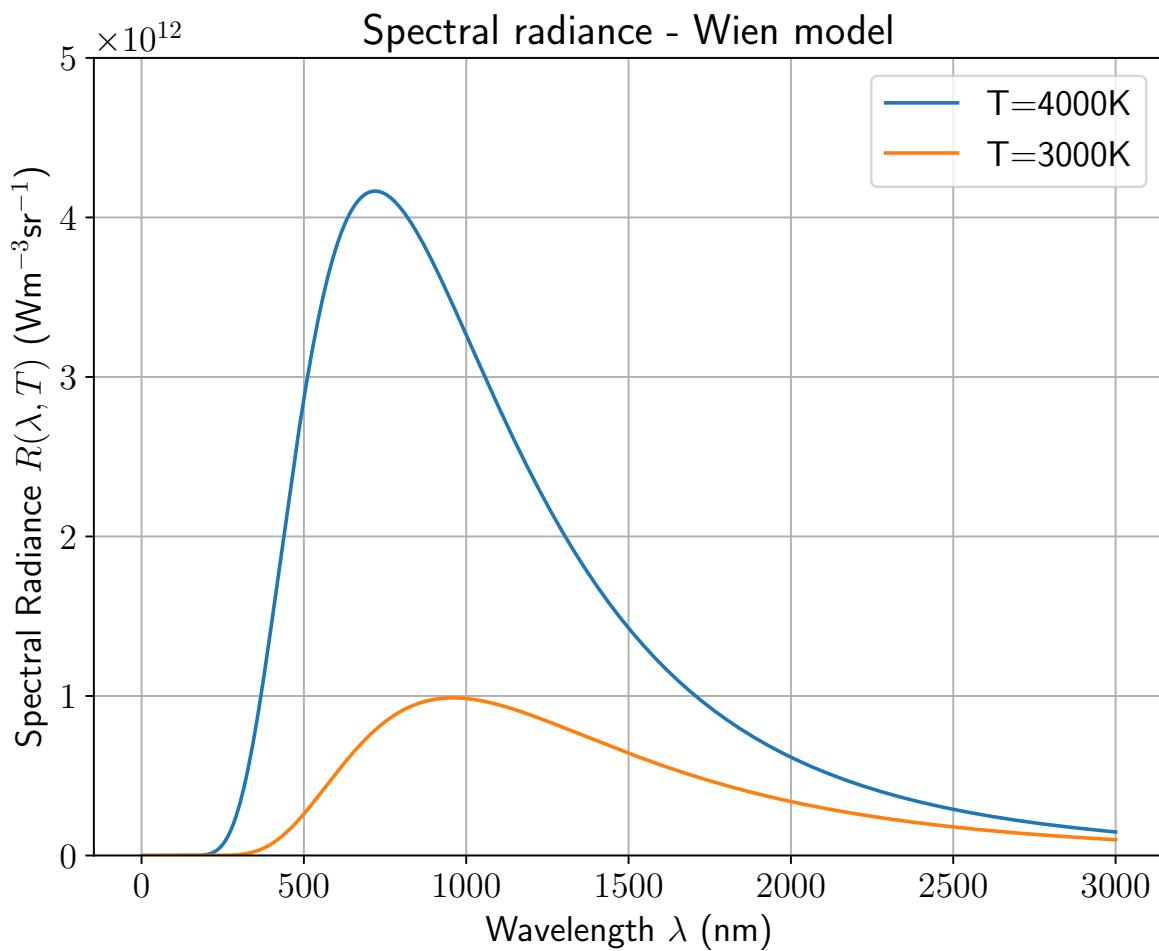


Figure 1.5: Spectral radiance plotted using the Wien's approximation (Equation 1.6), for $T=4000\text{K}$ and $T=3000\text{K}$.

Problems with the Wien's approximation:

- This model works well at small wavelengths but it is not very good at high wavelengths
- low frequencies.

- Thermodynamical reasoning is not sufficient to derive an accurate model.

1.2.2 Wien's displacement law

From Figure 1.5 we see that this model shows peak of the spectral emission at different wavelengths that depend on the temperature of the black body. The relationship between these two quantities is given by the Wien's displacement law (Equation 1.7).

Wien's displacement law

$$\lambda_{max} = \frac{b}{T} \quad (1.7)$$

λ_{max} - value of the wavelength corresponding to the peak

$b \approx 2.8978 \times 10^{-3} \text{ m} \cdot \text{K}$ - Wien's displacement constant

T - absolute temperature.

Wien's displacement law derivation

Demonstrate qualitatively the Wien's displacement law relation between the maximum wavelength and the temperature, showing that this can be obtained from Wien's model of the energy density (Equation 1.5 or Equation 1.4).

💡 Hint

It may be useful to do a change of variables.

Solution - Wien's displacement law derivation

Starting from the energy density in Equation 1.5, the peak can be found by imposing $\frac{du}{d\nu} = 0$ (the procedure is equivalent if we consider Equation 1.4 and we derive it with respect to the wavelength instead).

Defining $x = \nu/T$, we obtain

$$\begin{aligned} \frac{du}{d\nu} &= \frac{d\nu^3 g(x)}{d\nu} \\ &= 3\nu^2 g(x) + \nu^3 \frac{dg(x)}{dx} \frac{dx}{d\nu} \\ &= 3\nu^2 g(x) + \frac{\nu^3}{T} \frac{dg(x)}{dx} \\ &= \nu^2 \left(3g(x) + \frac{\nu}{T} \frac{dg(x)}{dx} \right) = 0. \end{aligned}$$

Note that we obtained

$$3g(x) + x \frac{dg(x)}{dx} = 0,$$

which depends only on x and not ν or T individually. This simplifies things, as it means that the peak itself is going to be a function of this ratio ν/T only (we do not need to solve it explicitly, so do not need to know the form of $g(x)$.)

$$x\Big|_{\text{peak}} \equiv \frac{\nu}{T}\Big|_{\text{peak}} = \text{const} \implies \nu \propto T \implies \lambda_{\max} \propto 1/T$$

We cannot determine the constant b of Equation 1.7, but we can derive the relation between the wavelength at the peak and the temperature of the black body.

1.2.3 Rayleigh-Jeans model (1900)

Lord Rayleigh used classical physics to predict the shape of the black body spectrum, which was later refined by Sir James Jeans. The explanation is based on electromagnetic theory and standing waves (more details about the derivation are given in the subsection below).

The Rayleigh-Jeans law in Equation 1.8 is a good approximation to the *observed* spectrum at long wavelength (low frequencies) $hc \ll k_B T \lambda$, giving the following spectral radiance, plotted in Figure 1.8.

Rayleigh-Jeans law

$$R_{\lambda}^{\text{RJ}}(\lambda; T) = \frac{2ck_B T}{\lambda^4} \quad (1.8)$$

```
import warnings
warnings.filterwarnings('ignore')
import numpy as np
from astropy.constants import k_B, c, h
from matplotlib import pyplot as plt
from matplotlib import rc
rc('text', usetex=True)
rc('font', size=15)
rc('figure', figsize=(8,6), dpi=300)

l = np.linspace(0,3000,1000) # Wavelength, nm
T = 4000 # Temperature, Kelvin
T2 = 3000 # Temperature, Kelvin
```

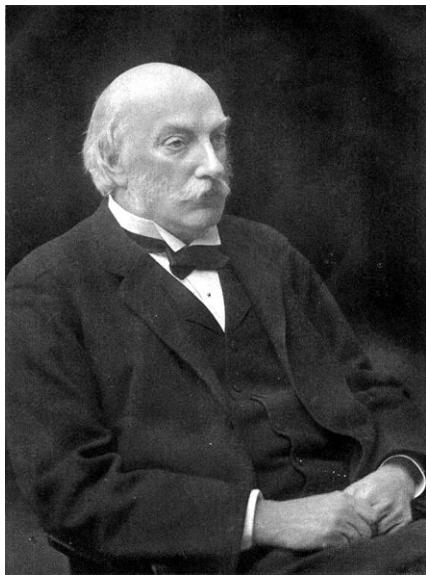


Figure 1.6: Lord Rayleigh (1842 - 1918)

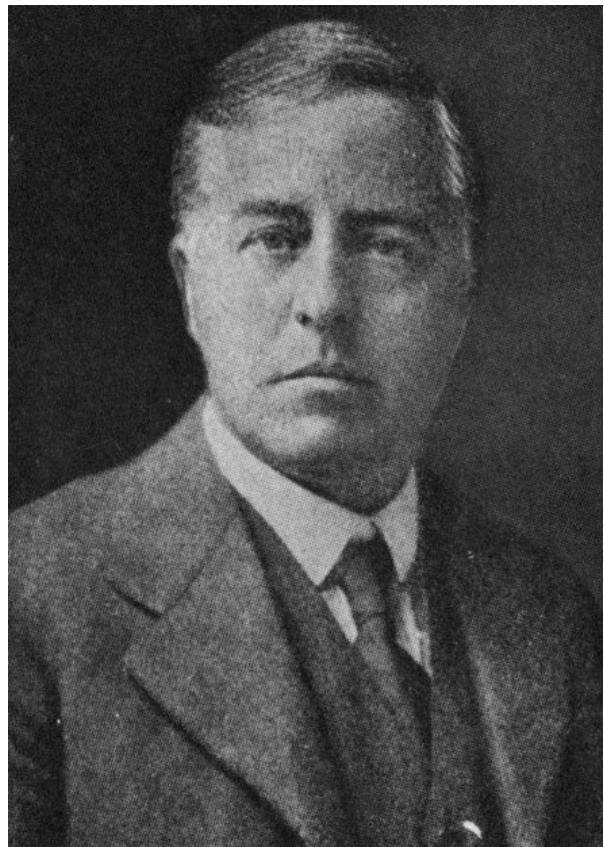


Figure 1.7: Sir James Jeans (1877 - 1946)

```

B_rj=c.value* 2*k_B.value * T /(1*1e-9)**4; # 1e-9 to convert from nm to meters
B_rj2=c.value* 2*k_B.value * T2 /(1*1e-9)**4; # 1e-9 to convert from nm to meters

plt.plot(l, B_rj, label="Rayleigh-Jeans law, T=4000K")
plt.plot(l, B_rj2, label="Rayleigh-Jeans law, T=3000K")
plt.xlabel(r'Wavelength $\lambda$ (nm)')
plt.ylabel(r'Spectral Radiance $R(\lambda, T)$ (Wm$^{-3}$sr$^{-1}$)')
plt.title('Spectral radiance')
plt.ylim([0,5e14])
plt.grid()
plt.legend();

```

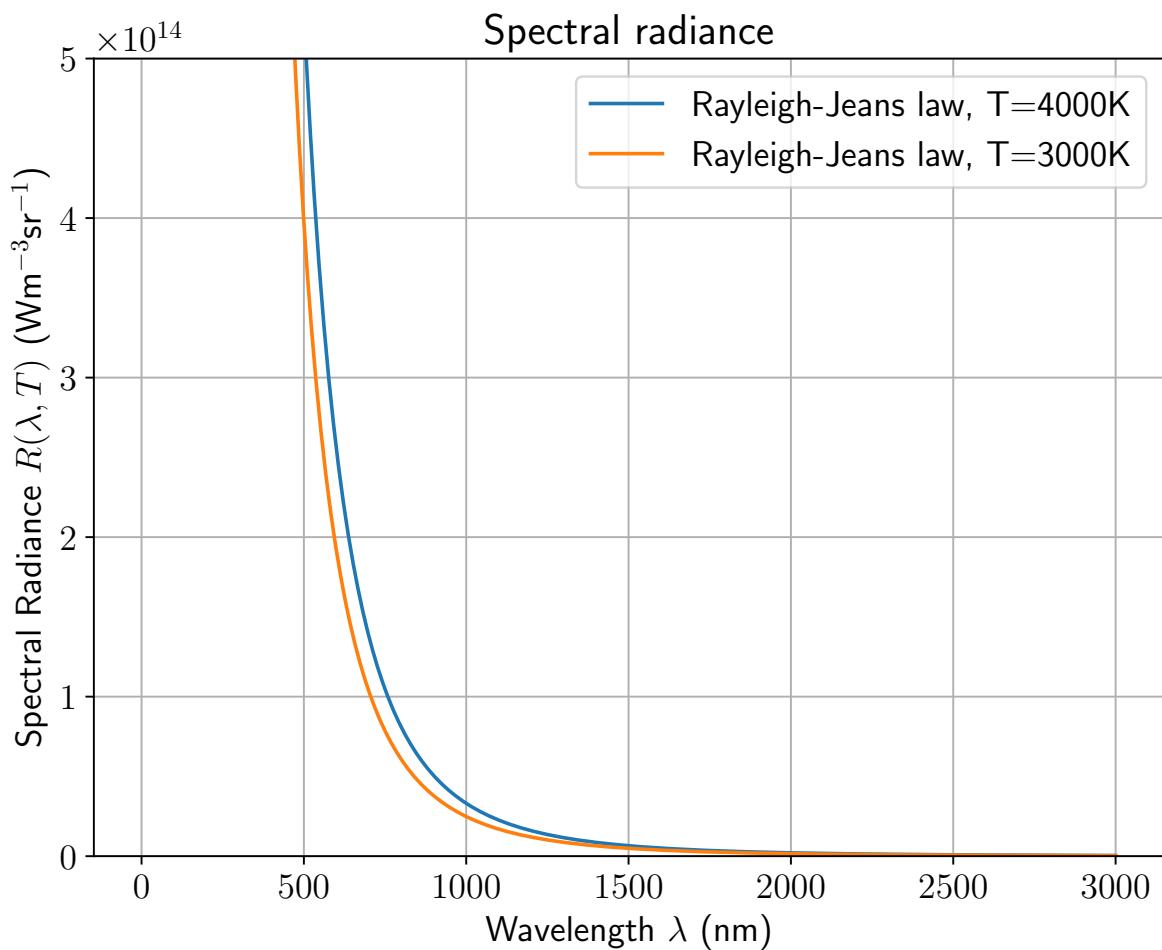


Figure 1.8: Spectral radiance plotted using the Rayleigh-Jeans law (Equation 1.8), for $T=4000\text{K}$ and $T=3000\text{K}$.

The Rayleigh-Jeans law has a big problem: the radiance keeps increasing indefinitely at short wavelengths (higher frequencies) - see Figure 1.8. If it were true, the power emitted at short wavelengths would be infinite! This is known as the **ultraviolet catastrophe**, and indicates the failure of classical physics to explain the behaviour of thermal radiation.

1.2.4 Derivation of the Rayleigh-Jeans law

The derivation can be found in [5] and follows principles from electromagnetism.

Imagine to have a 3D cubic cavity with side-length L . Inside the cavity we would expect to have standing waves (as in Figure 1.9), and the amount of radiation that the cavity can emit at a certain frequency is proportional to the number of modes it allows in that range. We are interested in the following three questions:

1. How many modes are there, per unit frequency, per unit volume, in the cavity?
2. How much energy is needed to excite each mode, and how is the energy density distributed?

1. Determining the number of modes in the cavity

To answer question 1, we proceed per steps:

- Consider the cavity and impose periodic boundary conditions (we must have zero amplitude at the walls): then the allowed modes have the form $\prod_i \sin(k_i x_i)$, with $k_i = n_i \pi L$, for integer n_i . Extending to each direction:

$$|\mathbf{n}|^2 = n_x^2 + n_y^2 + n_z^2 = (2\nu L/c)^2 = (2L/\lambda)^2. \quad (1.9)$$

- How do we find the total number of modes that satisfy this condition? We would need to count all the combinations of the integer values n_i , but we can use a trick, which consists in assuming that the number of combinations is the volume of something that we call and “n-space”: a 3D grid containing the values of n_i in each direction (Figure 1.10).

We consider the n-space as a sphere: to find the number of modes we need to find its volume (this is a good approximation since the cavity length is much larger than the wavelength), *but* we need to consider only the positive quadrant (1/8 of the volume of the sphere), *and* we can have waves polarized along two planes (multiply the volume by 2 to account for two polarizations). With all this in mind, the volume of the n-space (positive) sphere with two possible polarizations is:

$$dN(|\mathbf{n}|) = \frac{1}{8} \times 2 \times \frac{4\pi}{3} (|\mathbf{n}|^2)^{3/2} d|\mathbf{n}| \quad (1.10)$$

$$= \frac{\pi}{3} (|\mathbf{n}|^2)^{3/2} d|\mathbf{n}|. \quad (1.11)$$

From this, using Equation 1.9, we find that the *number of modes per volume* (in terms of the wavelength or the frequency) is

$$dN(\lambda) = \frac{8\pi L^3}{3\lambda^3} d\lambda, \quad dN(\nu) = \frac{8\pi\nu^3 L^3}{3c^3} d\nu \quad (1.12)$$

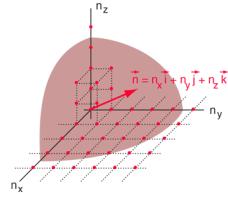
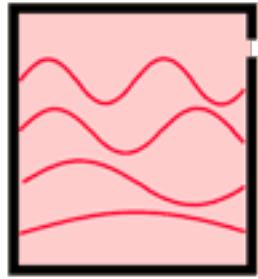


Figure 1.10: Rayleigh scheme used for counting modes. Credits: [5]

Figure 1.9: Standing waves in a cavity.

- We now want to find the *number of modes per volume, per unit frequency (or wavelength)*, so we divide by L^3 and differentiate with respect to ν (or λ), obtaining respectively:

$$\left| \frac{dN(\nu)}{d\nu} \frac{1}{L^3} \right| = \frac{8\pi\nu^2}{c^3} \quad (1.13)$$

$$\left| \frac{dN(\lambda)}{d\lambda} \frac{1}{L^3} \right| = \frac{8\pi}{\lambda^4}. \quad (1.14)$$

Note that $\frac{dN}{d\nu}$ gives a negative result because the number of modes decreases with increasing wavelength.

These results hold general validity in both classical and quantum physics, but we did not address the second and third questions yet to determine the probability of occupying the modes and their corresponding energy. This is where classical and quantum physics start diverging, and this is where this difference becomes important. The Rayleigh-Jeans model uses a classical physics approach in addressing the above questions.

2. Determining the energy associated to the modes

To find the energy associated to the modes we need to first ask: what is the probability that the modes are occupied? Based on classical physics all the modes have the *same probability* of being occupied, i.e. all modes have the same chance of being produced and their number increases according to Equation 1.13 and Equation 1.14.

According to classical (statistical thermal) physics, the principle of *equipartition of energy* states that, for a body in thermal equilibrium at temperature T , the energy associated to each

degree of freedom is $1/2k_B T$. So the energy needed to excite each mode (accounting for two polarizations) is $\bar{E} = \frac{1}{2} \times 2k_B T = k_B T$, and this *does not depend on the frequency*.

What's the result for the energy density, if we put these things together?

$$u(\nu, T) = \frac{8\pi\nu^2}{c^3} k_B T. \quad (1.15)$$

i Note

Note that this means that the function $g(x)$ introduced in Equation 1.5 is then $g(\nu/T) = \frac{k_B T}{\nu} = \frac{1}{x}$.

If we integrate Equation 1.15 over all the frequencies to find the total energy,

$$E_{tot} = \int_0^\infty u(\nu, T) d\nu = \infty$$

we obtain *infinite energy* (not physically possible) - this is the **ultraviolet catastrophe**.

1.2.5 Planck's Law (1900)

We saw that Wien's approximation was good at low wavelengths (high frequencies), while the Rayleigh-Jeans law works better in the opposite regime of high wavelengths (low frequencies).

In 1900 Planck was able to derive a spectrum that fits the observed data of the blackbody spectral emission by assuming that energy comes in discrete quanta of energy

$$E = h\nu = hc/\lambda, \quad (1.16)$$

with $h = 6.626 \times 10^{-34} \text{ J} \cdot \text{Hz}^{-1} = 4.136 \times 10^{-15} \text{ eV} \cdot \text{Hz}^{-1}$ being Planck constant.

This marked the birth of quantum physics! And believe it or not, Planck's idea started from a lightbulb (see [video](#) below).

<https://www.youtube.com/watch?v=i1TVZIBj7UA>

The Planck radiation spectrum for a black body's spectral radiance is:

Planck's law for the spectral emittance

$$R_\lambda(\lambda; T) = \frac{2hc^2}{\lambda^5} \left(\frac{1}{\exp\left(\frac{hc}{\lambda k_B T}\right) - 1} \right) \quad (1.17)$$

The Wien's approximation, the Rayleigh-Jeans law and Planck's law are compared as a function of wavelength in Figure 1.11 and as a function of frequency in Figure 1.12.

```
B_planck = 2*h.value*c.value**2 * (1.0/(np.exp(h.value*c.value/(k_B.value*T*(1*1e-9)))-1))
plt.plot(l, B_planck, label="Planck's law")
plt.plot(l, B_wien, label="Wien's approximation")
plt.plot(l, B_rj, label="Rayleigh-Jeans law")

plt.xlabel(r'Wavelength $\lambda$ (nm)')
plt.ylabel(r'Spectral Radiance $B_\lambda(\lambda, T=4000K)$ (Wm$^{-3}$sr$^{-1}$)')
plt.title('Spectral radiance')
plt.ylim([0,5e12])
plt.grid()
plt.legend();
```

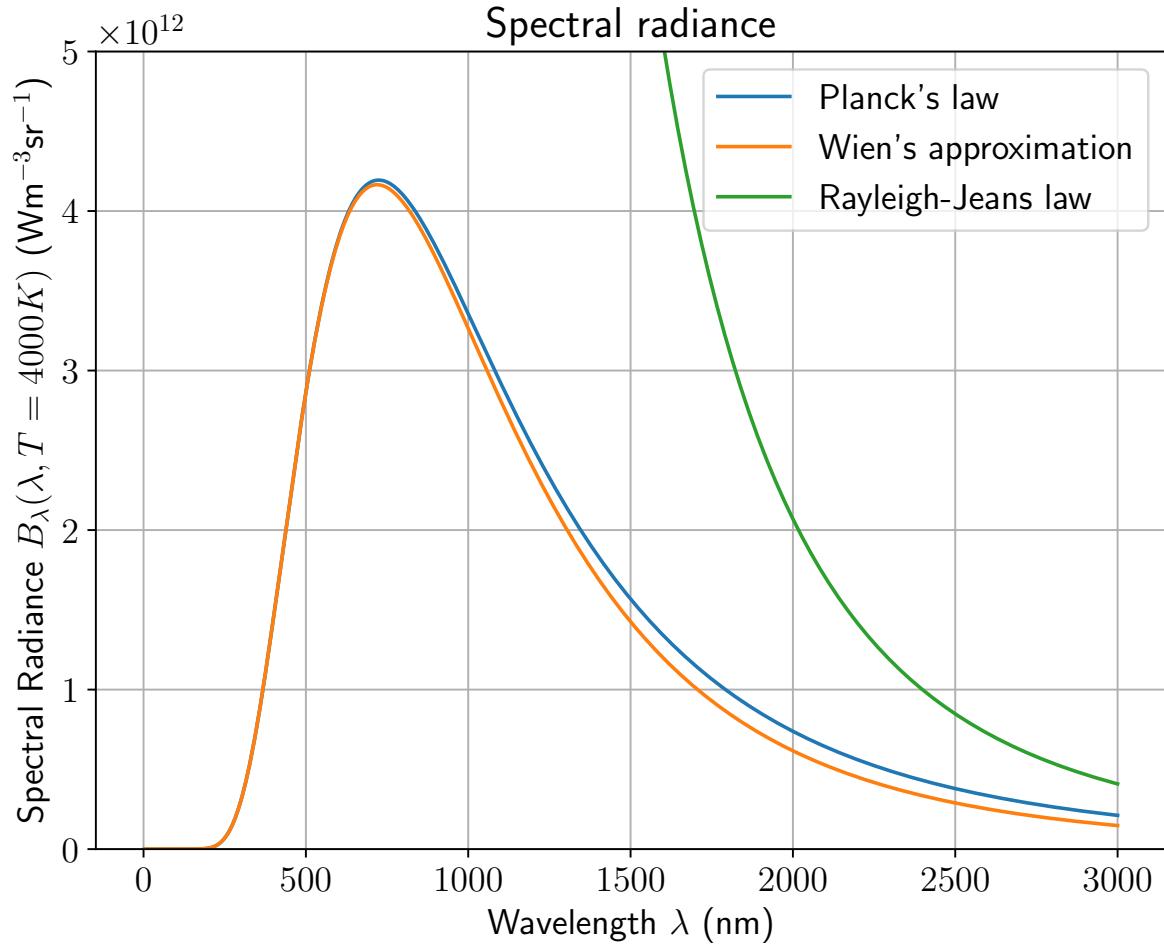


Figure 1.11: Comparison of the spectral radiance plotted using the Rayleigh-Jeans law (Equation 1.8), Wien’s approximation (Equation 1.6) and Planck’s law (Equation 1.17) for $T=4000\text{K}$.

Interlude: Max Planck (1858-1947)

Personal life:

- Music prodigy on piano, organ and cello, but chose physics. Later played regularly with Einstein in Berlin.
- Advised against physics: “Almost everything already discovered, and all that remains is to fill a few holes” – Planck “did not wish to discover new things” (!)
- Academic career in Germany, ended up as theoretical physics prof in Berlin, specialising in thermodynamics & entropy

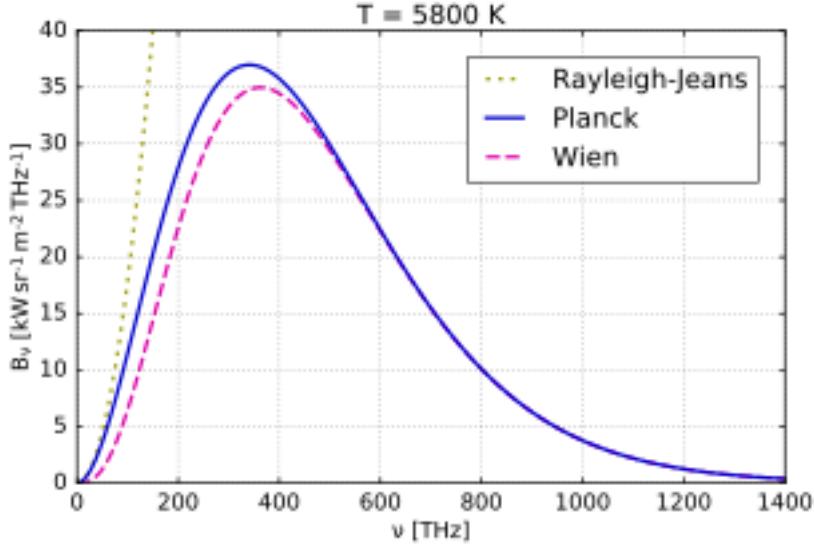


Figure 1.12: Comparison of the spectral radiance plotted using the Rayleigh-Jeans law (Equation 1.8), Wien's approximation (Equation 1.6) and Planck's law (Equation 1.17), as a function of frequency, for $T=5800\text{K}$.

- Planck was patriotic and remained in Germany during the Nazi era, despite opposing some government policies, especially the persecution of Jews.
- His home was destroyed by bombing near the end of World War II, leading to significant hardship
- He was married twice, first to Marie Merck, and after her death, to her cousin, Marga von Hösslin.
- He faced personal tragedy, losing three of his children, including one son, Erwin, who was executed by the Gestapo for unsuccessfully attempting to assassinate Hitler in 1944.

Physics considerations and contributions:

- Grand figure of German physics, Nobel Prize, father of DPG and Solvay conferences. He resisted Nazi takeover of DPG and helped Jewish physicists. Died shortly after his son Erwin was executed.
- Rayleigh and Jeans kept setting $h \rightarrow 0$: “I amunable to understand Jeans' stubbornness – he is an example of a theoretician as should never be existing. So much the worse for the facts if they don't fit.”

1.2.5.1 Derivation of the Planck's model

What are the ‘classical’ assumptions of the Rayleigh-Jeans model that were dropped by Planck?

In the previous section on the derivation of the Rayleigh-Jeans law we determined the number of modes and the energy associated to those. The number of modes counting was correct, but the mistake is in assuming that the equipartition of energy leads to each mode having an energy of $k_B T$, and that exciting each of these modes is equally probable.

According to the classical theory, the average energy per mode can be obtained starting from the **Maxwell-Boltzmann distribution**

$$p(E) = e^{-E/k_B T}, \quad (1.18)$$

and using the classical *equipartition of energy*:

$$\langle E \rangle = \frac{\int E p(E) dE}{\int p(E) dE} = \frac{k_B^2 T^2}{k_B T} = k_B T \quad (1.19)$$

This result is the one we previously determined in Section 1.2.4, however the integration relies in a continuum distribution of energies.

Planck however used the discretisation of the energy levels to derive (in 2 months!) the correct form of the spectral radiance.

Using discretisation of the energy E into discrete energy steps $\epsilon = h\nu$, from Equation 1.19 we obtain

$$\langle E \rangle = \frac{\sum_{n=0}^{\infty} n \epsilon e^{-n\epsilon/k_B T}}{\sum_{n=0}^{\infty} e^{-n\epsilon/k_B T}} = \frac{\epsilon}{e^{\epsilon/k_B T} - 1}. \quad (1.20)$$

1.2.6 Exercise

Demonstrate the result of Equation 1.20.

 Hint

You can use the result of the geometric series $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$ and its consequence $\sum_{n=0}^{\infty} (n+1)x^n = \frac{1}{(1-x)^2}$.

1.2.6.1 Consequences of Planck's model and comparison to the Rayleigh-Jeans model

Planck's introduction of discrete quanta of energy therefore affected the probability of occupation of the modes and their energy, as follows:

- In classical physics the occupation of each mode was *equally possible*, but since modes are quantised and each quanta has energy $E = h\nu$, exciting *higher modes is less probable* because it requires more energy. The probability that a mode will be occupied is given by the **Bose-Einstein distribution** function:

$$p_\nu = \frac{1}{e^{h\nu/k_B T} - 1} \quad (1.21)$$

Note that this distribution was introduced only in 1924; Planck's law can be re-derived using this from quantum statistics arguments.

- The classical theory stated that each mode needed equal energy of $k_B T$ to be excited, in accordance with the *equipartition of energy*. However, the average energy per "mode" (or "quantum") is given by its energy ($h\nu$) times the probability that this will be occupied (the Bose-Einstein distribution function of Equation 1.21):

$$\langle E \rangle = \frac{h\nu}{e^{h\nu/k_B T} - 1}, \quad (1.22)$$

which is now *dependent on the frequency*.

The energy density is then given by the number of modes, per unit volume, per unit frequency (Equation 1.13), times the average energy per mode of Equation 1.22:

Planck spectral distribution law

$$u(\nu, T) = \frac{8\pi\nu^2}{c^3} k_B T \left(\frac{h\nu/k_B T}{e^{h\nu/k_B T} - 1} \right) = \frac{8\pi h}{c^3} k_B T \left(\frac{\nu^3}{e^{h\nu/k_B T} - 1} \right). \quad (1.23)$$

1.2.6.2 Understanding the Planck's distribution

- Compare Equation 1.23 with Equation 1.15: you can see that the term related to equipartition energy has now a modification factor which reduces to the mode-counting form at low ν and returns Wien's approximation at high ν (demonstrate it in the problems below).
- Note that the $n = 0$ mode does not contribute to the value of $\langle E \rangle$ in Equation 1.20, while in the smooth equipartition integrap in is a dominant term: this is why the growth of energy density at high frequencies is suppressed!
- The first non-zero contribution starts at $E = \epsilon = h\nu$, which increases with frequency.
- For small $h\nu/k_B T$ (upper plot of Figure 1.13), the discrete steps are small enough that the distribution is pseudo-continuous: recover classical Rayleigh-Jeans form (show it in the problems below).
- At large $h\nu/k_B T$, the first non-zero mode is highly Boltzmann-suppressed: finite energy in the cavity \rightarrow solves the ultraviolet catastrophe!
- But... is the energy fundamentally quantized, or just emitted that way?

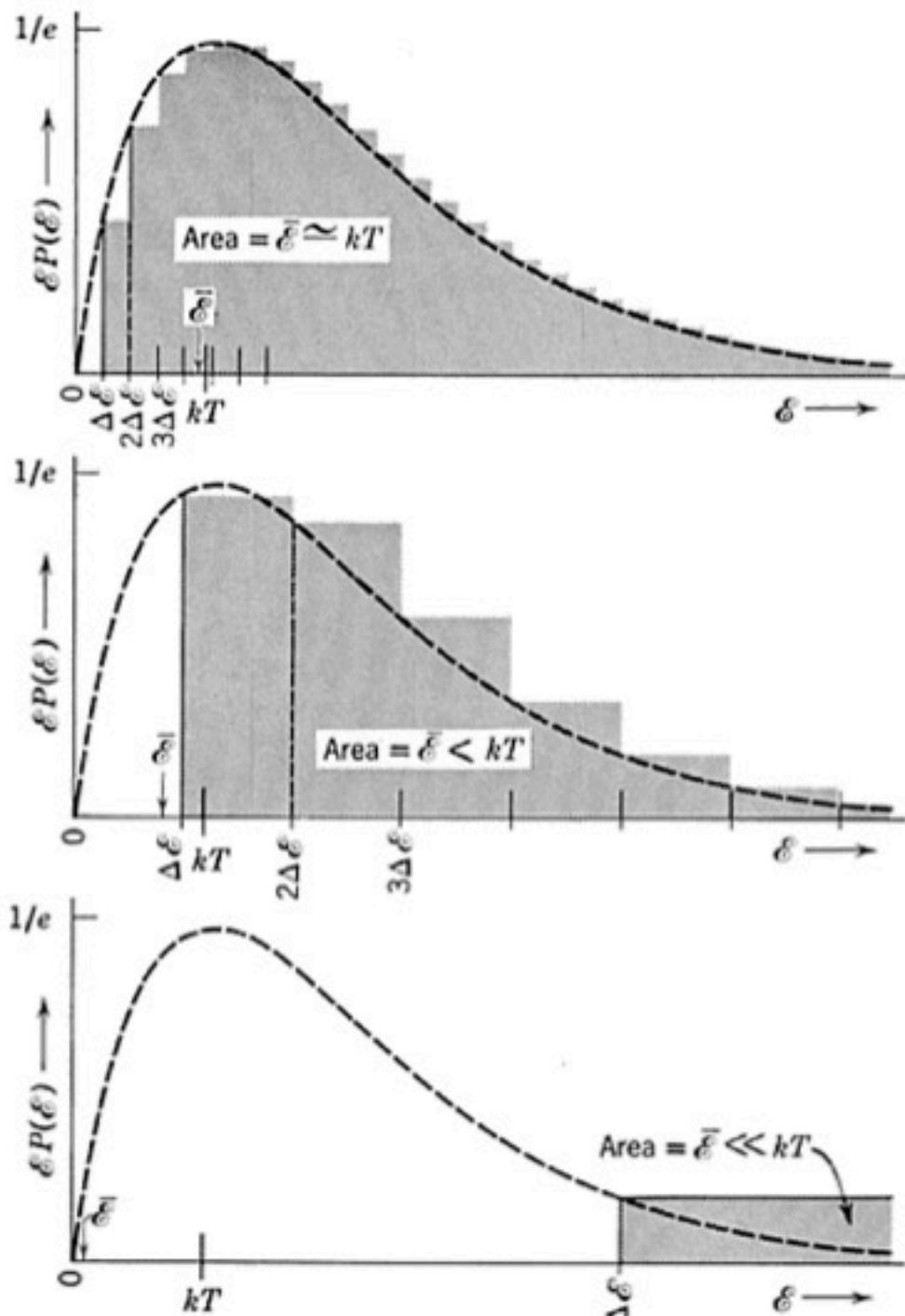


Figure 1.13: Planck's distribution, with explicit discretisation (grey steps). The energy step is $\Delta\epsilon = h\nu$.

1.2.7 Exercise

1. Show that for low frequencies ($hc \ll \lambda k_B T$) the Rayleigh-Jeans law can be derived from the Planck law.
2. Show that for high frequencies ($hc \gg \lambda k_B T$) Wien's law can be derived from the Planck law.
3. Show that Stefan's law can be found by integrating Equation 1.17
4. Find Wien's displacement law (and the Wien's displacement constant) using the Planck law.

💡 Hint

You can use the approximation $\lim_{x \rightarrow 0} \exp(x) = 1 + x$

1.2.8 Blackbody radiation and the cosmic microwave background

The cosmic microwave background (CMB) is the aftermath of the Big Bang cosmological model, discovered by accident in 1964 by Penzias and Wilson. Using a radio telescope, they detected an isotropic radio ‘noise’ incident on Earth from all directions with the same intensity and filling the universe uniformly. The intensity was then measured at other wavelengths and the spectral distribution was observed to be consistent with a near-perfect 2.7 K with $\sigma(T) \sim 20\mu K$ and a peak of emission in the microwave region (see Figure 1.14).

Research conducted with the Far Infrared Absolute Spectrophotometer (FIRAS) on the COBE satellite measured a temperature of 2.725 ± 0.002 K. Earlier experiments had detected some anisotropy in the background radiation due to the motion of the solar system, but COBE’s data revealed fluctuations in the radiation (Figure 1.15).

Later, the WMAP mission produced a higher-resolution map of these anisotropies in the cosmic background radiation. The Planck satellite has since further refined this mapping, offering the most precise measurements of the key cosmological parameters to date.

Why do we observe these temperature fluctuations? At the very beginning, during the Big Bang, the Universe experienced an extremely rapid period of inflation that lasted only a tiny fraction of a second—around 10^{-32} s. By the end of this inflationary phase, the Universe had expanded by a staggering factor of about 10^{30} (that’s *a lot!*).

At the beginning of the inflation and expansion, random quantum fluctuations were present, what happened to those? They were stretched to enormous, cosmological scales during this process! As a result of these quantum fluctuations, the post-inflation universe now consists of vast regions with slightly different properties from one another, with certain areas having a slightly higher density of matter than others. Such fluctuations are therefore necessary in Big Bang cosmology to create the non-uniformities required for galaxy formation. Quantum fluctuations are thought to have seeded large-scale structure in the universe!

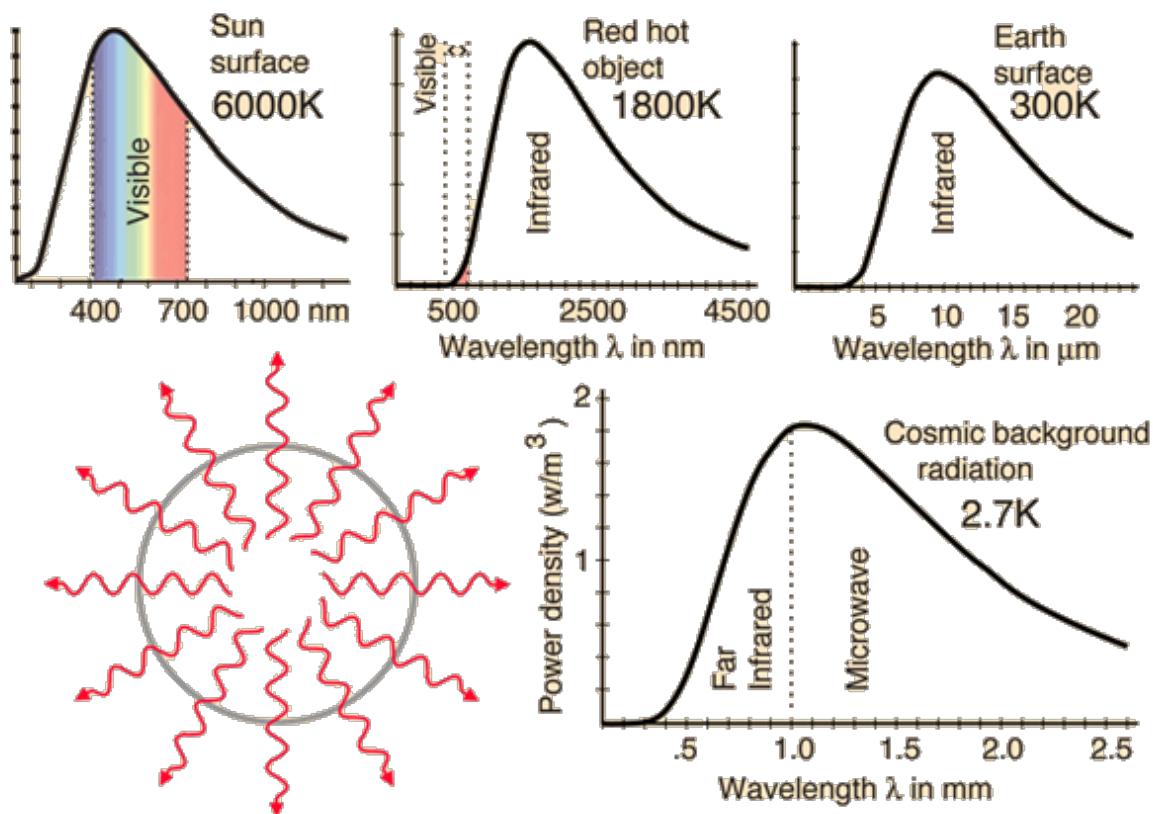


Figure 1.14: Based on their temperatures, different bodies emit at different ranges of frequency. The cosmic microwave background emits in the far infrared/microwave with a Wien's peak in the microwave.

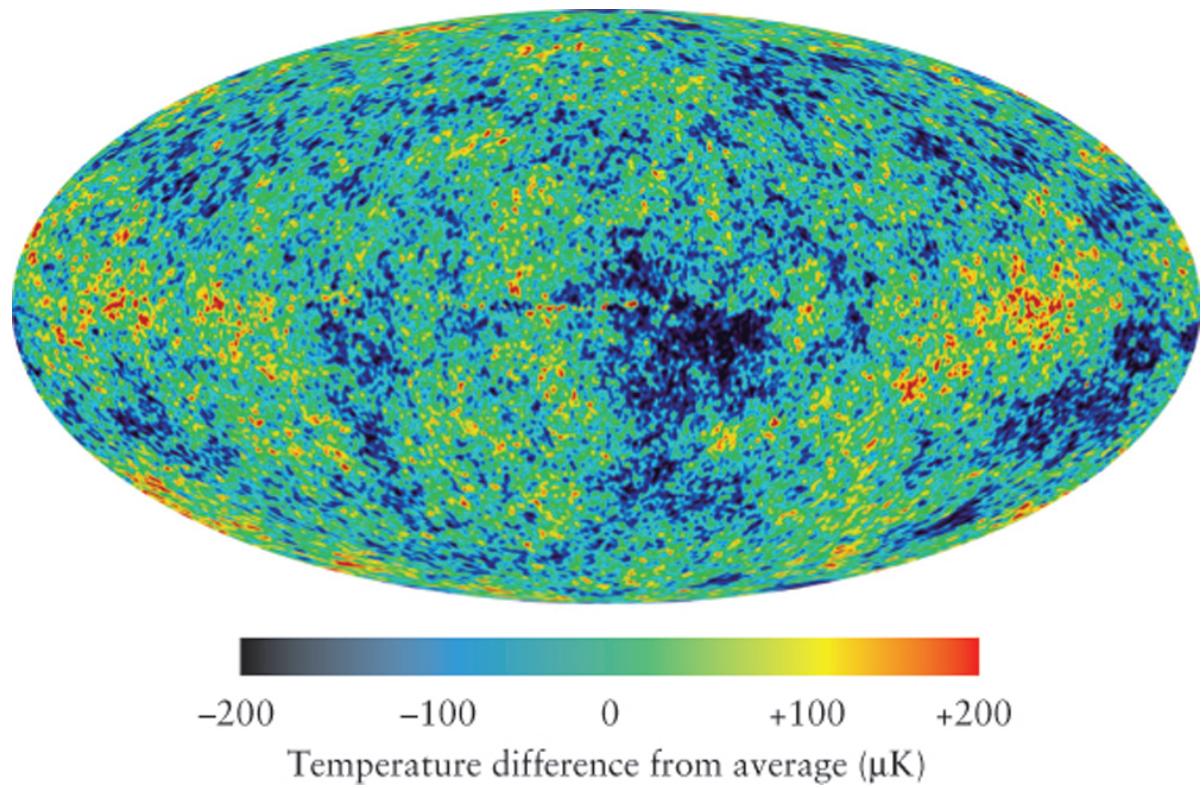


Figure 1.15: Cosmic microwave background, showing the fluctuations around the average temperature of 2.74K.

Ok, but how are the density fluctuations and the thermal fluctuations related?

Before the cosmic microwave background (CMB) was released, photons and particles were tightly linked, forming a unified ‘fluid’ of matter and radiation. At the time of recombination, around 380,000 years after the Big Bang, the two decoupled, allowing photons to travel freely across the Universe. These photons carried with them a record of how matter and radiation were distributed at that moment. Photons that came from denser regions lost more energy (since they needed to give away more energy to escape from a higher gravitational attraction), making them cooler, while those from less dense regions lost less energy and appeared warmer. As a result, the temperature fluctuations in the CMB reflect the density fluctuations in the early Universe, offering a ‘snapshot’ of its structure at the time of recombination.

Let’s now go back to discussing how quantum mechanics was discovered. The ultraviolet catastrophe was one of the pieces of the puzzle, but not the only thing that could not be explained without introducing the concept of quantisation of energy.

1.3 Photoelectric effect

The photoelectric effect was discovered by Heinrich Rudolf Hertz in 1887 during his work on radio waves. He observed that ultraviolet light affected the voltage at which sparking occurred between metal electrodes. Philipp Lenard clarified this in 1902, showing that light striking a metal surface releases electrically charged particles (later identified as electrons), discovered by J.J. Thomson in 1897. The photoelectric effect was then fully explained by Einstein in 1905, who got the Nobel prize for this (not for the relativistic theory!).

To explain this, let’s introduce the setup: a vacuum chamber containing a metallic plate (it can be of different materials, e.g. sodium) hit by radiation at a certain wavelength and intensity, and a collector on the opposite side. Both sides are connected to a battery that can apply a voltage difference, and the collector is connected to an ammeter that measures the current of the charged particles hitting the collector. Use the following [simulation](#) (embedded below in the html file) and note what you observe when varying the following quantities:

- Wavelength of incident light
- Intensity of the incident light
- Voltage of the battery
- The material of the metallic plate

It is advisable to change only one parameter per time at first to better determine its effects. Show the relevant graphs on the simulator after varying the corresponding quantities indicated on the axis.

What should we expect classically? Classically, if we treat light as a wave, we know that its energy depends on the intensity, and therefore on the amplitude of the EM wave. The expectation is therefore that the energy of an electron emitted does not depend on the

frequency of the light, but on its intensity. Furthermore, classically there is no reason why electrons should not be emitted if enough energy (intensity) is sent to the metallic plate, this should not depend on the frequency. What do we observe instead?

What do we observe? **Simulated experiment 1: changing the material and the wavelength** Let's start with sodium, setting the voltage to zero and the intensity to 50%, and let's change the wavelength from higher to lower (from 850nm, in the IR, to the UV), what do we notice?

- Nothing happens until the wavelength reaches a certain threshold ~ 530 nm, where then electrons start being emitted by the metallic plate hit by the light
- Below the threshold frequency (above the threshold wavelength) we do not see emitted electrons even if we change the intensity
- If we change the material, the critical frequency changes
- The more we decrease the wavelength (increase the frequency) the faster the electrons move - this means they have more kinetic energy

Simulated experiment 2: changing the intensity and the voltage Let's now fix a wavelength at which electrons are released and change first the intensity and then the voltage to see how the current is affected.

- Let's start by setting a null voltage. We can see that the current increases with increasing intensity, and in fact we can observe that more electrons are released and all of them reach the collector at higher intensities.
- **Positive voltage:** Let's now ramp up the voltage to the maximum positive value: we see that now the electrons are moving faster because of the potential, and the system reaches a saturation current
- **Negative voltage:** If the value of the voltage is then decreased until it becomes negative, however, we can see that increasingly higher proportions of electrons start travelling towards the collector but come back, as now the potential is opposed to the initial motion of the electrons. Since the intensity changes the number of electrons emitted and therefore the photocurrent, at a fixed negative voltage the photocurrent will decrease with decreasing intensity. However, *regardless of light intensity*, we can eventually reduce the photocurrent to 0 with a negative **stopping voltage** V_0 (see Figure 1.16). The stopping voltage decreases linearly with the light frequency, until it reaches 0: there is no photocurrent below a threshold frequency.

Why we need a quantum explanation Classical physics does not explain why no photocurrent is recorded above the threshold wavelength: it should depend on the amplitude and therefore on the intensity, according to classical waves physics.

We need the quantisation of energy, as Einstein proposed, for this.

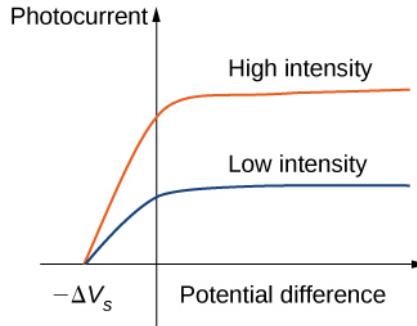


Figure 1.16: Photocurrent vs the potential difference, for two different values of the intensity and the same frequency. The stopping potential here is indicated with ΔV_s . For positive values of the potential difference, the photocurrent saturates for both intensities, while for negative values the stopping potential is eventually reached, regardless of the intensity.

- Electrons are emitted when a work of at least the *binding energy* E_b is paid to free them: this value depends on the material of the metallic plate. If we consider that the energy coming from light is quantised and follows Planck's definition of quanta of energy $E = h\nu$, then the minimum work W that needs to be done to emit electrons is related to the kinetic energy of the electrons through the Einstein's photoelectric equation:

$$E_b = h\nu - W, \quad (1.24)$$

and the minimum (threshold) frequency necessary to eject electrons is $\nu_{th} = W/h$.

The extra energy results in kinetic energy, which determines a maximum possible velocity (see Figure 1.17)

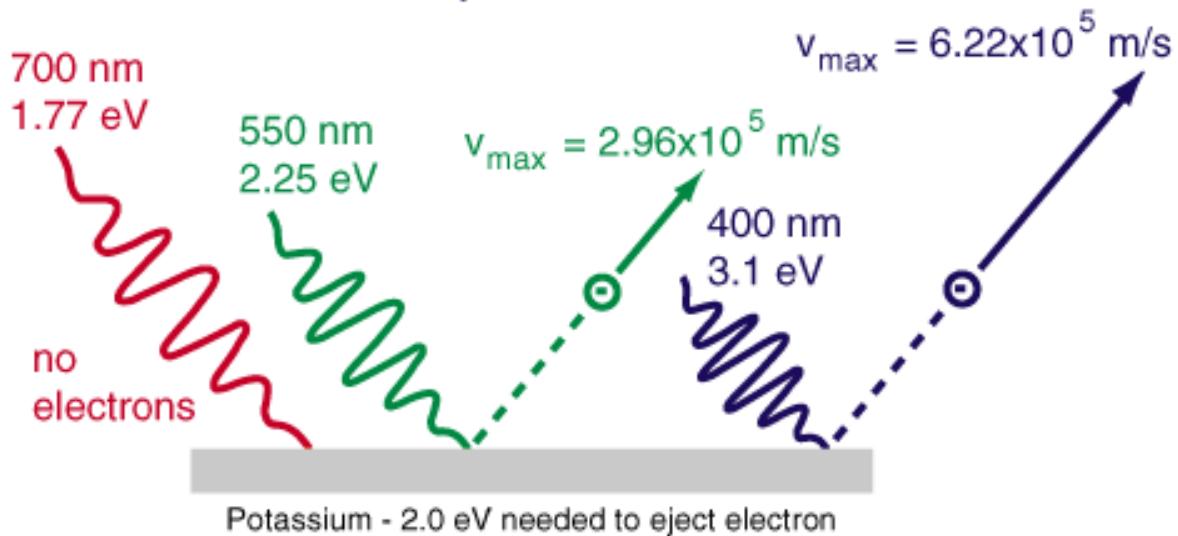
- The intensity determines the number of photons present in the light sent to the metallic plate, so higher intensity means that more electron can be extracted, if the frequency is above the threshold, and the photocurrent increases
- The voltage can be tuned to increase the photocurrent by facilitating the motion of electrons increasing the voltage. A negative voltage can otherwise invert the motion of the electrons, and consequently reduce the photocurrent all the way to zero (see Figure 1.18).

1.4 The nuclear atom

The description of the structure of the atom has evolved according to the timeline in Figure 1.19.

Where did quantum mechanics come into play in the structure of the atom?

$$E_{\text{photon}} = h\nu$$



Photoelectric effect

Figure 1.17: Some observations of the photoelectric effect. Electrons are ejected only below a certain wavelength, and the lower the wavelength, the higher is the kinetic energy (max velocity).

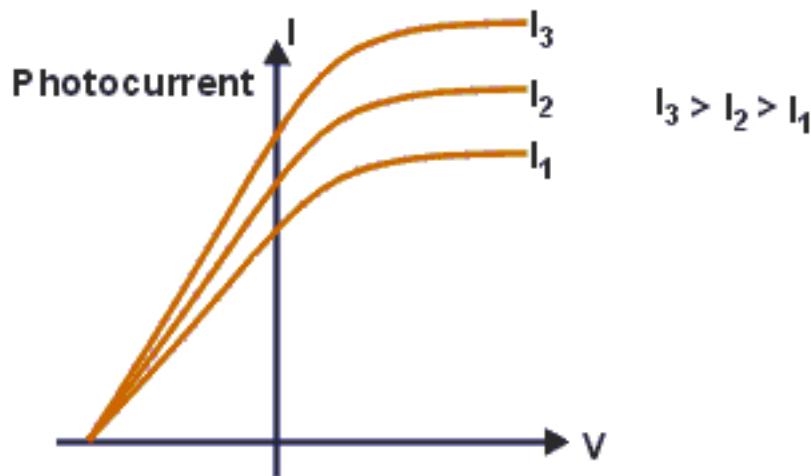


Figure 1.18: Photocurrent vs the voltage for different values of the intensity. For each intensity, increasing the voltage from 0 leads to an increase of intensity up to a plateau, where all the electrons emitted get to the collector. When the voltage is negative, the more it gets smaller, the less is the photocurrent.

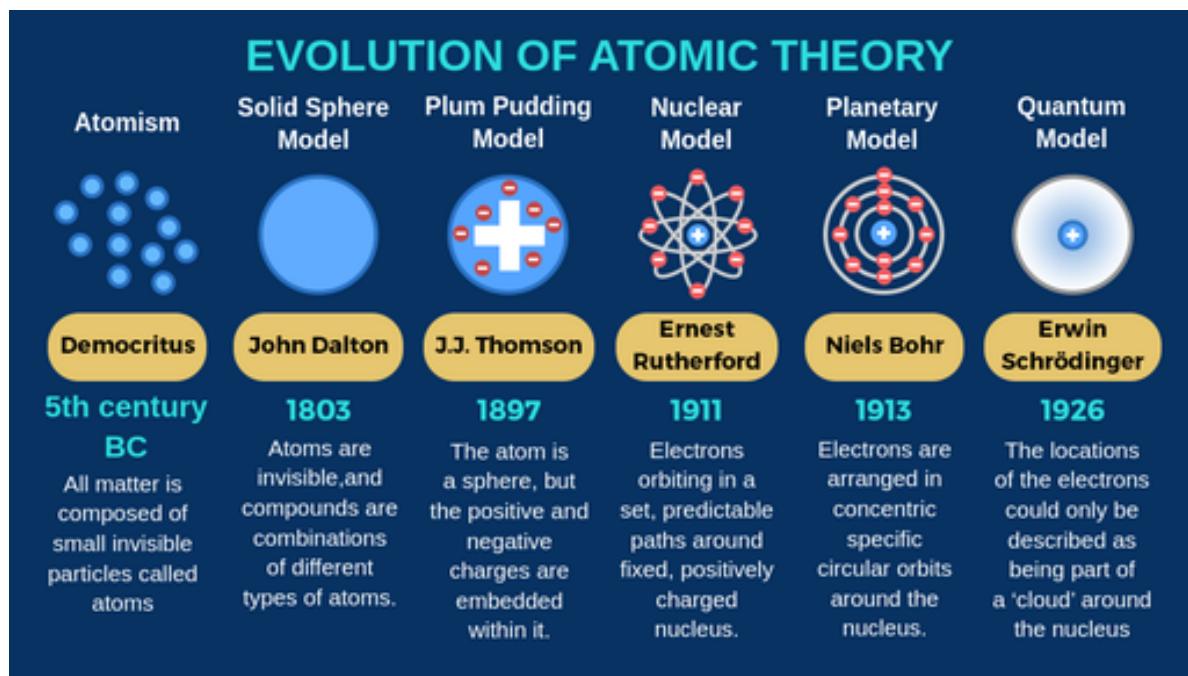


Figure 1.19: History of the structure of the atom.

1.4.1 The ‘classical’ problems with Rutherford’s model

Geiger, Marsden, and Rutherford experiments on α -particle scattering (1908–1913) revealed that most atomic mass and all positive charge is concentrated in a central nucleus. This marks the end of the “plum pudding” model... but there is a problem. According to classical physics, electrons in orbits (accelerating charges) should emit radiation and therefore lose energy and spiral in to the nucleus. What does prevent this from happening? Classical physics could not provide a valid reason.

1.4.2 The Bohr model

Bohr (1913) solved this problem with a theory based mainly on classical mechanics, but with some new quantum ideas. Bohr’s postulates were:

1. Electrons move in *circular orbits* determined by Newton’s and Coulomb’s laws.
2. These circular orbits are *quantised*: the electron can’t occupy *any* region around the nucleus in a continuous way, but only stable orbits, hence it can’t spiral into the center. Since the energy is constant with time, these are called “*stationary states*” (we’ll get back to this concept later).
3. To go from one orbit to another, the electron must emit or absorb quanta of energy: this explained the discrete absorption and emission spectral lines!

For example, a spectral line at frequency f is emitted when the atom transitions from the orbit at energy E_i to the orbit at energy $E_f > E_i$: $\$ hf = E_i - E_f \$$

4. The atomic angular momentum is quantized and the only orbits allowed are the ones for which the angular momentum is a multiple of $\hbar = \frac{h}{2\pi}$ (reduced Planck constant), i.e. $L = mvr = n\hbar$. This idea came from the fact that the Planck constant h has dimension of angular momentum.

Starting from that, Bohr predicted what the angular momentum and allowed orbits should have been for outer electrons in atoms with large mass number, and observed consistency with classical electrodynamics predictions.

1.4.3 Bohr’s correspondence principle

Why did the quantum and classical results agree in atoms with large mass number? And why is angular momentum quantized?

Because in large systems classical physics is a good approximation of quantum physics! This is the **correspondence principle**, and we will encounter it again later in the course.

Bohr's correspondence principle

Predictions of quantum theory must correspond to the predictions of classical physics in the region of sizes where classical theory is “known to hold”.

Quantum systems are usually described by quantum numbers n , and the classical limit should be recovered from the quantum results, for $n \rightarrow \infty$, i.e.

$$\lim_{n \rightarrow \infty} [\text{quantum physics prediction}(n)] = [\text{classical physics prediction}]$$

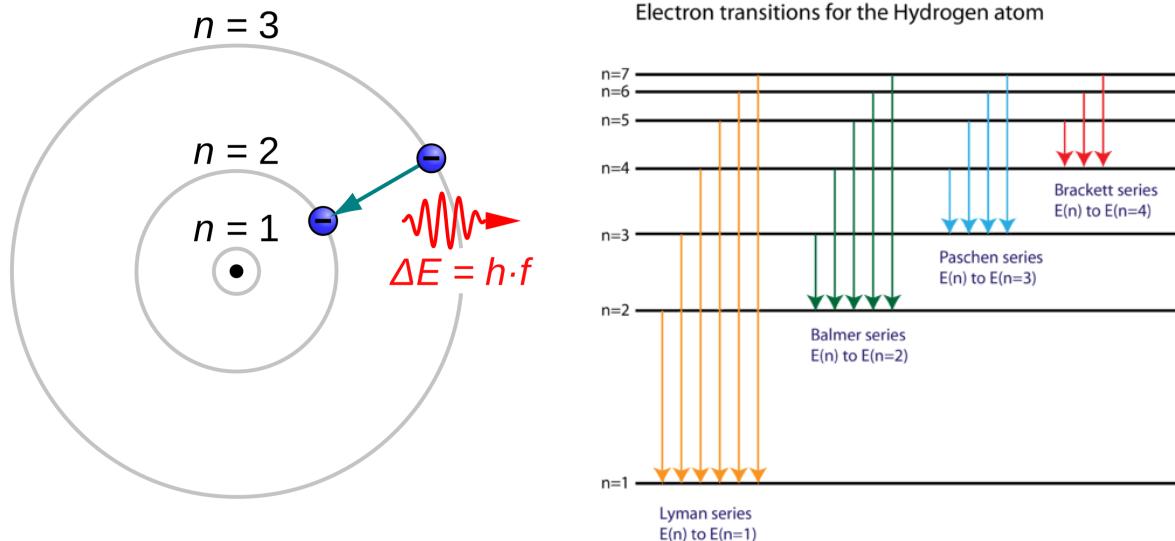


Figure 1.20: Quantized electron orbitals. To move from one orbital to one at a lower energy, the electron must emit a photon of energy equal to the energy gap between the two orbitals. To move to a higher energy orbital, it must absorb a photon with the energy difference between the two orbitals.

Figure 1.21: First three series of spectral lines of the Hydrogen atom.

Historical remark: Arthur Erich Haas

In reality, the first quantum atomic model with quantised orbitals was proposed by **Arthur Erich Haas** in 1910 [4] (3 years before Bohr). Haas presented his work at a lecture in Austria held during carnival time, and his idea was at that time rejected and ridiculed. As Haas recalled: “When I lectured to the Chemical-Physical Society of Vienna ... Lecher ... referred to the presentation during open discussion as “a carnival joke”. Haas was in fact the first one

to estimate the Bohr radius and to create a connection between atomic orbitals and Planck constant! It is unclear what changed a shift in the opinion and led to the acceptance of Bohr's theory only three years later, but probably the times were not mature enough to accept quantum concepts more widely when Haas proposed his idea.

1.4.4 Compton's scattering

We saw that the Bohr's model explained the emission and absorption spectral lines. What happens if electrons interact with light at high frequencies such as X-ray? In this case it is reasonable to consider the electrons *free*, since the binding energy of the atom is very small compared to the X-ray energy, so the interaction between the high-frequency radiation and the electron would result in scattering.

Classical expectation: when the EM wave is scattered off atoms, the wavelength of scattered radiation should be the same as the wavelength of the incident radiation. X-rays had been discovered in 1895 by W. K. Rontgen and were thought to be electromagnetic radiation of high frequency. Based on Thomson's theory (Thomson's scattering), the oscillating electric field of the incident radiation would make the electrons of the target atoms vibrate with the same frequency, and these would radiate electromagnetic waves at the same frequency.

This is not what Compton observed though. In 1922 Compton studied inelastic scattering of X-rays from graphite, and found that, with increasing scattering angle, a second peak was appearing at longer wavelength compared to the one of the incident light (Figure 1.22). This change in wavelength between the peaks of the incoming beam and the scattered one is the **Compton shift** and was found to be

$$\Delta\lambda = \lambda_f - \lambda_i = \frac{h}{m_e c} (1 - \cos(\theta)), \quad (1.25)$$

where m_e is the mass of the electron and θ is the scattering angle.

The constant in Equation 1.25, $\frac{h}{m_e c} = 2.4 \times 10^{-12}$ m is the **Compton wavelength of the electron**.

Derivation of the Compton shift

Compton used both special relativity and the quantisation of light to derive a formula for the Compton shift. The derivation itself is quite straightforward as it only requires the application of *conservation of energy* and *conservation of momentum* for the collision of a photon with a stationary electron. The only difference with a pure classical physics approach is that we need to use the relativistic energy expression and the Planck formula.

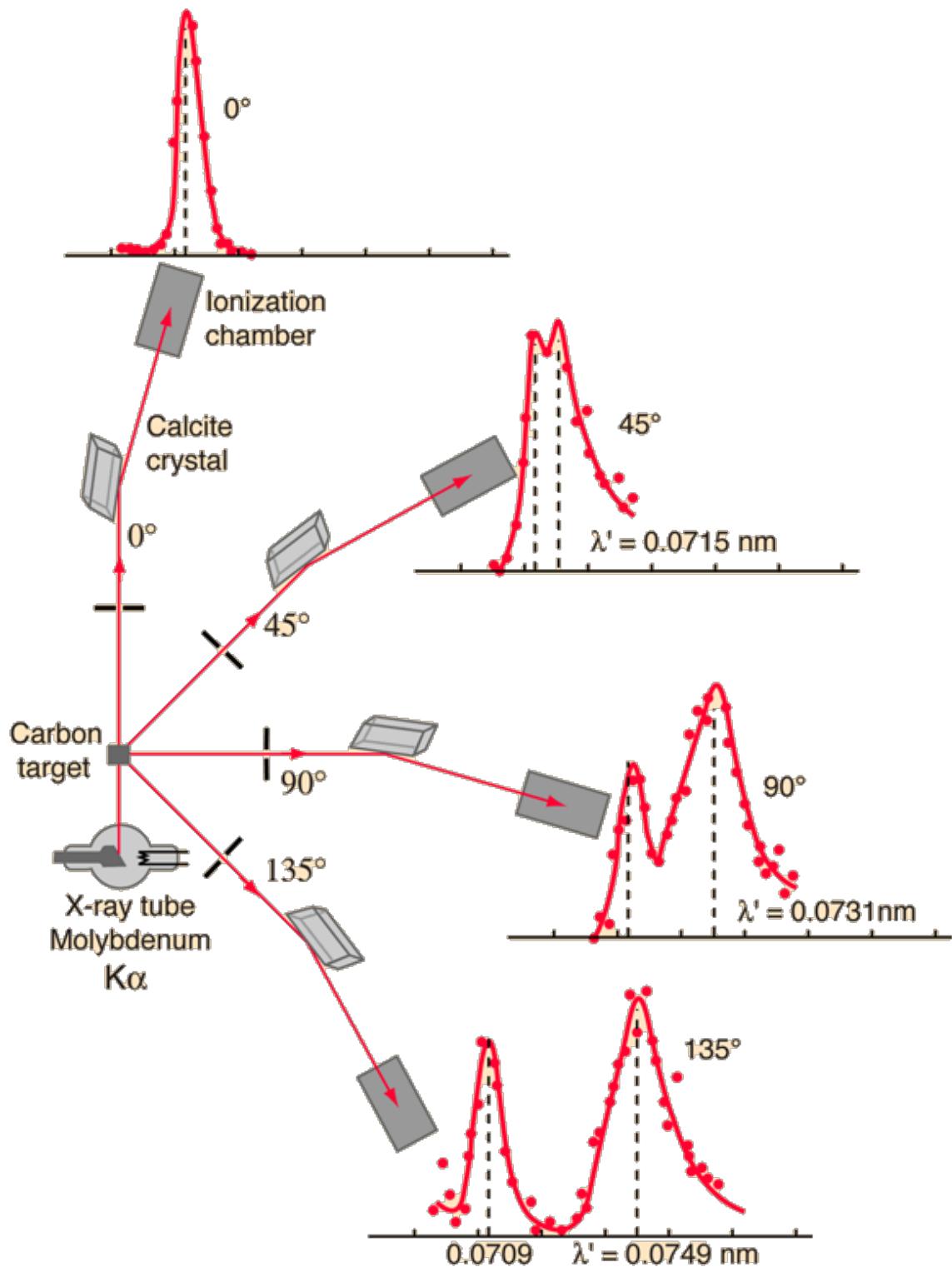


Figure 1.22: Compton scattering experiment: the secondary peak at longer wavelength starts appearing with increasing scattering angle.

Compton scattering

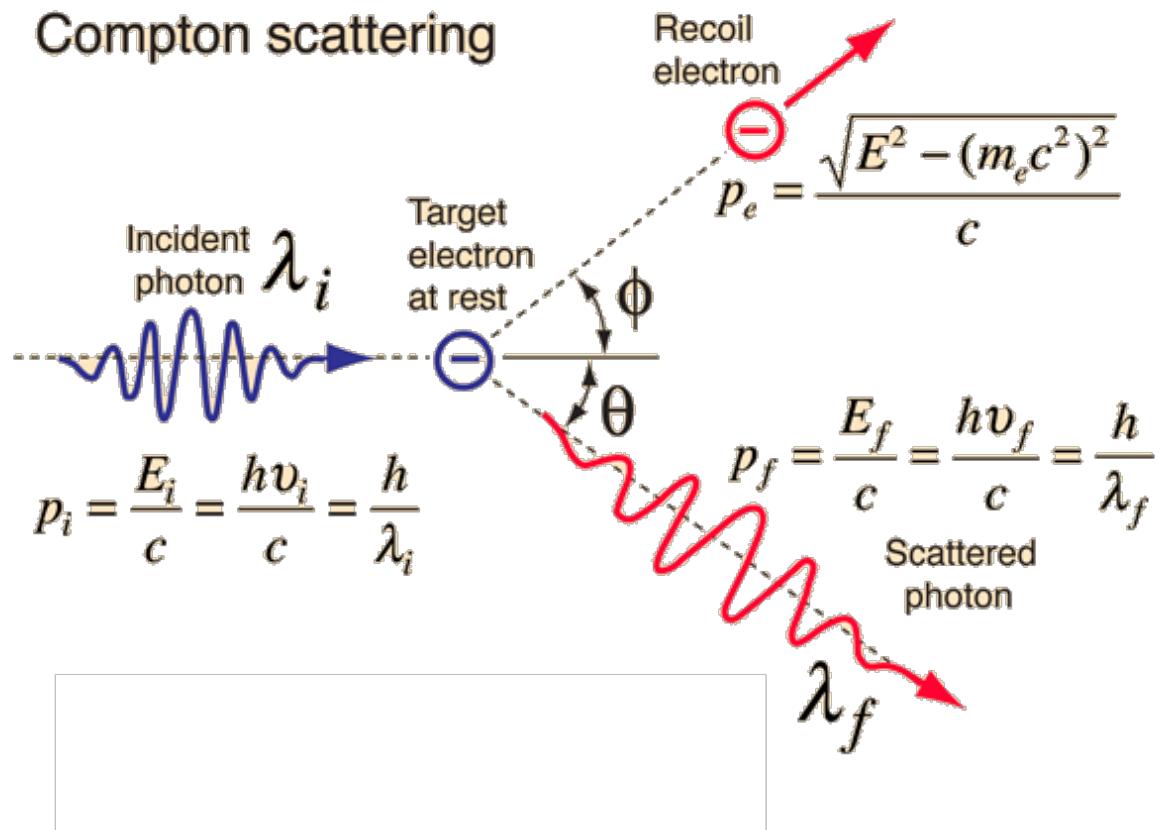


Figure 1.23: Compton scattering: a high energy photon collides with an electron at rest. The electron gains energy, so the scattered photon has a lower frequency.

- From *conservation of energy*:

$$h\nu_i + m_e c^2 = h\nu_f + \sqrt{p_e^2 c^2 + m_e^2 c^4}$$

Squaring this and rearranging the terms:

$$(p_e c)^2 = (h\nu_i)^2 + (h\nu_f)^2 - 2h^2 \nu_i \nu_f + 2m_e c^2 (h\nu_i - h\nu_f) \quad (1.26)$$

- From *conservation of momentum*:

$$\vec{p}_i = \vec{p}_e + \vec{p}_f$$

Rearranging this and using the definition of the scalar product:

$$p_e^2 = (\vec{p}_i - \vec{p}_f)^2 = p_i^2 + p_f^2 - 2p_i p_f \cos(\theta).$$

We can now multiply each side by c^2 and use the relativistic and Planck's energies ($pc = h\nu$):

$$\begin{aligned} p_e^2 c^2 &= p_i^2 c^2 + p_f^2 c^2 - 2c^2 p_i p_f \cos(\theta) \\ (p_e c)^2 &= (h\nu_i)^2 + (h\nu_f)^2 - 2h^2 \nu_i \nu_f \cos(\theta). \end{aligned} \quad (1.27)$$

- Equating Equation 1.26 and Equation 1.27 and rearranging, we obtain

$$\frac{1}{h\nu_f} - \frac{1}{h\nu_i} = \frac{1}{m_e c^2} (1 - \cos \theta),$$

which using $\nu = c\lambda$ gives the Compton shift of Equation 1.25:

$$\Delta\lambda = \lambda_f - \lambda_i = \frac{h}{m_e c} (1 - \cos(\theta)).$$

1.4.5 The problems with Bohr's model

In the years 1920s it was clear that Bohr theory still presented some unresolved problems:

- It failed to predict the observed intensities of spectral lines.
- Good for one-electron atoms, but very limiting for multi-electron atoms.
- It failed to provide an equation of motion governing the time development of atomic systems starting from some initial state.
- It overemphasized the particle nature of matter and could not really explain the origin of stationary states in the orbits or wave-particle duality.
- It did not supply a general scheme for “quantizing” other systems, especially those without periodic motion.

1.5 De Broglie's hypothesis and matter waves

In 1923 Louis de Broglie (Figure 1.24) introduced new ideas that answered some of the above open questions. In analogy with the light wave-particle duality, de Broglie suggested that, microscopically, matter may do the same thing, having both particle-like and wave-like properties. He suggested that this would for instance be the case for electrons: every electron was “accompanied” by a (non electromagnetic!) wave, which “piloted” the electron through space.



Figure 1.24: Louis de Broglie, physicist and aristocrat.

The wave-like and particle-like properties of light can be related by following equation for the momentum p of the photon, $\$ p = k = h / c = h / \nu \$$, where $k = 2\pi/\lambda$ is the wavenumber and ν and λ are respectively the frequency and wavelength of the light wave.

In analogy to this, de Broglie introduced the concept of wavelength associated to any object moving with a momentum p - the de Broglie wavelength.

De Broglie wavelength

$$\lambda_{dB} = \frac{h}{p}. \quad (1.28)$$

We can also use the form $p = \hbar k$, using the de Broglie wavenumber $k = 2\pi/\lambda_{dB}$.

1.5.1 Interpretation of the Bohr's circular orbits

The idea of de Broglie gives a *qualitative* explanation of Bohr's idea of quantized orbits, as the waves associated to the electrons would interfere, leaving areas of destructive interference where "electrons are forbidden", while only standing waves would be allowed to "fit" in the Bohr's circular orbits. These would be given by a discrete set of wavelengths that interfere constructively if an integral number of wavelengths fits exactly into the circumference of the orbit (see Figure 1.25), i.e.

$$n\lambda = 2\pi r,$$

where r is the radius of the circular orbit.

1.5.2 Exercise

Determine the de Broglie wavelength of a 54eV electron.

💡 Hint

What is the kinetic energy of the electron?

Solution

The kinetic energy of the electron is $\frac{p^2}{2m} = E = 54\text{eV}$, therefore $p = \sqrt{2mE} = 1.67 \times 10^{-10}\text{m} = 1.67 \text{ Å}$ (remember to convert energies to Joules!). This is the typical scale of spacing between atoms in a solid.

```
import numpy as np
from scipy.constants import m_e, h

E_eV = 54 #eV
```

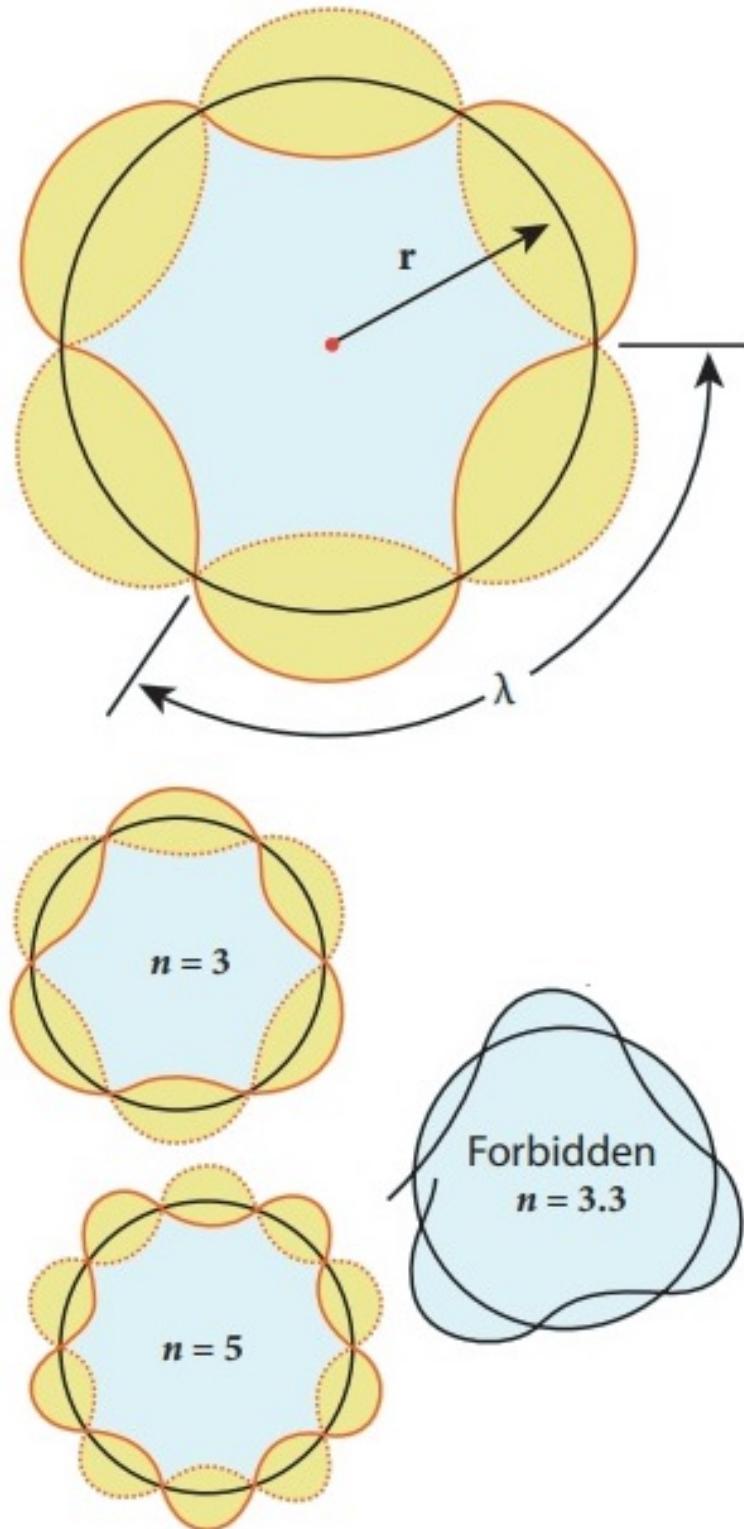


Figure 1.25: Standing waves fit to a circular Bohr orbit if the circumference is an integer multiple of the electron's de Broglie wavelength.

```

E = E_eV*1.60218e-19 #J. Converted the energy to Joules here
m_e #kg
p = np.sqrt(2*m_e*E)
lambdadB = h/p
print(f'De Broglie wavelength = {lambdadB:.2e} m')

```

De Broglie wavelength = 1.67e-10 m

1.5.3 Electron diffraction experiments

Designing and conducting an experiment to confirm de Broglie's hypothesis was not an easy task, why? The most straightforward thing to show that small entities like electrons can in fact behave like waves would be to try to observe wave-like behaviour through experiments already used in optics, for instance diffraction. But... how small should the distance between slits be in a diffraction experiment, to observe interference with such "electron waves"?

Think about the number that we obtained in the previous exercise for the de Broglie wavelength, which sets the scale at which the electron would exhibit quantum wave-like properties. This is of the order of nm, which is the typical distance between atoms in a molecules. How can we design a diffraction grating with such small distances? Any macroscopic diffraction slit would be too big for this purpose. Well, the only way would be to use the molecules themselves, using **atomic lattices!**

In 1925, Davisson & Germer were conducting electron scattering experiments to understand the disposition of atoms on a nickel surface, when they accidentally broke the vacuum in their experimental setup, and to get rid of the oxidation of the nickel surface they heated up the sample. Upon repeating the experiment, different results were found and electron Bragg scattering was accidentally measured! The reason was that the prolonged heating had changed the internal structure of the atoms in the nickel target, forming large single-crystal regions in the polycrystalline sample, which provided an ideal lattice to observe electron diffraction.

From the diffraction equation, $\lambda = 2ds\sin((\pi - \theta)/2)$ for Ni with $d = 0.091$ nm, the peak observed at $\theta = 50$ deg was consistent with the the Broglie wavelength of 54 eV electrons. At the same time and independently, George Thomson also observed electron diffraction with aluminium films, and in 1937 he shared the Nobel Prize with Davisson (not with Germer, who had a B plan for his career and became an offbeat rock climber - always think about a B plan...).

Such experiments have since been extended to neutrons, atoms, molecules, fullerene (C_{60}), etc.

1.6 (1926) Schrödinger's atomic model: a *really* quantum model of the atom

You may wonder what I mean with “*really* quantum model of the atom”, when Bohr already introduced the quantisation of energy levels, even managing to explain the spectral lines and de Broglie introduced the concept of waves. We’ve been very vague about the physical meaning we give to such “pilot waves that guide the electrons through space”, but it seems that we have all the ingredients for a really quantum model of the atom. What else do we need?

In the following chapter(s) we will see that the notion of representing a particle as a localized wave or wave group leads naturally to limitations on simultaneously measuring position and momentum of the particle (Heisenberg’s uncertainty principle).

The reality is that the Bohr’s model, even with the de Broglie interpretation, still doesn’t capture the *probabilistic* nature of quantum physics. This is where we really start getting into the more fundamental differences in nature between classical and quantum physics.

What is different about quantum physics?

We talked about quantised light (photons) and quantised circular orbits for the atom, and we saw that we *need* to introduce these to explain phenomena that can’t be explained by using classical physics. However, the most striking difference between classical and quantum physics lies in their fundamental nature.

Classical physics is *deterministic*: given a system, if you have enough information about it (i.e. about the initial configuration of an experiment) and know the physics to describe its dynamics, you can *determine* what will happen next to said system, i.e. what configuration it’s going to be in after some time. An important aspect is that the process of observing (*measuring*) a classical system does not affect the final configuration of the system. We’ll get into more details about this later.

Classically, an electron in Bohr’s model is still imagined as a particle orbiting about the nucleus. However, that’s not exactly the correct picture, as de Broglie pointed out. At such small scales, we have to abandon the classical description and resort to the quantum interpretation.

Quantum physics, in contrast to classical physics, is *probabilistic*. What does this mean for the electron? It means that actually the electron is not really a particle: just like light, it can behave like a wave and its description is given by a *wavefunction*, a probability distribution indicating the regions where the electron is more likely to be (observed after measurement) (Figure 1.26). The region of space around the nucleus where there is a high probability of observing the electron (~90-95%) is called an *orbital*, while ‘forbidden’ regions around the nucleus are the *nodes* (Figure 1.27).

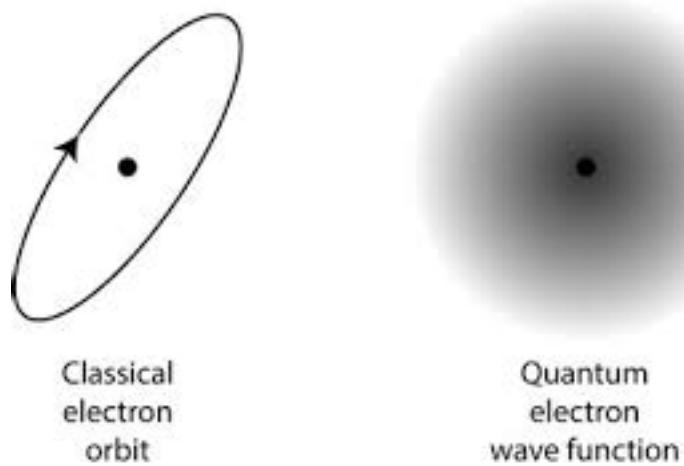


Figure 1.26: Classical vs quantum electron visualization.

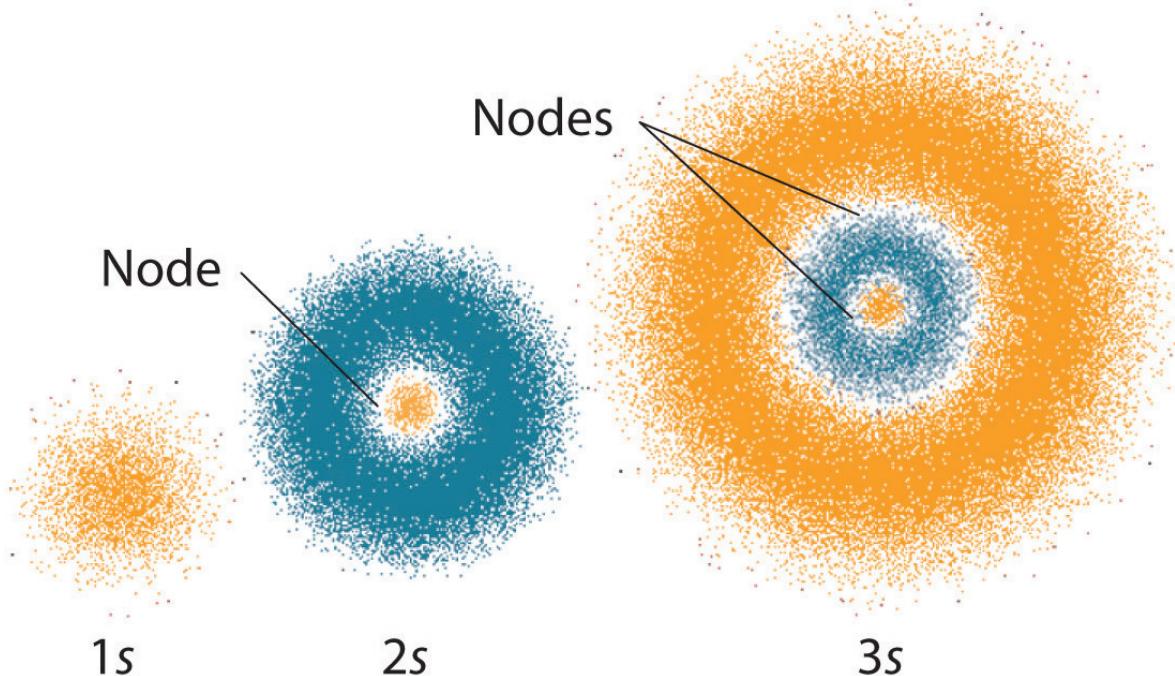


Figure 1.27: Probability distribution of electron for different orbitals (1s, 2s, 3s). More densely coloured regions indicate areas where electrons are more likely to be. White areas (nodes) are regions where electrons have zero probability of being found. Image credit: [UCDavis Chemwiki, CC BY-NC-SA 3.0 US](#)

1.6.1 Electron orbitals

Electron orbitals introduced by Schrödinger are characterised by quantum numbers. Schrödinger introduced the concept of electron orbitals and *quantum numbers* that describe them:

- **Principal Quantum Number (n):** Indicates the energy level and relative size of the orbital. It can take positive integer values (1, 2, 3, ...).
- **Angular Momentum Quantum Number (l):** Defines the shape of the orbital and can take values from 0 to $n - 1$ for each value of n . Each value of l corresponds to a specific type of orbital (s, p, d, f...).
- **Magnetic Quantum Number (m_l):** Describes the orientation of the orbital in space and can take integer values from $-l$ to $+l$, including 0.
- **Spin Quantum Number (m_s):** Specifies the electron's spin direction, which can be either $+1/2$ or $-1/2$.

We'll get back to these quantum numbers later on, when we study more in detail the orbital angular momentum and spin.

We are starting to introduce concepts that are very different (if not nonexistent) in classical physics, such as:

- *State*
- *Superposition*
- *Measurement* (its interpretation and implications in quantum physics)
- (and more...)

These concepts will be discussed in greater detail during the course (they will in fact be our 'bread and butter during' the semester!) and we will come back to explaining them in due time. Now, let's try to understand the probabilistic quantum nature with a known experiment: Young's double slit experiment.

In the next chapter we will see how the wave-particle duality can be observed and understood through the double slit experiment with electrons (and neutrons).

1.7 Recap

- Black body spectrum first investigated by Wien (1894) via thermodynamic arguments
- Rayleigh–Jeans model in 1900 argued from equipartition and mode density → ever-increasing modes with same energy: ultraviolet catastrophe!
- Wien model worked at high frequencies, Rayleigh at low frequencies

- Planck (1900) interpolated between the two, fitting data at all wavelengths and deriving Stefan's Law for total radiated power
 - Post-hoc understanding of Planck model: discretisation of EM mode energies as integer multiples of $\epsilon = h\nu$.
 - Photoelectric effect and Bohr atomic spectra evidence quanta of light and quantized atomic energy levels
 - Compton effect shows particle nature of light
 - de Broglie: wave–particle duality (and wave mechanics). No real radiation/matter distinction
 - Experimentally confirmed by electron (and more) diffraction
-

2 Wavefunctions and the uncertainty principle

Learning objectives

1. Understanding the concept of wave-particle duality and its experimental confirmations
2. Applying the concepts of wavefunctions and probability amplitudes to describe free particles as matter-waves
3. Understand the Heisenberg uncertainty principle's origin and its implications

The content of this lecture is mostly covered in Chapter 1 of Feynman's lectures [2].

2.1 Wave-particle duality

In the previous chapter we saw how classical physics is not sufficient to describe a series of phenomena, and that we need to introduce concepts such as *quantisation* of energy and *wave-particle duality* that applies not only to light, but also to matter.

In the following we will see that matter waves, in analogy to light, can be described by quantities called **wavefunctions**. The physical meaning of these is not immediate, as they differ from conventional water, sound or light waves, but we will discuss this further later in the chapter and it will become more clear after we go through different variations of the double-slit experiment.

2.2 The double-slit experiment

A typical experiment that shows the wave-particle duality is the double-slit experiment, which you have already seen in other courses such as “Classical and Quantum Waves” and will be briefly discussed here.

2.2.1 Feynman's lectures

If you want a thorough description of the wave-particle duality, the double-slit experiment and how this impacted the development of quantum physics, there are endless resources you can read, including the books in the suggested reading list. I have a personal preference in this regard, and it's provided by Feynman. His "Lectures on Physics" are freely available online [here](#) (also embedded below), and I recommend reading the first three chapters of the third book (or at least the first one - the third gets more technical). They are very descriptive, so do not be scared by the length, but they are very useful for a fundamental understanding of the physics behind the mathematical description!

Feynman did a pretty fine job at going through this extensively in three chapters, so forgive me if I don't make an effort at reinventing the wheel by trying to reformulate what he perfectly explains, I'll just summarise these concepts briefly and use them to introduce some notation that we will keep using for the rest of the course and that you need to familiarise yourself with.

The two chapters from Lectures on Physics can be found at the following links, also embedded below. I will mainly focus on the first chapter for the various double-slit experiments, while the second chapter is more focused on understanding of the implications of the previous experiments and how these relate to the Heisenberg's uncertainty principle.

- [Chapter 1 - Quantum behaviour](#)
- [Chapter 2 - The Relation of Wave and Particle Viewpoints](#)

2.2.2 Experiment 1: An experiment with bullets

Let's start with our double-slit setup for the first experiment with bullets: a machine gun shooting bullets, a wall with two slits (each just big enough to let one bullet go through), and a backstop with a detector, where the bullets stop and accumulate.

Let's break down the experiment in three phases:

1. Only slit #1 open

Let's suppose to divide the screen in squares with sides of length Δx , and of shooting a total of N bullets.

What we observe is a distribution of bullets $N_1(x)$, giving a probability $P_1(x) = N_1(x)/N$ of finding a bullet, coming from slit 1, at the position x (see P_1 in Figure 2.1). A bullet could hit the edge of the slit and be deflected in its trajectory, so we can't know *exactly* where the bullet would go, but can know what the probability that it will be found at the coordinate x of the backstop would be. This probability distribution will be a bell curve, centered around the coordinate aligned with the slit.

2. Only slit #2 open

If now we close slit 1 and open slit 2, we will have the same kind of result as for phase 1, but now with the bell curve $P_2 = N_2(x)/N$ centered around the coordinate aligned with the second slit (see P_2 in Figure 2.1).

3. Both slits open

What do we observe if both slits are open? For the purpose of this experiment, let's stick with the assumption that bullets are *indivisible* and macroscopic and arrive one per time at the backstop. They will either go through slit 1 or slit 2. The total probability that the bullet will hit a position x in the screen is the sum of the individual probabilities P_1 and P_2 obtained when only one of the slits is open (see P_{12} in Figure 2.1):

$$P_{12}(x) = P_1(x) + P_2(x).$$

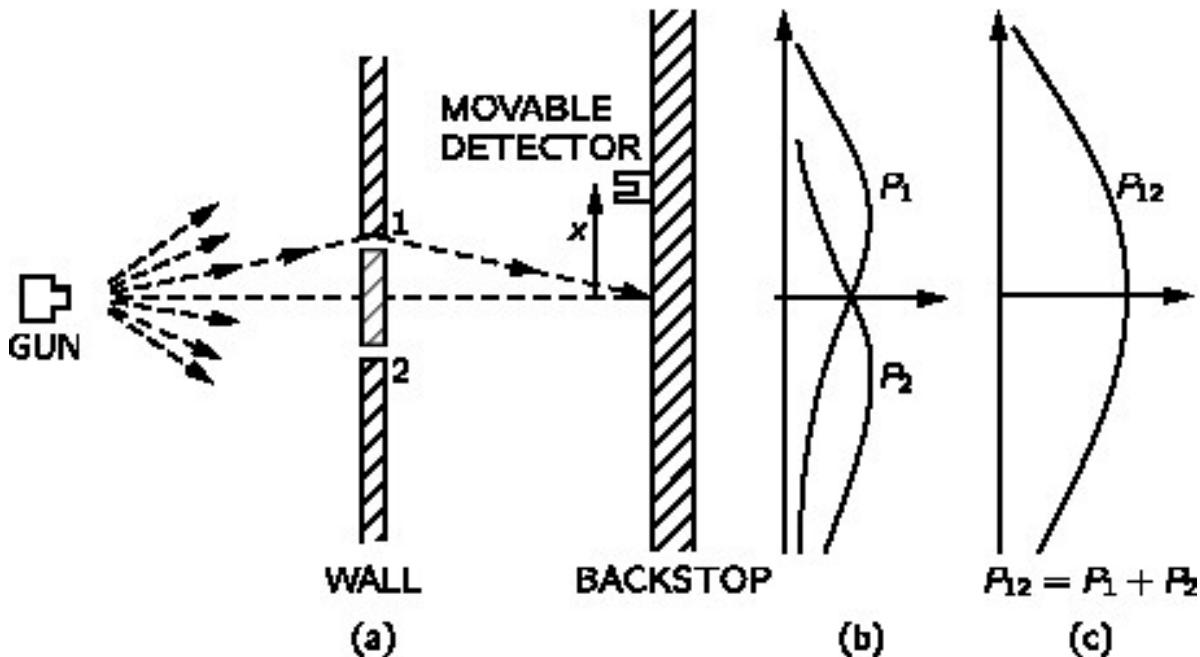


Figure 2.1: Double slit experiment with bullets. Source: Feynman, Leighton, and Sands [2]

No interference is observed in this experiment!

2.2.3 Experiment 2: An experiment with waves

Let's consider the same setup, but instead of a gun shooting bullets, let's now consider a wave source, such as water waves, and a screen which measures the “intensity” of the wave motion. Concentric wavefronts propagate from a point source and hit the two slits, which we can then

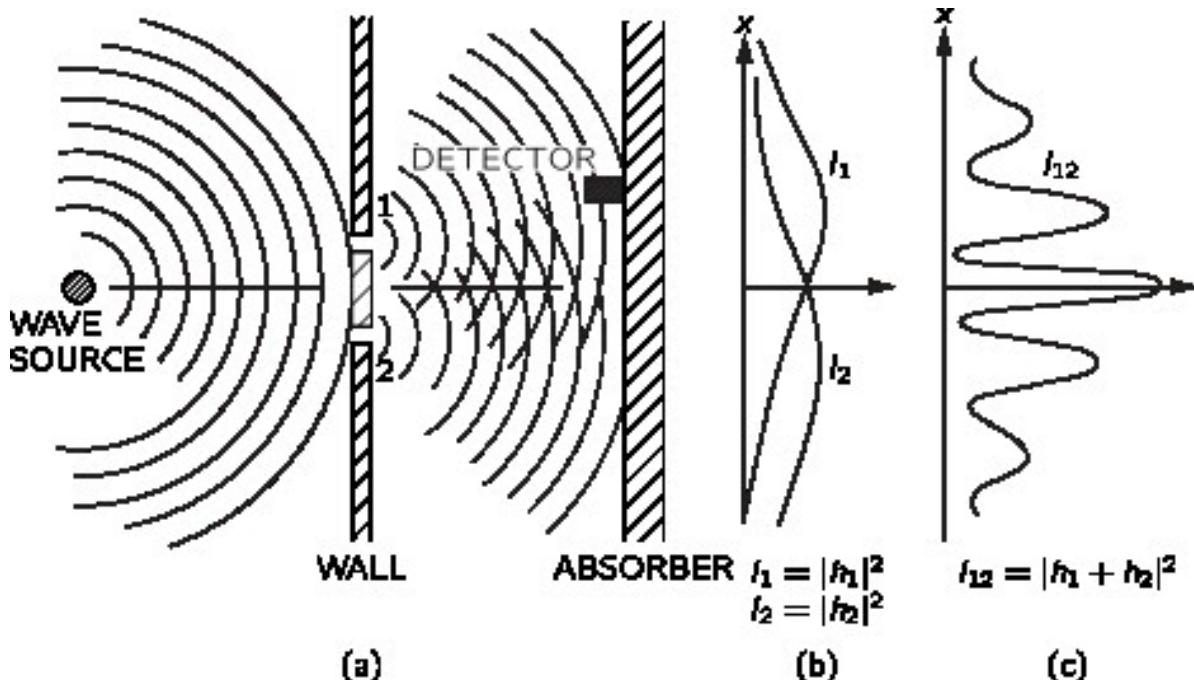


Figure 2.2: Double-slit experiment with waves. Source: Feynman, Leighton, and Sands [2]

consider as point-like sources of secondary waves (using Huygen's principle, which you should remember from the course on Waves). The setup is illustrated in (Figure 2.2).

What we observe is that if we cover one of the two slits, we have the intensity bell curves indicated with I_1 or I_2 in Figure 2.2. However, if both slits are open, the waves interfere and form a pattern with maxima and minima of intensity, corresponding to positions where they interfere constructively and destructively, respectively.

In contrast to the previous experiment with bullets, where the total probability as a function of space was the sum of the two probability distributions, here the intensity observed is not simply the sum of the two intensities obtained when one of the slits is closed: $I_{12} \neq I_1 + I_2$.

The key point here is that we cannot simply consider a sum of intensities, we need to consider the waves *amplitudes*: the intensity is in fact proportional the modulus square of the amplitude. This means that, if the amplitudes of the secondary waves coming from slits 1 and 2 are respectively h_1 and h_2 , the corresponding intensities with only one slit open are $I_1 = |h_1|^2$ and $I_2 = |h_2|^2$ (see Figure 2.2). However, when both slits are open, the total intensity will be given by the square of the *sum of the amplitudes*:

$$I_{12} = |h_1 + h_2|^2.$$

Interference is observed with waves!

2.2.4 Experiment 3: An experiment with electrons

What do we expect to observe in an experiment with electron guns that send one electron per time through two slits?

The set-up is similar to the one in the first experiment, except that now the gun is shooting electrons instead of bullets, and at the screen we have a detector of electron that “clicks” every time an electron hits it.

Since we are shooting electrons, one per time, we would expect something like the probability distribution in the first experiment with bullets...

...however, this is not what we observe (not always at least - we’ll get back to this soon).

What we observe is illustrated in Figure 2.3.

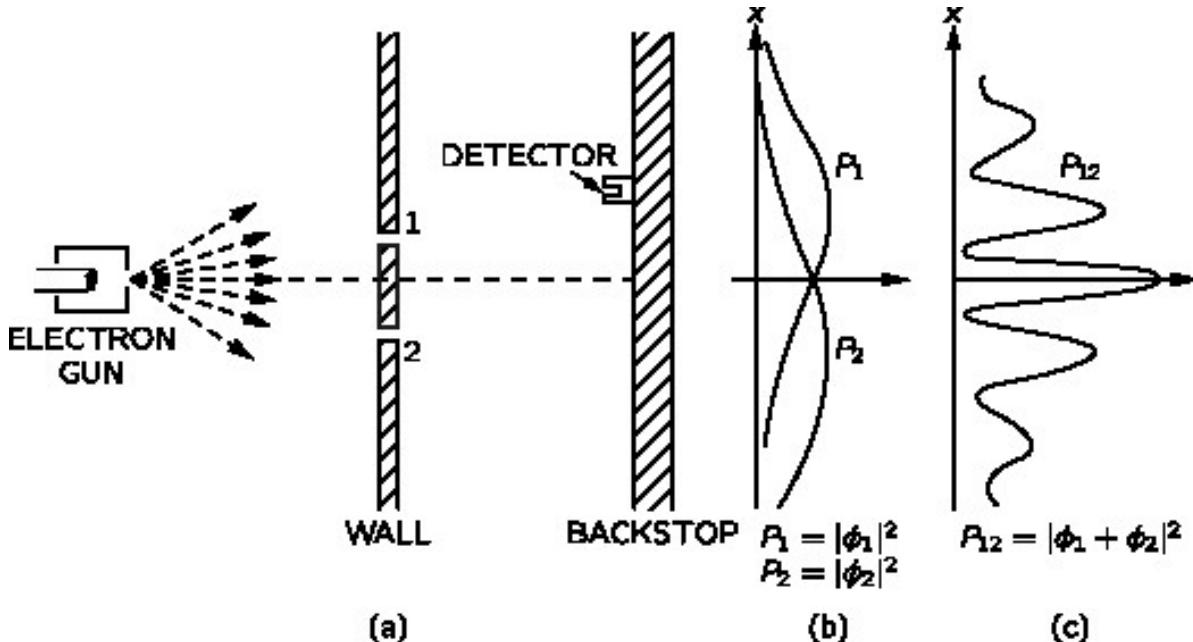


Figure 2.3: Double-slit experiment with electrons. Source: Feynman, Leighton, and Sands [2]

Do electrons split across the two slits?

Every “click” registered has always the same loudness - each electron detected is always a whole *localised* entity with its own mass and charge. No half electrons are detected, ever! This means that *electrons do not split across the two slits*.

So does this mean that each electron goes either through slit 1 or slit 2?

No, because if that was the case, the probability P_{12} would just be the same of P_1 and P_2 , as with experiment 1. Instead, $P_{12} \neq P_1 + P_2$: there is interference!

So in the same experiment electrons behave like *waves while in transit, before they get detected*, but behave like *particles upon detection*.

After all, the same experiment can be done with light, and in that case we observe the same interference pattern. In that case we can think of the interference of electromagnetic waves, in the light-as-wave description, but in the detection process light behaves as a particle, since it is detected by the photoelectric effect, which needs a (quantised) photons description to be explained (see previous chapter). Even sending one photon per time, the interference pattern is observed, so it is not a result of more photons interfering.

Ok, but maybe since electrons are charged they just interfere with each other and form that interference pattern?

No, actually the same pattern is observed also if you use neutrons, so it does not depend on the charge, it really depends on the *nature* of matter and how it manifests in different scenarios.

We can see where this is going: wave-matter duality. The interference pattern that we obtain in this experiment is analogous to the one we saw for intensity in experiment 2. In that case, we saw that the apparently complicated pattern of the intensity is actually very simple if we consider the sum of the wave amplitudes, which are physical and related to the oscillating electric and magnetic fields. But then, how do we explain this with electrons? We need to introduce a concept analogous to the classical wave amplitudes and to the classical waves, but for probabilities. So we come up with an analogous concept to describe the wave behaviour of matter: the *wavefunction*, introduced at the beginning of the chapter, playing the role of *probability amplitude*.

In this experiment we use the symbol ϕ for the wavefunction, so the probability is $P = \phi^* \phi = |\phi|^2$.

If an event - like the detection of a particle on the screen - can happen following different possible paths, we associate a wavefunction (or probability amplitude - a complex number) to each path. In this case, for two slits, we have two probability amplitudes ϕ_1 and ϕ_2 associated to each possible trajectory. The probabilities of detecting a particle at the screen, coming from slit 1 and slit 2 respectively (when one of the slits is closed), are therefore $P_1 = |\phi_1|^2$ and $P_2 = |\phi_2|^2$.

When both slits are open, the total probability of observing a particle at the screen is *not* the sum of these two probabilities. In analogy with the intensity for waves calculated from the square of amplitudes, the total probability is obtained from the modulus square of the *total probability amplitude* $\phi_1 + \phi_2$:

$$P_{12} = |\phi_1 + \phi_2|^2 = |\phi_1|^2 + |\phi_2|^2 + 2|\phi_1||\phi_2| \cos \theta. \quad (2.1)$$

In this case we say that the electron is in a **superposition state** $\phi = \phi_1 + \phi_2$.

Note

Note that the last term in Equation 2.1, $2|\phi_1||\phi_2| \cos \theta$, is an interference term that makes the difference between this experiment and the classical case with no interference. θ is the phase between the two complex wavefunctions ϕ_1 and ϕ_2 .

The mathematics is exactly the same as the one that we have for classical waves once we introduce the concept of probability amplitude!

Interference is observed in this experiment with electrons (or neutrons)!

2.2.5 So is the electron a wave, a particle or both?

There are different interpretations about the wave-particle duality of matter, and maybe the most accepted answer to the above question is... *neither* (not at the same time at least!).

Electrons behave like waves on Mondays, Wednesday and Fridays, like particles on Tuesdays, Thursdays and Saturdays, and like nothing at all on Sundays.

– Readapted version of Sir Lawrence Bragg's quote (originally for light).

Electrons (and other matter) can behave like either particles or waves, depending on the kind of experiment performed on them, but *never* behave like both waves and particles *simultaneously*.

This statement reflects Bohr's concept of **complementarity**.

Complementarity

According to the *complementarity principle*, pairs of *complementary quantities*, like wave and particle properties, can't be measured simultaneously. The principle of complementarity is not limited to the wave-particle duality, but can be applied to other complementary quantities such as position and momentum, therefore extending to the uncertainty principle.

The ultimate origin of the difficulty lies in the fact (or philosophical principle) that we are compelled to use the words of common language when we wish to describe a phenomenon, not by logical or mathematical analysis, but by a picture appealing to the imagination. Common language has grown by everyday experience and can never surpass these limits. Classical physics has restricted itself to the use of concepts of this kind; by analyzing visible motions it has developed two ways of representing them by elementary processes: moving particles and waves. There is no other way of giving a pictorial description of motions — we have to apply it even in the region of atomic processes, where classical physics breaks down. Every

process can be interpreted either in terms of corpuscles or in terms of waves, but on the other hand it is beyond our power to produce proof that it is actually corpuscles or waves with which we are dealing, for we cannot simultaneously determine all the other properties which are distinctive of a corpuscle or of a wave, as the case may be. We can therefore say that *the wave and corpuscular descriptions are only to be regarded as complementary ways of viewing one and the same objective process*, a process which only in definite limiting cases admits of complete pictorial interpretation.

– M. Born, Atomic Physics, fourth edition, New York, Hafner Publishing Co., 1946, p. 92.

2.2.6 Experiment 4: An experiment with electrons... but we monitor the electrons through the slits!

Strange things happen when we add a strong light source between the two slits (as in Figure 2.4), so when the electron passes through a slit, it scatters some light and we can know which one it went through by seeing which aperture the “flash” comes from.

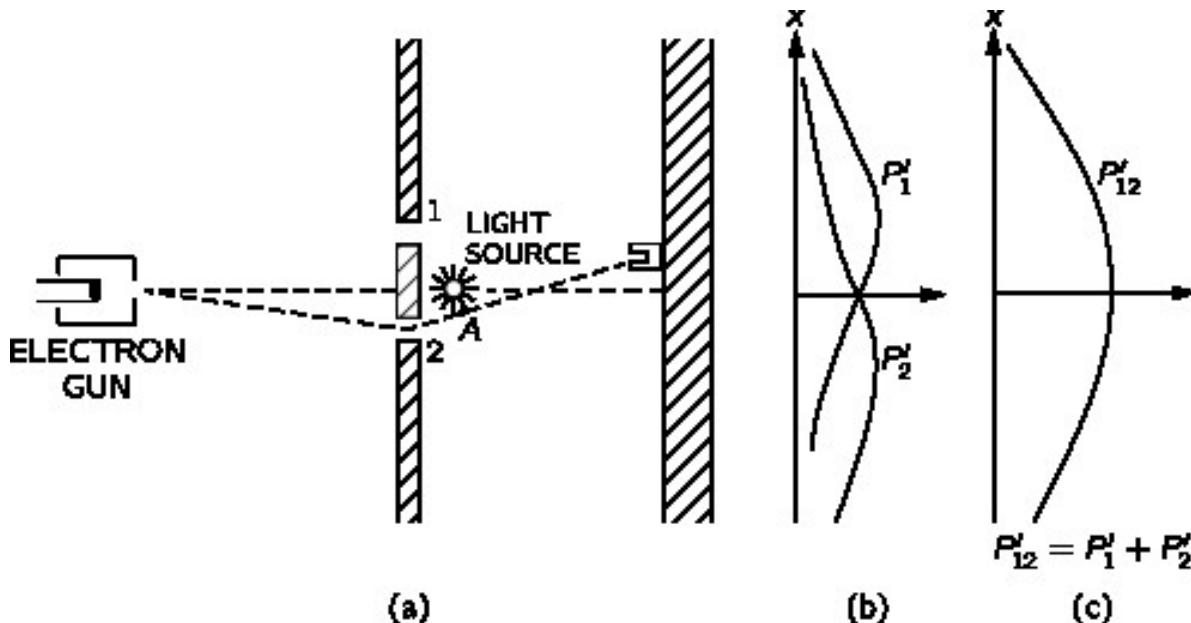


Figure 2.4: Double-slit experiment with electrons and a light source between the slits to observe the electrons’ path. Source: Feynman, Leighton, and Sands [2]

What do we observe once we start shooting electrons? Every time we hear a “click” (we detect an electron at the screen), we also see a flash, either coming from slit 1 or slit 2. However, we *never* observe two flashes coming from both slits!

So were we wrong when in the previous experiment we said that electrons do not go through either slit 1 or slit 2? Not really, because if we keep shooting electrons, we see something unexpected: the interference now disappears! We obtain the same result as in experiment 1 with bullets. However, if the light close to the slits is turned off, the interference pattern of experiment 3 is restored... What is going on here?

Something happens when we look at the electrons, their distribution changes, so their motion seems to be affected by the interaction with the light source used to “see” the electrons going through the slits. Indeed, the bright light source disturbs the electrons, “localising” them.

What if we try to use a less bright source, so that the electrons are less disturbed?

We can in fact make light dimmer, however the flashes of light observed will not get weaker, what changes is their frequency. With dim enough light, not all “clicks” will be preceded by a flash. In fact, reducing the *intensity* of the light source will not change the size of the photons, only their quantity, hence the rate at which they are emitted (remember the simulation on the photoelectric effect in the previous chapter). This explains why some electrons can go through a slit without scattering light at times, because at times there is no photon around while the electron travels! Let’s keep running this experiment with dim light anyway and let’s build two different histograms of the detections to reconstruct the probability distribution for the two different cases: when we see the flash and when we do not see the flash. If we keep the count of the detections where a flash is observed, we obtain the total distribution P'_{12} , without interference, of Figure 2.4. However, for the counts of detections for which a flash was not observed, we obtain again the interference pattern P_{12} of Figure 2.3! We can confirm that **if electrons are not seen, we have interference!**

What if we use a light with a lower frequency?

The kick given to the electrons by light depends on the momentum of the photons, $p = h/\lambda$, so if we want to decrease this, we should increase the wavelength using light at lower frequencies (e.g. infrared or radio), not lower intensities! If we try the experiment with increasingly lower frequencies, nothing seems to change, until we eventually try with light of wavelength comparable to the distance between the slits. Maybe you can guess what happens here though... if the wavelength is comparable to or longer than the distance between the slits, then we can’t distinguish these as two separate spots! So what we see in this case is just a big fuzzy flash, but we would not be able to tell whether it comes from slit 1 or slit 2, so we can’t tell which slit the electron went through. For such wavelengths then we start seeing again the interference pattern of Figure 2.3.

This phenomenon is actually a consequence of the *uncertainty principle*, which will see in the rest of the chapter, and for which we can give this alternative statement: *it is impossible to design an experiment to determine through which slit the electron passes without destroying the interference pattern by disturbing the electron.*

2.3 Wavefunctions and Born's probability interpretation

We introduce the concept of **wavefunctions** to describe the wave property of matter in analogy to light.

Wavefunction

A wavefunction $\Psi(\vec{r}, t)$ is a complex number associated to a particle (in the matter wave description of QM) and it is used to calculate the probability of finding the particle in a small region of space at a given time t , according to *Born's interpretation*.

Wavefunctions, used to describe matter waves, are quite different from waves like water, sound or electromagnetic waves: they are *not* measurable quantities, they are represented by **complex numbers**, so they *do not directly define probabilities*, but *probability amplitudes* that are used to find probabilities. To have physically “meaningful” measurable quantities, we need to obtain *real* values, so we use the modulus square of wavefunctions (or probability amplitudes) to obtain the probabilities.

! Probabilities from wavefunctions - Born interpretation

For a one-dimensional wavefunction $\Psi(x, t)$, the **probability density of finding a particle** in an infinitesimal region between x and $x + dx$ is

$$P(x, t)dx = |\Psi(x, t)|^2dx = \Psi(x, t)^*\Psi(x, t)dx. \quad (2.2)$$

- The probability of finding the particle in a finite region $a \leq x \leq b$ is

$$P(x, t) = \int_a^b |\Psi(x, t)|^2dx. \quad (2.3)$$

- Since the particle must be somewhere along the x -axis where the wavefunction is defined, the total probability of finding the particle in that region must be equal to 1:

$$\int_{-\infty}^{\infty} |\Psi(x, t)|^2dx = 1. \quad (2.4)$$

This is the **normalization condition** that we always need to make sure is satisfied by a wavefunction:

If you consider the wavefunction at a fixed time T , the probability of finding the particle in a certain interval along x is given by the area under the curve described by $|\Psi(x, T)|^2$.



Figure 2.5: Probability of finding a particle in the interval $a \leq x \leq b$, given the wavefunction $\Psi(x, T)$ at a given time T . The probability, given by Equation 2.3, is the highlighted orange area under the curve.

! Important

Because of the relation between the wavefunction and the probability, it is important to stress that a wavefunction $\Psi(x, t)$, function of x and t , must be:

- **single-valued** (i.e. for each value of x and t there is only one corresponding value of Ψ)
- **continuous** (i.e. you could “draw” it without lifting your pen from the paper)
- **smooth** (i.e. no sudden changes in the derivative - continuous derivative)
- **normalized** (see Equation 2.4)

We will see later more in detail why these conditions are required, but they will be important for the rest of the course, so remember them!

2.3.1 Exercise

The initial wavefunction of a particle is $\Psi(x, 0) = Ae^{-x^2/4}$.

1. Find the value of the coefficient A .
2. Sketch the probability of finding the particle against x .
3. What is the probability of finding the particle in the interval $[-x_0, x_0]$?

💡 Hints

- What properties must the wavefunction satisfy?
- It may be convenient to use the symmetry of the wavefunction.

Solution

1. You need to impose the normalization condition for the probability (i.e. the integral of the probability over the whole space must be 1):

$$\int_{-\infty}^{\infty} |\Psi(x, 0)|^2 dx = |A|^2 \int_{-\infty}^{\infty} e^{x^2/2} dx = |A|^2 \sqrt{2\pi} = 1,$$

which gives $A = (2\pi)^{-1/4}$.

2. The probability against x is given by the plot below:

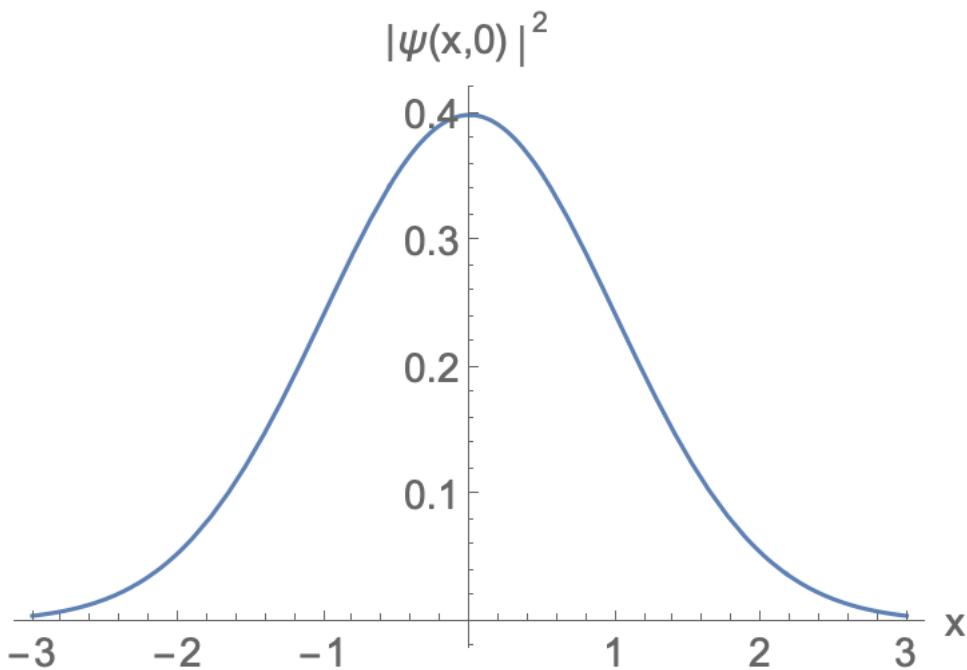


Figure 2.6: Probability of finding the particle described by the wavefunction $\Psi(x, 0)$, as a function of x .

2.4 The wave formalism for free particles

This part is based on the second chapter of Feynman's lectures.

We saw that the de Broglie wavelength (Equation 1.28) implies that a particle with a momentum p , like an electron, has to have an associated wavelength λ : $p = h/\lambda$.

So how do we represent this matter wave having this wavelength?

2.4.1 Free particle as plane wave?

The most simple thing to describe a wave with a given wavelength is to consider a traveling plane wave.

Plane wave representation

In one dimension and in the complex plane, a plane wave has the form

$$\Psi(x, t) = A e^{i(kx - \omega(k)t)} = A [\cos(kx - \omega(k)t) + i \sin(kx - \omega(k)t)], \quad (2.5)$$

where A is the amplitude, $k = 2\pi/\lambda$ is the wavenumber and $\omega = \nu/(2\pi)$ is the angular frequency.

From the Planck equation Equation 1.16 we can write the energy in the form

$$E = h\nu = \hbar\omega.$$

Note that in the next section, for simplicity of visualization, we will be using the real part of the plane wave

$$\Psi(x, t) = A \cos(kx - \omega t), \quad (2.6)$$

but we will get back to using Equation 2.5 later on. Considering *complex* wavefunctions in quantum mechanics is fundamental due to the probabilistic nature of quantum physics. While in classical mechanics the complex part of waves is “physical”, in quantum mechanics wavefunctions are intrinsically complex because they are used to define the probability and introduce phase factors that are not directly measurable, but have measurable consequences such as the interference in the double-slit experiment.

2.4.1.1 Dispersion relations

From Equation 2.5 we see that the frequency ω depends on the wavenumber k . The functions relating the time-dependence and space-dependence of physical entities, with either $\omega(k)$ or $E(p)$, is given by the *dispersion relations*.

Dispersion relation for non-dispersive systems For non-dispersive systems, the dispersion relation is *linear*

$$\omega(k) = vk, \quad (2.7)$$

where v is the wave speed.

Dispersion relation for light

The dispersion relation for light (in vacuum) is also a *linear* function:

$$c = \lambda\nu = \omega/k = E/p \implies \omega(k) = ck. \quad (2.8)$$

Dispersion relation for a non-relativistic particle

For a non-relativistic particle, using Equation 1.28, we can derive the relation between the momentum and the wavenumber:

$$p = \hbar k,$$

and combining this with the definition of kinetic energy and the Planck equation

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

we get the (*quadratic*) dispersion relation for a Newtonian (non-relativistic) particle

$$\omega(k) = \frac{\hbar k^2}{2m}. \quad (2.9)$$

Dispersion relation for a relativistic particle

For a relativistic particle, the energy is

$$E = \sqrt{p^2 c^2 + m^2 c^4} - mc^2.$$

For $p \ll mc$, this reduces to

Plane waves derivatives

The space and time derivatives of plane waves are

$$\begin{aligned} -i\hbar \frac{\partial \Psi}{\partial x} &= \hbar k \Psi = p\Psi, \\ i\hbar \frac{\partial \Psi}{\partial t} &= \hbar \omega(k) \Psi = E\Psi. \end{aligned}$$

2.4.1.2 Issues with free particles as plane waves

If the plane wave describes the wavefunction for a free particle, it means that since this must be continuous and extending to infinity with the same amplitude, we could find the particle *everywhere* in space with the same probability ($|\Psi(x, t)|^2 = A^2$), which does not seem right. In fact, if we consider the uncertainty on the position of the particle, Δx , as being (roughly) the region across which the wavepacket extends, we can see (Figure 2.7 (a)) that this is a *big* uncertainty - so big that it goes to infinity! The plane wave has the same intensity everywhere and it is *fully delocalised*.

There is another issue... Let's try to impose the normalisation condition considering a fixed time, e.g. $t = 0$ for simplicity:

$$\int_{-\infty}^{\infty} |\Psi(x, t)|^2 dx = |A|^2 \int_{-\infty}^{\infty} \cos^2(kx) dx \rightarrow |A|^2 \times \infty.$$

In theory this should be equal to 1 in order to be normalised, but we see that there is a problem: this integral does not converge, so this wavefunction *can't be normalised*. This means that we can't describe the free particle simply as a plane wave because it would be *unphysical*.

It seems that having one single wavelength (or wavenumber, or momentum value), does not really do the job. What we would need is something to make this wave more "localised" in space. Remember however that the wavefunction still needs to extend across the whole space, from $-\infty$ to ∞ , but naturally go to 0 beyond a certain range where we expect to find

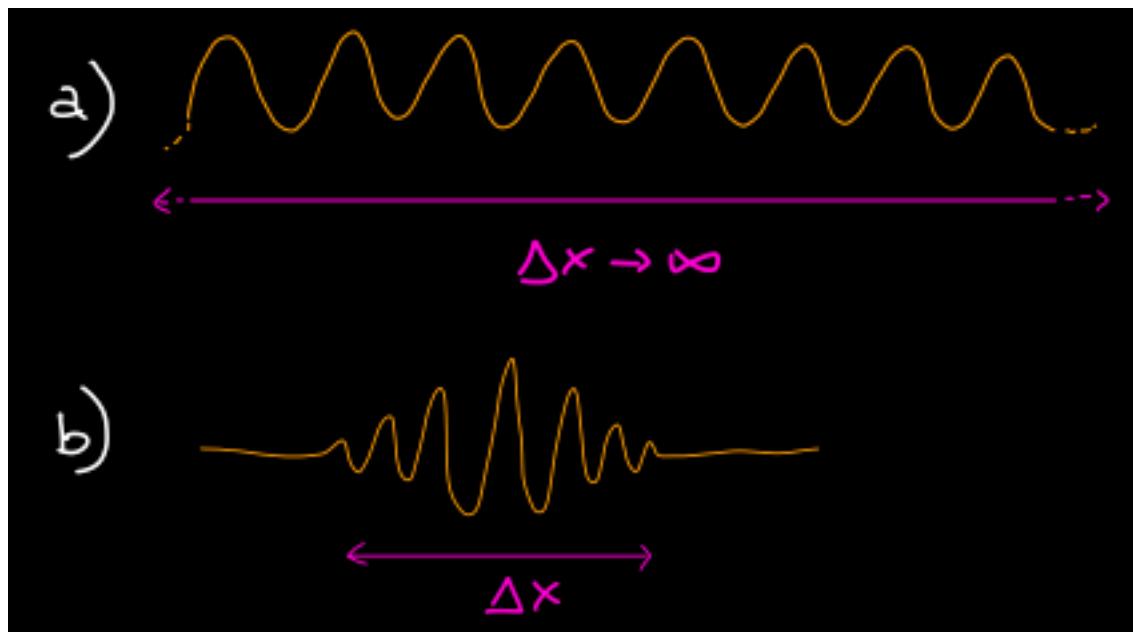


Figure 2.7: (a) Wavefunction of the free particle as a plane wave - the uncertainty in position extends to infinity; (b) Wavepacket “localised” in a length Δx .

the particle (remember that the wavefunction has to be also smooth, continuous and single-valued). How do we obtain something that has zero amplitude somewhere and is contained in a wavepacket of length Δx , like the one shown in Figure 2.7 (b)?

As Feynman says: let's add more waves!

2.4.2 Superposition of two plane waves

Let's try to add more waves having different wavelengths (corresponding to more values of momentum based on the de Broglie wavelength).

Let's start by adding two waves of different wavelengths. What do we obtain? Look at the example below that considers this, showing two waves and the resulting one obtained using the superposition principle.

Go to [this link](#) if the resource embedded below does not work.

The resultant wave is now made of “beats”, resulting in “wave groups” with smaller extension due to constructive and destructive interference. We're still not in the ideal case of having one *localised wavepacket*, but it seems that adding more waves (and more momenta) helps and reduces the uncertainty of the position. It is worth noting that by adding two waves of different wavelengths, because of the de Broglie wavelength, we have two possible values of momentum

that we could observe, which means that we *increased the uncertainty of momentum*, and this leads to a *reduced uncertainty of the position*. This is not a case, and we will make more clarity on this by the end of the chapter.

We also see that adding two waves with the same amplitude and different wavenumbers, we obtain a series of wave “envelopes” travelling at a group velocity that is different from the velocity of the individual waves. This will be clarified in the following.

Let’s see how mathematically we get these “envelopes” by considering the sum of (the real part of) two plane waves with equal amplitude A , different wavenumbers k_1 and $k_2 = k_1 + \Delta x$ and different frequencies ω_1 and $\omega_2 = \omega_1 + \Delta\omega$:

$$\Psi(x, t) = \Psi_1(x, t) + \Psi_2(x, t) = A[\cos(k_1 x - \omega_1 t) + \cos(k_2 x - \omega_2 t)]. \quad (2.10)$$

Using the trigonometric identity $\cos a + \cos b = 2 \cos(\frac{1}{2}(a - b)) \cos(\frac{1}{2}(a + b))$, from Equation 2.10 we obtain

$$\Psi(x, t) = 2 \cos\left[\frac{1}{2}(\Delta k x - \Delta\omega t)\right] \cos\left[\frac{1}{2}(2k + \Delta k)x - \frac{1}{2}(2\omega + \Delta\omega)t\right], \quad (2.11)$$

and since $\Delta k \ll k$ and $\Delta\omega \ll \omega$, the last cosine term in Equation 2.11, describing the wave *within the envelope* is $\cos[\frac{1}{2}(2k + \Delta k)x - \frac{1}{2}(2\omega + \Delta\omega)t] \approx \cos[kx - \omega t]$. Therefore the total wavefunction is

$$\Psi(x, t) = \Psi_1(x, t) + \Psi_2(x, t) = 2 \cos\left[\frac{1}{2}(\Delta k x - \Delta\omega t)\right] \cos[kx - \omega t], \quad (2.12)$$

giving the envelope illustrated in Figure 2.8 for $t = 0$.

2.4.2.1 Phase and group velocities

The wave within the envelope moves at the **phase velocity** given by the term $\cos[kx - \omega t]$

$$v_p = \frac{(2\omega + \Delta\omega)/2}{(2k + \Delta k)/2} \approx \frac{\omega}{k} = \frac{\omega(k)}{k}. \quad (2.13)$$

The other cosine term in Equation 2.12 gives the velocity of the whole envelope, also known as **group velocity**

$$v_g = \frac{\Delta\omega/2}{\Delta k/2} = \frac{\Delta\omega}{\Delta k}. \quad (2.14)$$

In the following section we will see a more accurate form of the velocity group, obtained in the limits for $\Delta\omega \rightarrow 0, \Delta k \rightarrow 0$.

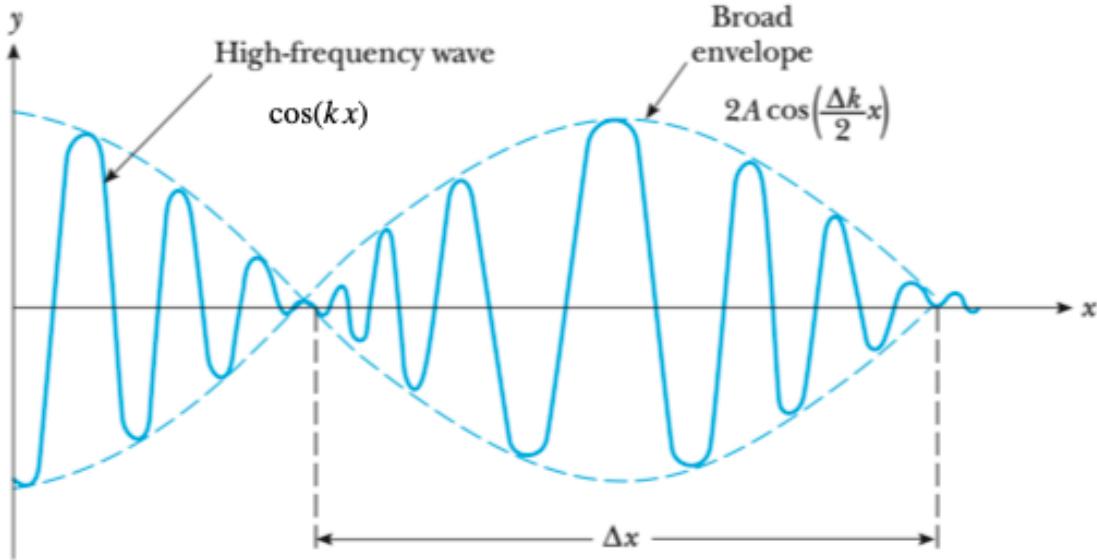


Figure 2.8: Resultant wavefunction of Equation 2.12, for $t = 0$.

2.4.3 Adding more waves to form a wavepacket

So how do we create a *single localised wavepacket* and reduce the surrounding noise?

- Add *more waves* of different frequencies (more values of momentum/wavenumber)
- Choose conveniently the *distribution of the amplitudes* of each wave component.

Play with the simulator below: try to find ways to have further wavepackets with small noise between them.

In the above figures we kept increasing the number of modes (waves at different wavenumbers) and see that adding more waves increases the wavepackets, while changing their weights in the sum (by changing their amplitudes) reduces the noise between wavepackets.

Now, you get the idea. We can keep increasing the number of waves to eventually isolate a wavepacket. The more we add, the better, so let's add up a continuum of waves in the k -space (the space of the wavenumbers), with amplitudes dependent on k , obtaining the integral below.

Particle represented by a localised wavepacket

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(k) e^{i(kx - \omega(k)t)} dk, \quad (2.15)$$

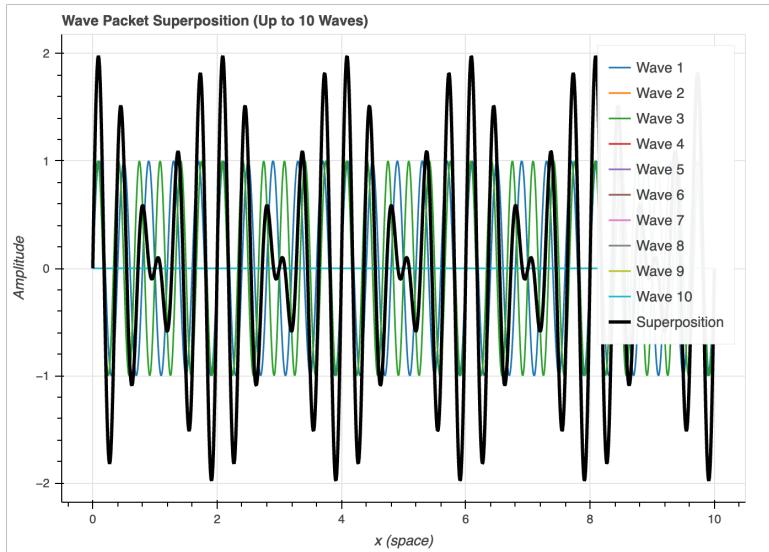


Figure 2.9: Superposition of 2 waves with wavenumbers and amplitudes of Figure 2.10.

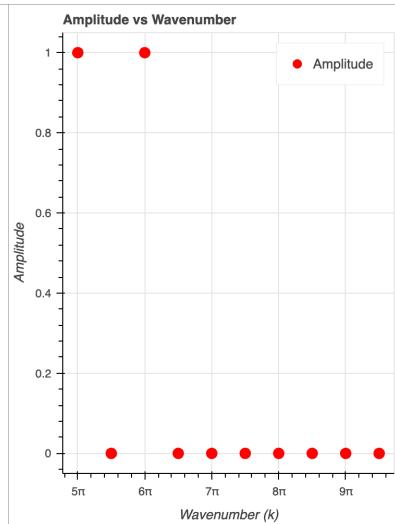


Figure 2.10: Amplitudes vs wavenumber used for the waves of Figure 2.10.

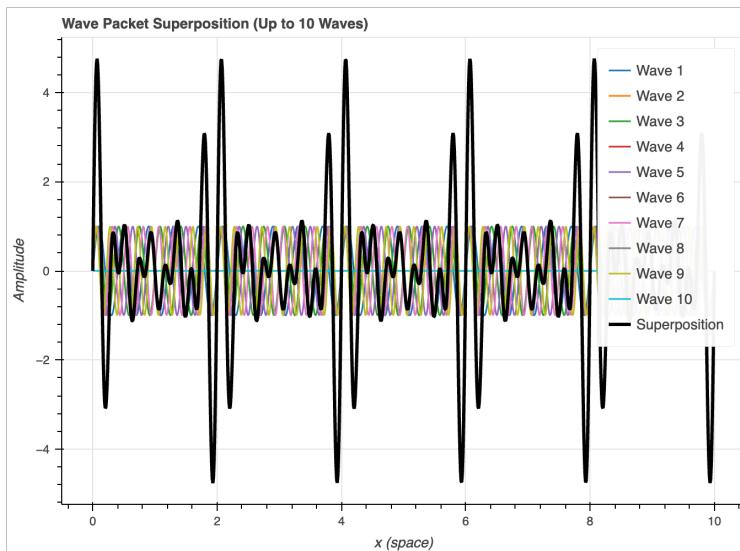


Figure 2.11: Superposition of 5 waves with wavenumbers and amplitudes of Figure 2.12.

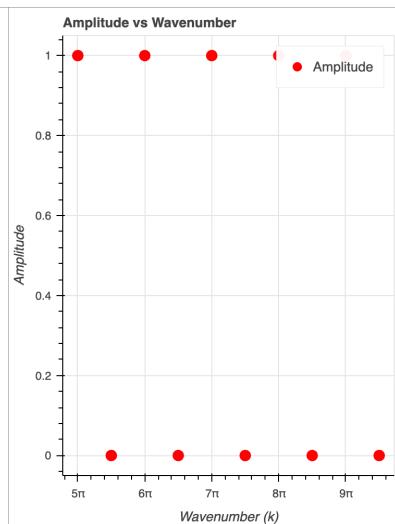
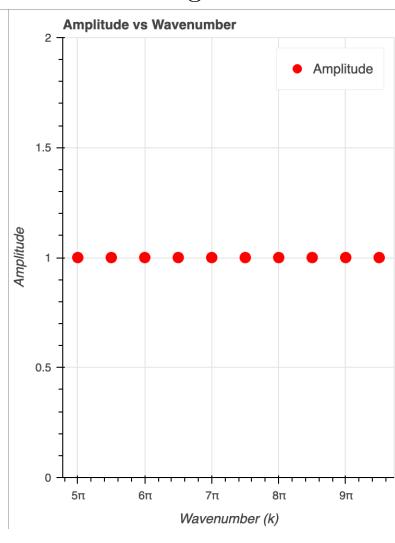
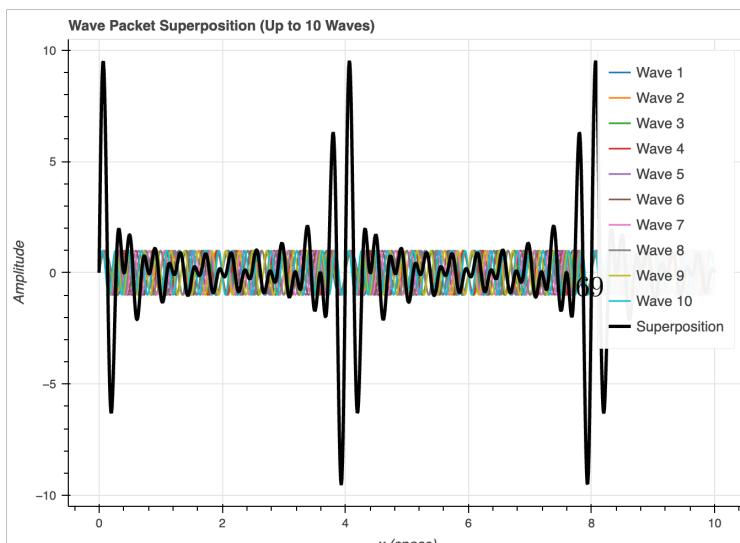


Figure 2.12: Amplitudes vs wavenumber used for the waves of Figure 2.12.



where $\phi(k)$ are the amplitudes in the k -space and provide the so-called *spectral content* of the wavepackets.

$\Psi(x)$ is the *Fourier transform (FT)* of $\phi(k)$, which in fact has the role of a wavefunction too, but in the k -space instead of the position space. So the FT is useful to go back and forth between these two spaces, without loss of information.

! Important

The broader the function $\phi(k)$, the narrower is the wavepacket $\Psi(x, 0)$, satisfying $\Delta k \Delta x \approx 1$ if $\phi(k)$ and $\Psi(x)$ are Gaussians of widths Δk and Δx respectively. This is not a coincidence, as position and wavenumber (or momentum) are **conjugate variables**, so increasing one will decrease the other. We will expand on this at the end of the chapter.

Note that at the time $t = 0$, Equation 2.15 gives

$$\Psi(x, 0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(k) e^{ikx} dk. \quad (2.16)$$

2.4.4 Wavepacket dispersion

We have seen that the velocity of each wave (or phase velocity) is, from Equation 2.13, $v_p = \frac{\omega(k)}{k}$.

The whole wavepacket however moves at the group velocity, which for the localised wavepacket of Equation 2.15 can be obtained from Equation 2.14 in the limit $\Delta k \rightarrow 0, \Delta\omega \rightarrow 0$:

$$v_g = \frac{d\omega(k)}{dk}. \quad (2.17)$$

For the wavepacket to travel without dispersion (i.e. no distortion), the phase and group velocities must be equal: $v_p = v_g$.

For a *non-dispersive medium* the dissipation relation Equation 2.7 gives a constant phase velocity equal to the group velocity: $v_p = v = v_g$ and the wavepacket doesn't disperse.

However, for a dispersive medium with $\omega(k) \propto k^2$ as in Equation 2.9, we would obtain $v_p = \frac{\hbar k}{2m}$, so the components of waves with smaller wavenumbers would move more slowly. In this case the group velocity would be

$$v_g = \left. \frac{d\omega(k)}{dk} \right|_{k=k_0} = \frac{\hbar k_0}{m} = \frac{p_0}{m}, \quad (2.18)$$

We can see that the group velocity coincides with that of a particle of mass m and momentum p_0 , for a wavepacket centered around $k = k_0$ in k -space.

Since this is made from a superposition of waves moving at different speeds, we can expect that the wavepacket becomes *distorted* with time, as the group and the phase velocities are not equal: $v_p \neq v_g$ and the wavepacket is dispersive.

2.4.5 Conservation of probability

Going back to the definition of the wavepacket of Equation 2.15, having seen that this is the FT of the wavefunction in the k -space, we can use the Parseval identity to relate the probabilities in position-space and k -space:

$$\int_{-\infty}^{\infty} |\Psi(x)|^2 dx = \int_{-\infty}^{\infty} |\phi(k)|^2 dk. \quad (2.19)$$

From this you can notice that the wavefunction must be normalised also in k -space. We can give to $|\phi(k)|^2$ the same meaning of probability in the k -space that we gave to $|\Psi(x)|^2$ in position space.

Starting from the total probability in position space $\int dx |\Psi(x, t)|^2$, we can demonstrate that the normalization in the wavenumber (or momentum) k -space is preserved even if we have time evolution:

$$\int \Psi^*(x, t)\Psi(x, t)dx = \frac{1}{2\pi} \int dx \int dk \phi^*(k) e^{-i(kx - \omega(k)t)} \int dk' \phi(k') e^{i(k'x - \omega(k')t)}, \quad (2.20)$$

and using the definition and property of the Dirac delta function

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dx e^{-i(k-k')x} = \delta(k - k') \quad (2.21)$$

$$\int f(x) \delta(x - x_0) dx = f(x_0), \quad (2.22)$$

Equation 2.20 gives

$$\int |\Psi(x, t)|^2 dx = \int dk \int dk' \delta(k - k') e^{i(\omega(k) - \omega(k'))t} \phi^*(k) \phi(k') = \int dk |\phi(k)|^2 = 1. \quad (2.23)$$

i Note

Note that the probability of finding the particle at a certain value of k , $|\phi(k)|^2$, is not changing with time: the evolution of each definite- k (or definite- ω) is just a phase $e^{i\omega(k)t}$ that cancels out using the Dirac delta properties.

2.5 Reciprocity in wavepackets and the Heisenberg uncertainty principle

We will now go more in details in concept that we anticipated in the previous few sections, and will see how the wave formalism that we introduced, with the FTs, leads to the Heisenberg's uncertainty principle.

To construct the wavepacket for a free particle, the first thing we tried was to use a single plane wave, but we saw that it had an ill-defined position, extending to infinity, i.e. $\Delta x \rightarrow \infty$. We then mentioned that the position and the wavenumber (and so the momentum too) are conjugate variables through the Fourier transform: if you increase one, the other one will decrease. Now we have introduced enough mathematical notation to see that with the FT properties: if we Fourier-transform a single-frequency plane wave, we obtain the Dirac delta-function of Equation 2.21.

The same goes, in theory, if we consider a plane wave in k -space: we will have a Dirac delta in position space (extremely localised).

However, once we start adding more components in k -space (or position space), we narrow down the localisation in the complementary position (or wavenumber/momentum) space, as shown in Figure 2.17.

So how do we get a good trade-off between localisation in position and momentum space, minimizing both uncertainties Δx and Δk (or Δp)?

Remember the simulation and the outputs of Figure 2.15, where we found that adding more components with different wavenumbers and a certain distribution of amplitudes $\phi(k)$ gives more distinct and “clean” wave groups.

It turns out that when we go to the continuum to build the localised wavepacket from the FT of plane waves in momentum space, *the best function of $\phi(k)$ that can be used to obtain a good wavepacket that minimizes uncertainty in both position and momentum space is a Gaussian!*

2.5.1 Gaussian wavepackets

Let's consider the following Gaussian wavepacket with a standard deviation $\sigma(k) = 1/\sqrt{4\alpha}$ centered about $k = k_0$ in momentum space:

$$\phi(k) \sim e^{-\alpha(k-k_0)^2}. \quad (2.24)$$

Why a Gaussian?

The FT of a Gaussian is still a Gaussian!

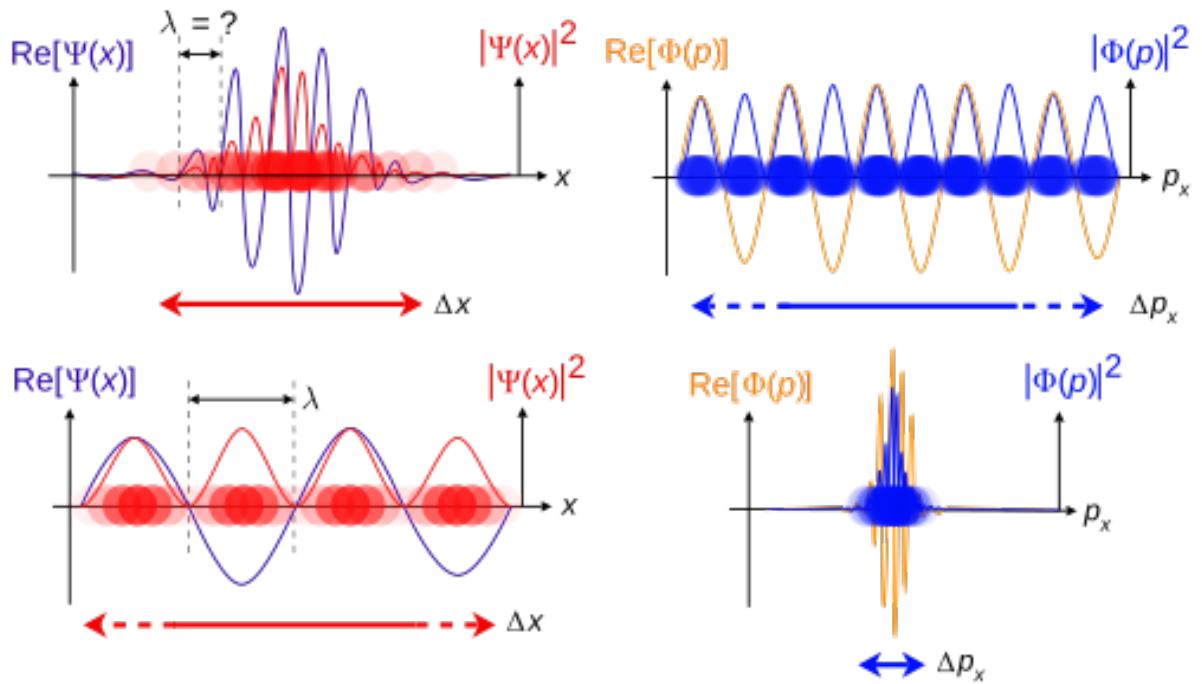


Figure 2.17: Position x and momentum $p = \hbar k$ wavefunctions corresponding to quantum particles. The colour opacity of the particles corresponds to the probability density of finding the particle with position x or momentum component p . Top: If wavelength λ is unknown, so are momentum p , wave-vector k and energy E (de Broglie relations). As the particle is more localized in position space, Δx is smaller than for Δp_x . Bottom: If λ is known, so are p , k , and E . As the particle is more localized in momentum space, Δp is smaller than for Δx . Source: [Wikipedia](#)

From Equation 2.15, using the Gaussian amplitude distribution of Equation 2.24:

$$\Psi(x) = \frac{1}{\sqrt{2\pi}} \int dk e^{-\alpha(k-k_0)^2} e^{ikx},$$

and defining $K = k - k_0$ this gives

$$\Psi(x) = \frac{e^{ik_0 x}}{\sqrt{2\pi}} \int dK e^{-\alpha(K^2 - iKx/\alpha)}.$$

Using the trick of “completing the square”, the above equation becomes

$$\Psi(x) = \frac{e^{ik_0 x}}{\sqrt{2\pi}} \int dK e^{-\alpha(K - iKx/2\alpha)^2} e^{-x^2/4\alpha}.$$

After another change of variables $K' = K - ix/2\alpha$:

$$\Psi(x) = \frac{e^{ik_0 x} e^{-x^2/4\alpha}}{\sqrt{2\pi}} \int dK' e^{-\alpha K'^2} = \frac{e^{ik_0 x} e^{-x^2/4\alpha}}{\sqrt{2\alpha}}, \quad (2.25)$$

where in the last passage we used the Gaussian integral $\int_{-\infty}^{\infty} dz e^{-\alpha z^2} = \sqrt{\frac{\pi}{\alpha}}$.

Note that the exponential term with the central momentum k_0 is just a phase factor, since all the k are integrated.

The resulting Gaussian of Equation 2.25 has a standard deviation $\sigma(x) = \sqrt{\alpha}$.

The product of the widths of the two Gaussian in k -space and x -space is therefore a constant independent of α , showing that the relation is reciprocal:

$$\sigma(x)\sigma(k) = \sqrt{\alpha} \frac{1}{\sqrt{4\alpha}} = \frac{1}{2}. \quad (2.26)$$

If we associate the uncertainty on position and momentum to the above standard deviations, i.e. $\sigma(x) \rightarrow \Delta x$ and $\sigma(k) \rightarrow \Delta k$, and use the wavenumber definition $k = p/\hbar$, Equation 2.26 becomes

$$\Delta x \Delta p = \frac{\hbar}{2}. \quad (2.27)$$

Now, this is something that looks like the well known Heisenberg's uncertainty principle, except for the fact that it is an equality! In fact, earlier on we mentioned that using Gaussian for the amplitude functions (either in position or momentum space) gives the *minimum uncertainty* in position and momentum, i.e. the equality of Equation 2.27.

If we use any other function, we would get less precision on the uncertainties, which are related by the Heisenberg uncertainty below.

Note

Heisenberg uncertainty relation

$$\Delta x \Delta p \geq \frac{\hbar}{2}. \quad (2.28)$$

The same kind of relation can be applied to *any other conjugate variables*, such as frequency ω (or energy E) and time t :

$$\Delta E \Delta t \geq \frac{\hbar}{2}. \quad (2.29)$$

In the next section we will focus on the physical meaning and interpretation of the Heisenberg uncertainty principle.

2.5.2 Heisenberg's uncertainty principle - implications

The Heisenberg's uncertainty relation is central in quantum mechanics and explains things such as the breakdown of the double slit on observation, or the natural spectral linewidths, through Equation 2.29. It was originally derived by Heisenberg in his matrix mechanics formalism and we shall see it again in the context of operator algebra. In some textbooks/sources you may find \hbar instead of $\hbar/2$: the factor of 2 is not deeply meaningful, the Planck constant is.

Qualitatively, the effects are profound: it tells us that if position is constrained then momentum must become uncertain and vice versa, and the general scaling relationship between the two. Many thought experiments can be constructed around attempts to circumvent this limit, always failing due to some consequence of the variables' conjugacy.

- It is also responsible for avoiding the in-spiral and collapse of the Bohr atom: as an electron falls in and becomes localised near the nucleus at $r = 0$, its root-mean-square momentum increases, and hence so does its kinetic energy: the ground state of minimum energy is not located at $r = 0$, but instead at a finite distance and electrostatic potential.
- The Hesenberg's uncertainty principle is also central in the natural spectral linewidths, which arises from Equation 2.29.

2.5.2.1 Uncertainties in the double-slit experiment

The Heisenberg uncertainty relation also explains why the interference pattern disappears in the double-slit experiment, if we *observe* the path of individual particles. This is in short because the observation of the position (which reduces the uncertainty Δx) will increase the uncertainty in the momentum sufficient to destroy the interference pattern. Let's see this a bit more in detail.

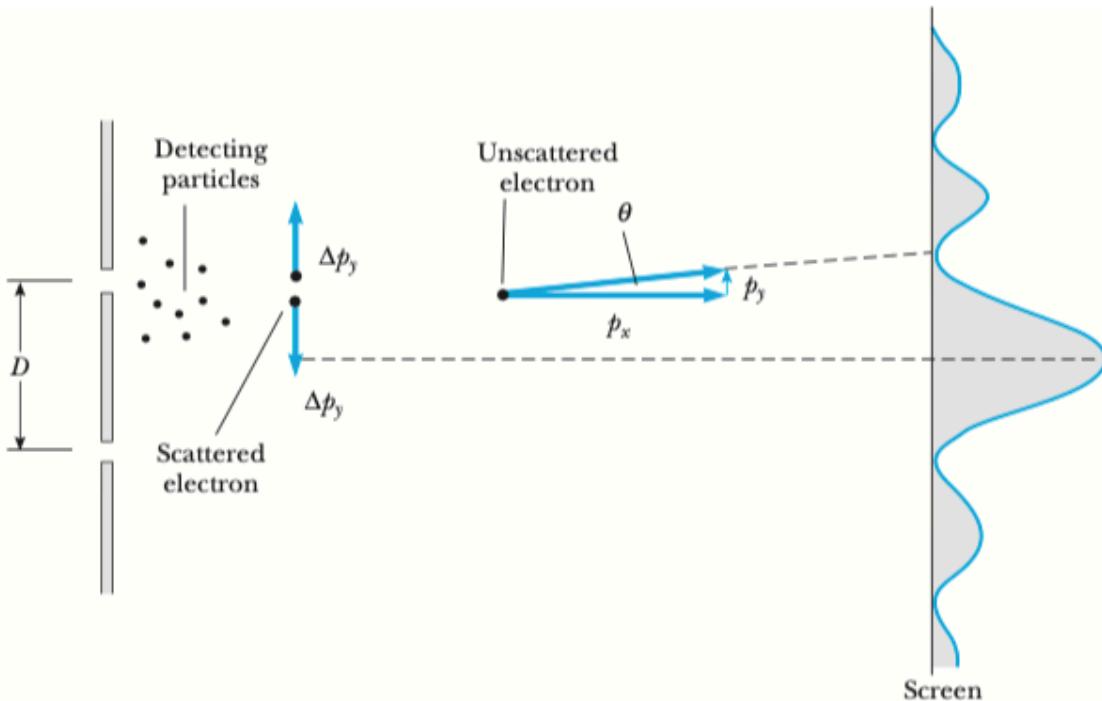


Figure 2.18: Setup of the double slit experiment to (try to) determine which slit the electron goes through.

In order to understand which slit the particle goes through, we place some particles right behind the slits, as shown in Figure 2.18, to use the recoil of small particles to have information on the path taken by electrons. To see which slit each electron goes through, we must have an uncertainty on the detecting particle's position

$$\Delta y \ll D. \quad (2.30)$$

Also, during the collision, the detecting particle changes its momentum of a quantity Δp_y , equal and opposite to the electron's change in momentum. If the electron is not deviated by an observation, it hits the screen producing the interference pattern having a first minimum given by (see Figure 2.18)

$$\tan \theta \approx \theta = \frac{p_y}{p_x} = \frac{h}{2p_x D}.$$

In order to preserve the interference pattern we should therefore have

$$\frac{\Delta p_y}{p_x} \ll \theta = \frac{h}{2p_x D} \Rightarrow \Delta p_y \ll \frac{h}{2D}$$

both for the electron and the detecting particle (since their change in momentum is the same). Therefore, from Equation 2.30 we would obtain

$$\Delta p_y \Delta y \ll \frac{h}{2D} D = \frac{h}{2},$$

which clearly violates the Heisenberg uncertainty principle Equation 2.28! In fact, any possible experiment we can imagine aimed at determining which slit the electron goes through, would give an uncertainty in momentum that is too big to still obtain the interference pattern.

2.5.2.2 Uncertainties in the harmonic oscillator

This will be more clear when we do the quantum harmonic oscillator, but for now we can anticipate that the ground state (state of minimum energy) of the quantum harmonic oscillator has non-zero energy, and this is due to the Heisenberg's uncertainty principle.

The (classical) harmonic oscillator is given by the following energy:

$$E = T + V = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2,$$

where the first term is the kinetic energy and the quadratic term is the potential energy.

Qualitatively: since the potential is harmonic, if the particle was exactly at the bottom of the potential with zero uncertainty, paradoxically the kinetic term related to the momentum would have to go to infinity because of the uncertainty principle.

For a quantitative description, we will come back to this after doing the quantum harmonic oscillator.

2.6 Summary

- Matter and light exhibit both wave-like and particle-like properties, depending on the experimental conditions (e.g., electrons in the double-slit experiment).
 - The wavefunction $\Psi(x, t)$ is a complex function describing the quantum state of a particle, and the probability density of finding the particle at the position x at time t is proportional to $|\Psi(x, t)|^2$.
 - Quantum states can exist in a superposition, where the total probability amplitude is the sum of individual amplitudes (e.g., $\Psi = \Psi_1 + \Psi_2$ in the double-slit experiment).
 - A single plane wave is not physically realistic to describe a free particle as this would be infinitely delocalised; we need to add an infinite number of weighted plane waves to form a wavepacket localised in space, obtaining the FT of a wavefunction in momentum space.
 - There is a fundamental limit to how precisely two *conjugate variables* A and B (e.g. momentum and position) can be measured *simultaneously*: $\Delta A \Delta B \geq \frac{\hbar}{2}$, i.e. when one is wide, the other is narrow.
-

2.7 References

3 The Schrödinger equation and some QM formalism

Learning objectives

1. Understand the role of the time-dependent and time-independent Schrödinger equations and their applications
2. Familiarise yourself with the concepts of operators in quantum mechanics, their mathematical structure and physical meaning
3. Understand the conservation of probability and its implications for the wavefunction
4. Evaluating expectation values and their time evolution
5. Understand the meaning of expectation values and the connection between quantum physics and the classical limit

The material of this lecture is mostly covered in Chapter 3 of Bransden-Joachain and [1] and Chapter 2 of Griffiths [3].

3.1 The importance of the superposition principle

Before getting into this chapter, we need a premise to clarify why I'll go back and talk about single plane waves after all we discussed in the previous chapter. In short: **why will we keep using single plane waves in the following?**

In the previous chapter we saw that we could in principle describe a free particle using a single plane wave (Equation 2.6), but this leads to complication, such as infinite delocalisation in space. We saw that to get around this (and other) issues, we can consider a sum (or *linear combination*) of wavefunctions with different weights, and we saw that in the continuum, for infinitely many waves, this leads to a Fourier transform of a wavefunction in momentum space (Equation 2.15).

There is one important principle that we need to keep in mind.

! Important

If two wavefunctions Ψ_1 and Ψ_2 are both solutions of a wave equation which is **linear** and **homogeneous**, the **superposition principle** ensures that any linear combination $c_1\Psi_1 + c_2\Psi_2$ (with constants c_1 and c_2), is also a solution of the same wave equation.

This can be extended to as many waves as we want! So in general it is enough to find the solution to a wave equation as a plane wave than as a wavepacket.

3.2 The time-dependent Schrödinger equation

The energy of a free particle is given by its kinetic energy. For a particle of mass m , moving in one direction x with momentum p_x , we can write

$$E = \frac{p_x^2}{2m}. \quad (3.1)$$

Let's consider a plane wave to describe it...

$$\Psi(x, t) = Ae^{i(kx - \omega t)} = Ae^{i(p_x x - Et)/\hbar}, \quad (3.2)$$

where we used the relations $k = p_x/\hbar$ and $\omega = E/\hbar$.

If we consider the wavefunction at the time $t = 0$, $\Psi(x, 0) = Ae^{ikx} = Ae^{ip_x x/\hbar}$ we can see from Equation 3.2 that the time evolution of a wavefunction $\Psi(x)$ of definite energy can be obtained from

$$\Psi(x, t) = e^{-iEt/\hbar}\Psi(x, 0), \quad (3.3)$$

so during the time evolution of a quantum state having a defined energy *only the phase is evolving*, and the state $\Psi(x, t)$ is called a **stationary state**, satisfying the condition

$$|\Psi(x, t)|^2 = |\Psi(x, 0)|^2. \quad (3.4)$$

If we differentiate Equation 3.2 with respect to time, we obtain

$$\frac{\partial \Psi(x, t)}{\partial t} = -\frac{iE}{\hbar}\Psi(x, t), \quad (3.5)$$

and if we differentiate twice with respect to x , we find

$$\frac{\partial^2 \Psi(x, t)}{\partial x^2} = -\frac{p_x^2}{\hbar^2}\Psi(x, t). \quad (3.6)$$

Using Equation 3.1, we see that Equation 3.2 satisfies the partial differential wave equation

$$i\hbar \frac{\partial}{\partial t} \Psi(x, t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi(x, t). \quad (3.7)$$

This is the **one-dimensional time-dependent Schrödinger equation for a free particle**.

Now, this wave equation is *linear* and *homogeneous*, so a linear superposition of plane waves, such as the wavepacket of Equation 2.15, will still be a solution of Equation 3.7.

We can generalise the Schrödinger equation of Equation 3.7 to the case where a particle is not free, but it is subject to a time-independent potential $U(x)$, hence it is acted on by a force $F = -\frac{dU}{dx}$.

In this case, we get the following form of the *one-dimensional time-dependent Schrödinger equation for a particle moving in a potential*.

One-dimensional time-dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \Psi(x, t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi(x, t) + U(x) \Psi(x, t), \quad (3.8)$$

where $U(x)$ is a time-independent potential.

Whilst we are considering a time-independent potential here, we can have an even more general form of the Schrödinger equation with a time-dependent potential $U(x, t)$. In this course however we will mostly focus on the case of time-independent potentials.

i Note

Note that the time-dependent Schrödinger equation (TDSE)(Equation 3.8) contains a *first order time derivative* $\partial/\partial t$, which means that if the initial value of the wavefunction is given at an initial time $t = t_0$, i.e. $\Psi(x, t_0)$, then the value of the wavefunction at any later time t can be found by solving the TDSE.

Therefore the fundamental task is to obtain a solution to $\Psi(x, t)$ to the Schrödinger equation (Equation 3.8) that satisfies the given initial conditions.

3.2.1 Operators in the Schrödinger equation

Let's start to introduce some words that we will use during the whole course and that we will describe more in detail later in the chapter.

The first thing it is useful to introduce here is the concept of **operator**, which are mathematical entities that transform one wavefunction into another.

The term on the left hand side of Equation 3.7, $i\hbar \frac{\partial}{\partial t}$, applied on the wavefunction, is the *total energy operator* - a mathematical function (of time, t) acting on the wavefunction that returns the total energy, and we indicate with the symbol \hat{E} . To indicate operators we use the hat symbol.

Total energy operator

$$\hat{E} = i\hbar \frac{\partial}{\partial t} \quad (3.9)$$

This operator yields the same result of another operator that also gives the total energy of the system, but acts on the spatial coordinates instead of time. This is the the *Hamiltonian operator*, \hat{H} , which describes the sum of the potential and kinetic energy operators, respectively $\hat{T} = \frac{-\hbar^2}{2m} \frac{\partial^2}{\partial x^2}$ and \hat{U} .

Hamiltonian operator

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + U(x, t). \quad (3.10)$$

For the right hand side of Equation 3.7 we can also define the *momentum operator* \hat{p} (in 1D) as below.

Momentum operator (1D)

$$\hat{p}_x = -i\hbar \frac{\partial}{\partial x}. \quad (3.11)$$

We will talk more about operators later in the chapter.

Having introduced this formalism, we can cast the TDSE of Equation 3.8 in the following alternative forms:

$$i\hbar \frac{\partial}{\partial t} \Psi(x, t) = \hat{H} \Psi(x, t), \quad (3.12)$$

$$\hat{E} \Psi(x, t) = \hat{H} \Psi(x, t). \quad (3.13)$$

3.2.2 Exercise

Test the time-dependent Schrödinger equation with a plane wave of amplitude A and a flat potential $U(x) = U$. What is the energy obtained?

Solution

$$\Psi(x, 0) = Ae^{i(kx - \omega t)}.$$

– Add rest of solution –

3.2.3 Continuity condition

The following is important and clarifies why in the previous chapter we imposed certain conditions on the wavefunctions.

If the potential $U(x)$ is a continuous function of x , then every function in Equation 3.8, i.e. $\Psi(x, t), \partial\Psi/\partial t, \partial\Psi/\partial x$, must be continuous.

If $U(x)$ shows any finite discontinuities (jumps) with x , then $\partial^2\Psi(x)/\partial x^2$ should also have corresponding finite jumps. For this to happen, we need Ψ to be continuous with x : if this were not the case, $\partial^2\Psi(x)/\partial x^2$ would be infinite at points in which Ψ changed discontinuously. The continuity of Ψ on the other end also implies that $\Psi(x, t)$ and $\partial\Psi/\partial t$ have to be continuous.

What about the continuity with time?

If $U(x, t)$ is a continuous function of t , so will be $\Psi(x, t)$ and $\partial\Psi/\partial t$. However, if $U(x, t)$ has a finite jump with time, $\partial\Psi/\partial t$, while $\Psi(x, t)$ will still be a continuous function of t .

3.2.4 The time-dependent Schrödinger equation in 3D

We can generalise the 1D TDSE to 3D as follows:

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}, t) = -\frac{\hbar^2}{2m} \nabla^2 \Psi(\mathbf{r}, t) + U(\mathbf{r}) \Psi(\mathbf{r}, t), \quad (3.14)$$

where

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad (3.15)$$

is the Laplacian operator.

It is useful now to refresh the definitions of gradient and divergence as well, as they will be used in the following section. The gradient ∇ is defined as

$$\nabla = \frac{\partial}{\partial x} \hat{i} + \frac{\partial}{\partial y} \hat{j} + \frac{\partial}{\partial z} \hat{k}, \quad (3.16)$$

and we can build the divergence from the gradient, by associating a unit vector to each partial derivative, for each direction. Remember that gradient and Laplacian act on functions,

whereas the divergence is a vector, so can be used for scalar or vector product with another vector.

3.3 Probability conservation

In the previous chapter we explained how, according to the Born's postulate, the wavefunction has the role of probability amplitude and it can be used to define the probability density (see Equation 2.2).

We have also seen that this leads to the normalization condition for the wavefunction (Equation 2.4), since the probability to find the particle over all space must be equal to one.

Let's now consider what happens to the probability as time changes. Since there is no loss of information, we expect this to be unchanged, so the wavefunction $\Psi(x, t)$ should be normalised at all times: $\int |\Psi(x, t)|^2 dx = 1$. This means that **the probability is conserved** over the whole volume, and mathematically we can write this condition, that we will demonstrate in the following, as

$$\frac{d}{dt} \int_V P(\mathbf{r}, t) d\mathbf{r} = 0.$$

Using the definition of probability density (Equation 2.2), we can write the left hand term above as

$$\frac{d}{dt} \int_V P(\mathbf{r}, t) d\mathbf{r} = \int_V d\mathbf{r} \frac{d}{dt} (\Psi^*(\mathbf{r}, t) \Psi(\mathbf{r}, t)) = \int_V d\mathbf{r} \left[\Psi^* \left(\frac{d\Psi}{dt} \right) + \left(\frac{d\Psi^*}{dt} \Psi \right) \right],$$

where we could take the derivative d/dt inside the integral because the probability function is continuous with continuous partial derivatives.

Applying the TDSE (Equation 3.8) we can cast this in the form

$$\frac{d}{dt} \int_V d\mathbf{r} P(\mathbf{r}, t) = \int_V d\mathbf{r} \frac{i\hbar}{2m} \left[\Psi^* \frac{d^2}{d\mathbf{r}^2} \Psi - \left(\frac{d^2}{d\mathbf{r}^2} \Psi^* \right) \Psi \right]. \quad (3.17)$$

Using the definitions of gradient and divergence (Equation 3.16), we can rewrite Equation 3.17 in the form

$$\frac{d}{dt} \int_V d\mathbf{r} P(\mathbf{r}, t) = \int_V \cdot \frac{i\hbar}{2m} [\Psi^* \nabla \Psi - (\nabla \Psi^*) \Psi]. \quad (3.18)$$

For convenience we can introduce the *vector field*

$$\mathbf{j}(\mathbf{r}, t) = -\frac{i\hbar}{2m} [\Psi^* \nabla \Psi - (\nabla \Psi^*) \Psi], \quad (3.19)$$

which is known as **probability current** or **probability flux**, and Equation 3.18 can be written as

$$\frac{d}{dt} \int_V d\mathbf{r} P(\mathbf{r}, t) = - \int_V \nabla \cdot \mathbf{j}(\mathbf{r}, t) d\mathbf{r} = - \int_S \mathbf{j} \cdot d\mathbf{S} = 0, \quad (3.20)$$

where in the last passage we used the divergence theorem (aka Green's theorem), stating that the integral of the divergence of a vector, over a volume V , is equal to the surface integral of the component of that vector along the outward normal, taken over a closed surface S .

Now, the last integral over the surface S in the last integral of Equation 3.20 extends to infinity, and since the wavefunction is not fully delocalised (because it can be normalised), then its extension to infinity will be zero and the integral is overall equal to zero ($\Psi \rightarrow 0$ if $\mathbf{r} \rightarrow \infty$).

Notably, from this result in Equation 3.20 we can say that the probability current respects a **conservation law**, which has the form of a continuity equation, as below.

Conservation equation for the probability

$$\frac{\partial}{\partial t} P(\mathbf{r}, t) = - \nabla \cdot \mathbf{j}(\mathbf{r}, t). \quad (3.21)$$

Physically this means that any change in probability density at a point is balanced by a flow of probability current into or out of that (differential) region.

What do we learn from this? That physical wavefunctions must have finite probability fluxes: we cannot have infinite rates of probability density moving between spatial regions (or more generally between regions in any generalised coordinates of the system).

! Important - Implications for wavefunctions

1. The term $\nabla\Psi(\mathbf{r})$ in $J(\mathbf{r})$ means that the spatial wavefunction must be *continuous*, to avoid delta-function spikes in flux;
2. $\nabla \cdot \mathbf{j}$ in the probability conservation (Equation 3.20) means that the wavefunction must be *differentiable* everywhere, $\frac{\partial\Psi}{\partial x}$ must be continuous, otherwise there would be an unphysical divergence of $\frac{\partial P}{\partial t}$ (cfr Equation 3.21).
3. Square-integrability coming from the normalization condition means that $\Psi(x) \rightarrow 0$ for $|x| \rightarrow \infty$.

There is a caveat: for infinite potentials and infinitely fast changes, the second of these rules no longer holds.

So Ψ must be continuous and differentiable (have continuous derivatives), or smooth. These requirements, that we anticipated in the previous chapter, in addition to the established normalization requirement of square-integrability, are the key to deriving the boundary conditions that allow us to find wavefunctions for various quantum systems, seen in the following chapters.

Another important result is that the probability current (Equation 3.19) is identically zero if either:

1. the wavefunction $\Psi(\mathbf{r}, t)$ is real-valued; or
2. it has a complex phase which applies uniformly to all space positions x , hence unaffected by the space derivative ∇ .

The first of these is trivial to see: if Ψ is real valued then $\Psi^* = \Psi$ and $(\nabla\Psi^*) = \nabla\Psi$, so the two terms in \mathbf{j} (cf. Equation 3.19) cancel.

The second is slightly less obvious: if $\Psi = e^{i\Phi}\psi(\mathbf{r})$, then $\Psi^*\nabla\Psi = e^{i\Phi}e^{-i\Phi}\psi\nabla\psi = \psi\nabla\psi = \Psi\nabla\Psi^*$, and again the two terms cancel.

The consequence of this second case in particular is that *stationary states, whose time-evolution is a uniform, energy-coupled phase $e^{iEt/\hbar}$, have no probability flux*. Which is a good thing: they are stationary states precisely because their probability distribution is time-invariant, so there'd better not be any probability current flowing between regions!

3.4 The time-independent Schrödinger equation

Let's go back to considering the TDSE in 1D, Equation 3.8.

If the potential $U(x)$ is time-independent, we can use the method of separation of variables for the wavefunction:

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

Then

$$\frac{\partial\Psi}{\partial t} = \psi(x)\frac{dT}{dt}, \quad \frac{\partial^2\Psi}{\partial x^2} = \frac{d^2\psi}{dx^2}T(t),$$

and we can rewrite the TDSE, Equation 3.8, as

$$i\hbar\psi(x)\frac{dT}{dt} = -\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2}T(t) + V(x)\psi(x)T(t).$$

Dividing this by $\Psi = \psi T$, we obtain

$$i\hbar \frac{1}{T(t)} \frac{dT}{t} = -\frac{\hbar^2}{2m} \frac{1}{\psi(x)} \frac{d^2\psi(x)}{dx^2} + V(x), \quad (3.22)$$

where it is clear that all the time dependence is on the left hand side and all the space dependence on the right hand side, so this is of the form: function of t = function of x, for all t and all x. This means that both sides must be equal to a constant! We call this constant E , as it represents the total energy - we can see this by comparing the operators on each side to the total energy operator and the Hamiltonian operator defined in Equation 3.9 and Equation 3.10.

From the RHS of Equation 3.22, we obtain the (1D) **time-independent Schrödinger equation**.

Time-independent Schrödinger equation

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] \psi(x) = E\psi(x), \quad (3.23)$$

which can also be written as

$$\hat{H}\psi(x) = E\psi(x) \quad (3.24)$$

using the definition of the Hamiltonian operator \hat{H} (Equation 3.10).

In the following section we will see that this is an *eigenvalue equation*.

From the LHS of Equation 3.22, we can extract

$$i\hbar \frac{1}{T(t)} \frac{dT}{t} = E \implies i\hbar \frac{dT}{t} = ET(t), \quad (3.25)$$

which is easy to solve:

$$T(t) = Ce^{-iEt/\hbar}, \quad (3.26)$$

where C is an arbitrary constant and we can set it to $C = 1$.

We can then write the general solution of a wavefunction with separable variables (again, this holds for U independent of t), as a stationary state.

Stationary states

A stationary state is a state with a *defined total energy*, E , and has the following form:

$$\Psi(x, t) = \psi(x)e^{-iEt/\hbar}. \quad (3.27)$$

The state (i.e. the wavefunction) itself is evolving with time, but the time evolution is

just a phase, hence *the probability density is time-independent*:

$$|\Psi(x, t)|^2 = \Psi^*(x, t)\Psi(x, t) = \psi^*(x)e^{iEt/\hbar}\psi(x)e^{-iEt/\hbar} = |\Psi(x, 0)|^2 \quad (3.28)$$

In fact, we have already anticipated this at the beginning of the chapter when we looked at a plane wave, Equation 3.3.

3.5 Operators, expectation values and eigenfunctions

In this section we will introduce some notation and concepts that will be *fundamental* for the whole course (and for other courses next year).

3.5.1 Operators

In Section 3.2.1 we have already anticipated what operators are.

Operators are mathematical entities that transform a wavefunction onto another wavefunction.

In fact, we have already defined the *Hamiltonian operator* (Equation 3.10), which gives the total energy of the system, and the *momentum operator* (Equation 3.11). Both of these quantities are physical *observables*, i.e. physical quantities that we can measure. In general, observable such as momentum, position, spin, angular momentum, etc, are described by operators. However, the opposite is not true! *Not all operators are observable*, and we will see an example of this when we do the quantum harmonic oscillator.

! Hermitian operators

Observables are represented by *Hermitian* operators.

Hermitian operators are operators that return *real* eigenvalues, therefore the following is valid:

$$\int \psi^* \hat{A} \psi dx = \int (\hat{A}^* \psi^*) \psi dx. \quad (3.29)$$

Physical quantities like momentum or energy can be “pulled out” of wavefunctions by applying the operators on them.

In the following you can see some examples of this, with the momentum and total energy operators (Equation 3.11 and Equation 3.9), applied on a plane wave $\Psi(x, t) = Ae^{ikx-\omega t}$.

$$\hat{p}\Psi = -i\hbar \frac{\partial \Psi}{\partial x} = -i\hbar \frac{\partial Ae^{i(kx-\omega t)}}{\partial x} = \hbar k A e^{i(kx-\omega t)} = p\Psi, \quad (3.30)$$

$$\hat{E}\Psi = i\hbar \frac{\partial \Psi}{\partial t} = i\hbar \frac{\partial Ae^{i(kx-\omega t)}}{\partial t} = \hbar\omega\Psi = E\Psi. \quad (3.31)$$

3.5.1.1 More operators

The **canonical operators**, in position-space representation, are:

- **Total energy:** $\hat{E} = i\hbar \frac{\partial}{\partial t}$
- **Position:** \hat{x} (1D) or $\hat{\mathbf{r}}$ (3D)
- **Momentum:** $\hat{p}_x = -i\hbar \frac{\partial}{\partial x}$ (1D) or $\hat{\mathbf{p}} = -i\hbar \nabla$

From these, all other operators can be constructed, including:

- **Kinetic energy:** $\hat{T} = -\frac{\hbar^2}{2m} \nabla^2$
- **Hamiltonian:** $\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + U(x, t)$
- **Angular momentum:** $\hat{L} = \hat{\mathbf{r}} \times \hat{\mathbf{p}} = -i\hbar \mathbf{r} \times \nabla$

We will encounter many more operators during the course.

3.5.2 Expectation values

In Chapter 2 we have discussed the Born's interpretation of the wavefunction, which is related to the probability density (Equation 2.2).

For each run of an experiment in quantum mechanics we can't predict exactly the outcome in a deterministic way, due to the probabilistic nature of QM, but if we run the experiment many times, we can have an *average of many measurement*, which is what we would expect to obtain in the classical limit: this is called the **expectation value**, and it is defined as follows

! Expectation values

The expectation value of an operator \hat{A} (described by a function $f(x)$), calculated on the wavefunction $\Psi(x)$, is

$$\langle \hat{A} \rangle = \int_{-\infty}^{\infty} dx \Psi^*(x) \hat{A}(x) \Psi(x) \quad (3.32)$$

Just make a “sandwich” of the operator between the wavefunction and its complex conjugate, as in Figure 3.1!

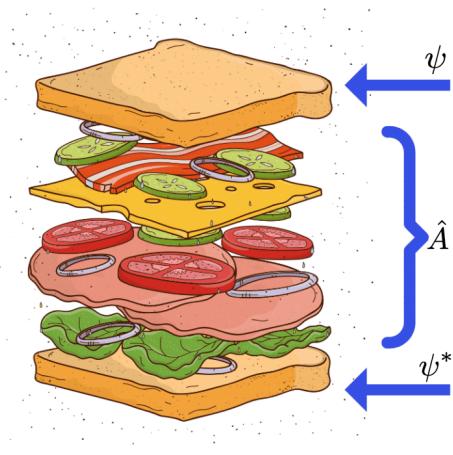


Figure 3.1: A way to think about the expectation value and remember it!

For some “normal” operators $A(x)$ (such as \hat{x}), A and Ψ commute, i.e. their order is not relevant: $A(x)\Psi(x) = \Psi(x)A(x)$. In such cases: $\langle \hat{A} \rangle = \int_{-\infty}^{\infty} dx \Psi^*(x) \hat{A}(x) \Psi(x) = \int_{-\infty}^{\infty} dx A(x) P(x)$.

However, this is not always the case! For example, the momentum operator p_x is defined by a derivative over x , so it will change the function that follows it, and the order in this case matters!

⚠ Warning

The ordering in $\langle \hat{A} \rangle = \int_{-\infty}^{\infty} dx \Psi^*(x) \hat{A}(x) \Psi(x)$ matters!

Variance of an operator

The variance of an operator \hat{A} is an important quantity that we will often encounter, and it is defined as

$$\Delta^2 A = \sigma_A^2 = \langle A^2 \rangle - \langle A \rangle^2. \quad (3.33)$$

This is useful also to demonstrate uncertainty relations.

3.5.2.1 Exercise

Calculate the expectation value $\langle x \rangle$ of a Gaussian wavefunction $\psi(x)$.

💡 Hints

A Gaussian wavefunction has the form $\psi(x) = \frac{1}{\sqrt{Z}}e^{-x^2/4\sigma^2}$, where $1/Z$ is the normalization factor for $|\Psi|^2$.

Solution

$$\langle x \rangle = \frac{1}{Z} \int_{-\infty}^{\infty} dx |e^{-x^2/4\sigma^2}|^2 = \frac{1}{Z} \int_{-\infty}^{\infty} dx e^{-x^2/2\sigma^2} = 0,$$

because the function has odd symmetry.

Similarly, you can demonstrate that $\langle x^2 \rangle = \sigma^2$. Do it as a homework!

We will evaluate expectation values quite a lot, and it will be useful to remember that the variance Δ^A (or $\sigma^2(A)$) of an operator \hat{A} is given by Equation 3.33.

3.5.2.2 Exercise

Calculate the expectation value $\langle p \rangle$ of a Gaussian wavefunction $\psi(x)$.

💡 Hints

Hint 1

The starting point is the same as the example above.

Hint 2

Be careful to the order, now it matters!

Solution

– ADD SOLUTION –

3.5.3 Eigenfunctions

Before we explain what eigenfunctions are, let's refresh some good ol' linear algebra stuff (please revise this as you will need it a lot!).

3.5.3.1 Remember eigenvectors and eigenvalues?

Let's go back to some maths that you should know well, about matrices, eigenvectors and eigenvalues.

Consider a matrix \mathbb{A} acting on a vector \mathbf{v} . If

$$\mathbb{A}\mathbf{v} = \lambda\mathbf{v}, \quad (3.34)$$

with λ a constant, then this is called the **eigenvalue equation**, where:

- \mathbf{v} is an eigenvector of \mathbb{A} ,
- λ is the eigenvalue.

What does this mean? It means that the application of the matrix A on the vector v returned the vector v (in the same direction), but rescaled, with a scaling constant λ .

You can check a visualization of this in the interactive resource embedded below, or found at [this link](#).

3.5.3.2 Eigenfunctions (and eigenstates)

Why do I dig out now the eigenvectors and eigenvalues stuff from the “forgotten-maths-I-thought-I-no-longer-needed” pit? (Spoiler: you will need this a lot...)

Because earlier I mentioned that the time-independet Schrödinger equation (Equation 3.24) is an *eigenvalue equation*.

Compare this with the eigenvalue equation in Equation 3.34. Do you see the similarities? Instead of having a matrix \mathbb{A} and a vector \mathbf{v} , we have an operator \hat{H} and a wavefunction $\psi(x)$, but the TISE tells me that if I apply the Hamiltonian operator \hat{H} on the wavefunction $\psi(x)$, I get the same wavefunction $\psi(x)$, but *rescaled* by a real number which is given by the energy E !

So in analogy with eigenvalues and eigenvectors, we can say that in the TISE $\psi(x)$ is an **eigenfunction** and E is the **eigenvalue** (for the energy in this case).

In the previous section we saw two more examples of eigenvalue equations: Equation 3.30 and Equation 3.31.

Below is a generalised definition of the eigenfunction, analogous to the one for eigenvectors given above. This is *very* important: the eigenfunctions satisfying eigenvalue equations are the building blocks of quantum states.

Eigenfunction

An **eigenfunction** ψ of an operator \hat{A} is a wavefunction that respects the *eigenvalue equation*

$$\hat{A}\psi = \lambda\psi, \quad (3.35)$$

for some scalar *eigenvalue* λ .

This means that the action of ψ does not change the wavefunction itself, it only rescales it by some scalar value λ .

Note that we use the term eigenfunction when we talk about wavefunctions in the *wave formalism* of QM, but we could also talk about **eigenstates**, which is usually referred to the same concept but in the matrix formalism of QM.

3.6 General solution of the TDSE - expansion in energy eigenfunctions

So now we are getting to the point... the TISE is an eigenvalue equation, and when we say that we want to *solve* it, for a given Hamiltonian, what we mean is that we want to find the set of eigenfunctions and eigenvalues for that Hamiltonian! But why is that important?

Now that we added some extra tools to our quantum toolbox, we can *expand on that* (pun intended, spoilers in the section title) and gain a better understanding of what we can extract from the Schrödinger equation and from its eigenfunctions and eigenvalues.

First, let's get back to some properties of the stationary states. Given the TISE (Equation 3.24), the expectation value of the total energy is

$$\langle H \rangle = \int \psi^* H \psi dx = E \int |\psi|^2 dx = E,$$

which results from the normalization of the wavefunction $\psi(x)$. Nothing really new so far. Let's now calculate the variance:

$$\Delta H = \langle H^2 \rangle - \langle H \rangle^2 = E^2 - E^2 = 0.$$

What does it mean that the variance is 0? It means that every time we measure the energy on the eigenstate ψ , the energy we find is going to be *always* exactly E , because the variance of the measurement is 0! Now it is clear why we say that stationary state are states with a *well defined total energy*.

So it seems that stationary states are pretty boring after all: they have a well defined energy, their probability is not evolving with time, they only have a time-evolving global phase, and what do we even do with them? How can we describe any system that evolves with time and does not have a well determined energy? Remember how we opened this chapter, emphasising the importance of the superposition principle... this is going to give us something useful here.

Stationary states, or eigenstates, are not that useless after all, and this is because *the general time-dependent solution for any system can be obtained from a (weighted) sum of stationary states* - such a system will have *indeterminate energy*.

So, the TISE is useful because it gives a set of eigenvalues, E_1, E_2, E_3, \dots and eigenstates, $\psi_1(x), \psi_2(x), \psi_3(x), \dots$, from which we can obtain a (total) wavefunction for each allowed value of the energy, i.e.

$$\Psi_1(x, t) = \psi_1(x)e^{-iE_1 t/\hbar}, \quad \Psi_2(x, t) = \psi_2(x)e^{-iE_2 t/\hbar}, \dots \quad (3.36)$$

! Important

The eigenstates $\psi_1(x), \psi_2(x), \dots$ of Hermitian operators are *orthogonal*, meaning that the projection, or *overlap*, of one on the other is zero, or in mathematical terms:

$$\int dx \psi_i(x) \psi_j(x) = \delta(i - j), \quad (3.37)$$

so the integral is 0 if $i \neq j$ and 1 if $i = j$ (due to the normalization condition).

Now we can use the superposition principle to solve the TDSE, since if each of the wavefunction in Equation 3.36 is a solution of the TDSE, then a *linear combination* (i.e. weighted sum) of them is also a solution to the TDSE! We therefore obtain the **general solution** of the TDSE.

General solution of the time-dependent Schrödinger equation

Given a set of eigenstates ψ_i and eigenvalues E_i of the TISE, the general solution of the TDSE is given by the linear combination

$$\Psi(x, t) = \sum_{n=1}^{\infty} c_n \psi_n(x) e^{-iE_n t/\hbar} = \sum_{n=1}^{\infty} c_n \Psi_n(x, t), \quad (3.38)$$

where c_n are the **amplitudes** - have the function of “weights” for the set of eigenfunctions - and need to be found to determine the correct general solution.

The general solution has *indeterminate energy* because the time evolution is not just a global phase now, there is *phase-interference* arising from the constituent stationary states.

Equation 3.38 means that the eigenfunctions ψ_n constitute a **complete basis**, so every wavefunction at the time $t = 0$ can also be represented as a weighted sum of the eigenstates:

$$\Psi(x, 0) = \sum_{n=1}^{\infty} c_n \Psi_n(x, 0). \quad (3.39)$$

Since the time-dependance in the general solution of the TDSE Equation 3.38 is not merely a global phase, the probability of the general state, $|\Psi(x, t)|^2$ is now going to be also time-dependent, in contrast with the probability of stationary states:

$$\begin{aligned} P(x, t) &= |\Psi(x, t)|^2 = \sum_i c_i^* e^{iE_i t/\hbar} \psi_i^*(x, 0) \sum_j c_j^* e^{-iE_j t/\hbar} \psi_j^*(x, 0) \\ &= \sum_{i,j} c_i^* c_j e^{-i(E_j - E_i)t/\hbar} \psi_i^*(x) \psi_j(x). \end{aligned}, \quad (3.40)$$

where you can now see that the *local* phases $e^{-i(E_j - E_i)t/\hbar}$ create interference.

The probability of Equation 3.40 still has to be normalised (its integral must be one), i.e. $\int P(x, t) dx = 1$, and it follows that

$$\int P(x, t) dx = \sum_{i,j} c_i^* c_j e^{-i(E_j - E_i)t/\hbar} \int dx \psi_i^*(x) \psi_j(x) = \sum_{i,j} c_i^* c_j e^{-i(E_j - E_i)t/\hbar} \delta(i - j) = 1,$$

where we used the property of orthogonality of the eigenstates ψ_i (cfr. Equation 3.37). Therefore, imposing $i = j$, the normalization condition reduces to the following normalization condition for the amplitudes:

$$\sum_i |c_i|^2 = 1. \quad (3.41)$$

You can see get a better understanding of how the principle of superposition works by using the resource at [this link](#), also embedded below. Given a general solution, we can express this as a sum of eigenstates, weighted by appropriate coefficients (tick the “show expansion in energy eigenfunctions box to see it”). There is also an “expansion game” where you can try to find the correct coefficients, given some wavefunctions!

3.7 Commutators

Going back to operators, something important to know is that **the order of the operators matter**, and this property is reflected mathematically in the **commutators**.

In fact, we have already seen the importance of *commutation* when we stressed that in the expectation values, the ordering of the functions matters!

The commutator between two operators is defined as

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}. \quad (3.42)$$

The order matters for both operators *and* wavefunctions.

If the commutator between two operators is non-zero, swapping the order of the two operators will result in different final wavefunctions:

$$[\hat{A}, \hat{B}] \neq 0 \implies (\hat{A}\hat{B})\psi \neq (\hat{B}\hat{A})\psi.$$

If the commutator is zero, it does not matter which order we use to apply \hat{A} and \hat{B} , i.e.

$$[\hat{A}, \hat{B}] = 0 \implies AB = BA. \quad (3.43)$$

3.7.1 Properties of commutators

These properties are *very* important, please remember them, as they will be essential during the whole course, but they will also be your bread and butter in any other course tackling quantum mechanics applications next year.

! Important - Useful commutation relations

Transpose: $[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}]$

Expansion: $[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B}$

Jacobi identity: $[\hat{A}, [\hat{B}, \hat{C}]] + [\hat{B}, [\hat{C}, \hat{A}]] + [\hat{C}, [\hat{A}, \hat{B}]] = 0.$

Hermitian conjugate: $[\hat{A}, \hat{B}]^\dagger = [\hat{B}^\dagger, \hat{A}^\dagger]$

The dagger symbol indicates the *transpose conjugate* - it makes more sense to use this name if we use the matrix formalism of QM instead of the wave formalism, but for the wave formalism consider the complex conjugate of the operators, so treat the \dagger symbol as a $*$.

3.7.2 The canonical commutation relation

An example of non-commuting operators is given by the conjugate variables \hat{x} and \hat{p}_x , which is particularly important, so important that it has its own name: the **canonical commutation relation**.

Canonical commutation relation

In 1D:

$$[\hat{x}, \hat{p}_x] = i\hbar. \quad (3.44)$$

More generally, indicating with subscripts i, j, k the different directions for coordinates and momenta in 3D, we have:

$$[\hat{r}_i, \hat{p}_i] = i\hbar, \quad [\hat{r}_i, \hat{p}_j] = 0 \quad \text{for } i \neq j \quad (3.45)$$

$$[\hat{r}_i, \hat{r}_j] = 0, \quad [\hat{p}_i, \hat{p}_j] = 0 \quad \forall i, j \quad (3.46)$$

3.7.3 The generalised uncertainty principle

The uncertainty for position and momentum can be formalised into a very important result between other pairs of conjugate variables. We will skip the demonstration of how to get to this result here (it is in the slides if you are curious), but the **generalised uncertainty principle** for two *incompatible* operators \hat{A} and \hat{B} is:

$$\Delta A \Delta B \geq \frac{1}{2i} |\langle [\hat{A}, \hat{B}] \rangle| \quad (3.47)$$

3.7.4 Exercise

Demonstrate the canonical commutator of Equation 3.44.

💡 Hints

Consider the commutator applied to a generic wavefunction $\psi(x)$ in 1D, and use the form of the momentum operator.

Solution

Let's start with the commutator applied to a wavefunction $\psi(x)$ in 1D:

$$[\hat{p}_x, \hat{x}] \psi(x) = (\hat{p}_x \hat{x} - \hat{x} \hat{p}_x) \psi(x).$$

To lighten the notation and avoid misunderstandings, in the following I write $\psi(x)$ as ψ , so the x -dependance is implicit.

Using the momentum operator definition of Equation 3.11:

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

Using the transpose properties of the commutators, then

$$[\hat{p}_x, \hat{x}] = -i\hbar \implies [\hat{x}, \hat{p}_x] = i\hbar.$$

3.8 Time evolution of expectation values

We have seen the time-evolution of stationary state wavefunctions (Equation 3.27) and the directly related probability densities; now we wish to see how these map into the time evolutions of expectation values, for values which may or may not share a compatible eigenstate basis with the Hamiltonian. Taking the time-derivative of an expectation value and being careful (as ever) to apply to product rule correctly, the key result is shown below.

$$\frac{d\langle \hat{A} \rangle}{dt} = \frac{d}{dt} \int dx \Psi^* \hat{A} \Psi = \int \frac{\partial \Psi^*}{\partial t} \hat{A} \Psi + \int dx \Psi^* \frac{\partial \hat{A}}{\partial t} \Psi + \int dx \Psi^* \hat{A} \frac{\partial \Psi}{\partial t}$$

Using the TDSE (Equation 3.12), we can cast the above equation in the form

$$\begin{aligned} \frac{d\langle \hat{A} \rangle}{dt} &= -\frac{1}{i\hbar} \int dx (\hat{H}\Psi)^* \hat{A} \Psi + \int dx \Psi^* \frac{\partial \hat{A}}{\partial t} \Psi + \frac{1}{i\hbar} \int dx \Psi^* \hat{A} \hat{H} \Psi \\ &= \frac{1}{i\hbar} \int dx \Psi^* (AH - HA) \Psi + \int dx \Psi^* \frac{\partial \hat{A}}{\partial t} \Psi, \end{aligned}$$

and using the definition of the commutator (Equation 3.42), we finally obtain:

$$\frac{d\langle \hat{A} \rangle}{dt} = \frac{1}{i\hbar} \langle [\hat{A}, \hat{H}] \rangle + \left\langle \frac{\partial A}{\partial t} \right\rangle. \quad (3.48)$$

This means that the rate of change of an expectation value is proportional to the commutator of its operator and the Hamiltonian operator \hat{H} , plus the time-dependence of the measurement operator itself. The second term in Equation 3.48 is usually zero: the most obvious way to introduce it would be for the system environment to time-evolve, e.g. for $U(r, t)$ to be time-dependent, which we don't consider in this course.

For “normal” time-independet operators, the driver of change in expectation values is whether the operator \hat{A} commutes with the Hamiltonian: this makes intuitive sense since if the operator \hat{A} and the Hamiltonian are compatible, then the stationary states will be eigenstates of \hat{A} and hence the expectation value will also be stationary, i.e.

$$[H, A] = 0 \implies H(A\Psi) = A(H\Psi) \implies H(A\Psi) = EA\Psi$$

and since the TISE needs to be satisfied, we need $A\Psi = \lambda\Psi$.

In short, what does this mean? *The commutation of an operator with the Hamiltonian implies time-invariance.*

3.8.1 The Ehrenfest theorem and the correspondence principle

At the beginning of the course we said that, in the classical limit, i.e. when we recover the probability distribution after many repetitions of the experiment, and take the average of the physical quantities observed, quantum mechanics should return the classical physics predictions.

We have introduced the expectation values and have seen that physically these correspond to the value that on average we would obtain after repeating a measurement many times, and the standard deviation corresponds to the width of the distribution of the measurement outcomes around the expectation value. We can then guess that if we evaluate the time evolution of the expectation value for some quantities (e.g. position or momentum operators), we should obtain the classical equations of motion for those quantities. Let’s see this.

If we apply Equation 3.48 to the operators $\hat{A} = \hat{x}$ and $\hat{A} = \hat{p}$, considering their commutators with the non-relativistic Hamiltonian (Equation 3.10)

$$\begin{aligned} [\hat{x}, \hat{H}] &= [\hat{x}, \hat{p}^2/2m + U(x)] = [\hat{x}, \hat{p}^2/2m], \\ [\hat{p}, \hat{H}] &= [\hat{p}, \hat{p}^2/2m + U(x)] = [\hat{p}, U(x)], \end{aligned} \tag{3.49}$$

using the commutation properties, we obtain:

$$\frac{d\langle x \rangle}{dt} = \frac{1}{2i\hbar m} \langle [\hat{x}, \hat{p}^2] \rangle = \frac{1}{2i\hbar m} \langle \hat{p}[\hat{x}, \hat{p}] + [\hat{x}, \hat{p}]\hat{p} \rangle = \frac{1}{2i\hbar m} \langle 2i\hbar\hat{p} \rangle,$$

which generalised to 3D is

$$\langle \mathbf{p} \rangle = m \frac{d\langle \mathbf{r} \rangle}{dt}. \tag{3.50}$$

Similarly, for the time evolution of the momentum expectation value, we obtain:

$$\frac{d\langle \mathbf{p} \rangle}{dt} = -\langle \nabla U \rangle, \quad (3.51)$$

which is Newton's second law!

This recovery of classical dynamics in the expectation values of quantum mechanics is **Ehrenfest's Theorem**. It is an important example of the general correspondence principle introduced by Bohr in his first nuclear-atom construction, which applied to large atomic number Z : that in the large-system statistical limit, quantum mechanics is consistent with classical physics.

QM is the fundamental physics, with classical behaviours being a handy approximation in systems where quantum features like wavefunction collapse and amplitude interference are “washed out” by the incoherent superposition of many different contributions to the total wavefunction.

Only a few special systems such as carefully prepared superfluids and Bose-Einstein condensates are able to propagate quantum effects to macroscopic scales — although the consequences of fundamentally quantum effects are immensely important in ensuring a few trifling details like the stability of matter, the distribution of structure in the universe, and solar fusion.

3.9 Summary

- In general the Hamiltonian is the generator of time-evolution for a (quantum) system.
- A state of definite energy is an eigenstate of the Hamiltonian, and its time-evolution is just a phase $\propto Et/\hbar$.
- For time-independent potentials, the time-independent Schrödinger equation can be obtained from the time-dependent one; obtain stationary states by assuming wavefunction separable into r and t parts.
- The TISE is an eigenvalue equation, and the corresponding eigenfunctions and eigenfunctions are used to obtain stationary states.
- The general solution for the TDSE can be obtained from a linear combination of the stationary states with a definite energy.
- Wavefunctions must be square-integrable, continuous everywhere, and smooth except at infinite potential steps.
- Expectation values also time-evolve with the Hamiltonian: static operators which commute with the Hamiltonian have time-invariant expectation values.
- Ehrenfest's theorem results from this: expectation values of position and momentum obey Newton's Laws, cf. correspondence principle.

4 Solutions of Schrödinger Equations in 1D - Bound states

Learning objectives

1. Understanding the procedure to find the general solution of the TDSE
 2. Understanding the difference between bound and scattering states
 3. Calculating the eigenfunctions of the TISE for different potentials in 1D, i.e. infinite well, square well, delta well, harmonic well
 4. Executing calculations of expectations values of operators on eigenstates and their combinations, for different potentials
 5. Familiarising yourself with the Dirac (bra-ket) notation
-

4.1 Solving a general problem with the TDSE

In this Chapter we will focus on solving the Schrödinger equation for different kinds of potential, in one dimension. Let's recap some of the things we've seen in the previous chapter.

The general problem that we will encounter in this course is the following: *given a time-independent potential $U(x)$ and the initial wavefunction $\Psi(x, 0)$, find the wavefunction $\Psi(x, t)$ at any following time.*

The strategy for this kind of problem is the following:

1. Solve the *time-independent Schrödinger equation* (Equation 3.23) to find the eigenstates $\psi_n(x)$ and eigenvalues $E_n(x)$;
2. With the eigenstates, “reproduce” $\Psi(x, 0)$ using the superposition principle, i.e.

$$\Psi(x, 0) = \sum_{n=1}^{\infty} c_n \psi_n(x),$$

with the appropriate amplitudes c_n .

3. Build the stationary states from the eigenstates and eigenvalues found at the first step

$$\Psi_n(x, t) = \psi_n(x) e^{-iE_n t/\hbar},$$

and use the superposition principle to find the general time-dependent Schrödinger equation (Equation 3.38), i.e.

$$\Psi(x, t) = \sum_{n=1}^{\infty} c_n \psi_n(x) e^{-iE_n t/\hbar} = \sum_{n=1}^{\infty} c_n \Psi_n(x, t).$$

In the rest of this course we will focus on the first two points, i.e. how to solve *time-independent* Schrödinger equations, as unfortunately we do not have enough time to deal with time evolution (ironically the structure of the course is time-dependent... **ba-dum-ts**).

4.2 Solving the TISE

Before going ahead and solving the 1D TISE with different potentials, let's see what are the conditions we need to impose to have physical solutions and what their characteristics are, both for the eigenfunctions and energy eigenvalues.

4.2.1 Properties of the eigenfunctions

Let's rewrite the 1D TISE (Equation 3.23):

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + U(x)\psi(x) = E\psi(x) \implies \frac{d^2\psi}{dx^2} = -\frac{2m}{\hbar^2}[E - U(x)]\psi(x). \quad (4.1)$$

From the RHS of the last equation we can make some observation based on the sign of $E - U(x)$ and its effect on the curvature of the wavefunction (since the sign of the second derivative defines whether a function is concave or convex):

- $E - U(x) > 0 \implies E > U(x)$ - *classically allowed region*; the wavefunction solution of the SE bends toward the x axis; *wavelike solutions*
- $E - U(x) < 0 \implies E < U(x)$ - *classically forbidden region*; the solution of the SE curves away from the x axis; *exponential solutions*
- $E = U(x)$ - *turning points*.

When solving the SE for any given potential, remember the following conditions for the wavefunctions, that we've been stressing in the previous chapters:

- **single-valued** (i.e. for each value of x (and t) there is only one corresponding value of Ψ)
- **continuous** (i.e. you could “draw” it without lifting your pen from the paper)
- **smooth except at infinite-sized potential steps** (i.e. no sudden changes in the derivative
- continuous first derivatives *if there are no singularities in the potential*)

On top of these, impose the normalization condition.

In the following of this chapter we will apply these fundamental rules to the SEs for different (easy - and useful) potentials.

4.2.2 Properties of the eigenenergies - Bound states and scattering states

The set of values of the total energy, E , for which the TISE has physically admissible solutions, determines the **energy spectrum**.

The Schrödinger equation has two kind of solutions: *bound states* and *scattering states*. Their definition in quantum physics is a bit more “relaxed” from the one in classical physics, due to the effect of *tunneling* through any finite potential barrier, which we will see later.

In classical physics, bound states and scattering states are defined comparing the total energy to the potential energy in an interval: if there are turning points ($E = U(x)$ for some x - the kinetic energy is 0), then the system is in a bound state in the interval between turning points; otherwise it is in a scattering state (see Figure 4.1).

The definition of bound states in quantum physics is a bit more “relaxed” from the one in classical physics, due to the effect of *tunneling* through any finite potential barrier, which we will discuss later. Overall we have a similar condition as for classical bound states but, instead of using the turning points, we compare the total energy with the potential $U(+\infty)$ and $U(-\infty)$.

Bound states

Bound states satisfy the condition: $E < U_{\pm\infty}$

- Classically they cannot “escape” to infinity, and even with quantum relaxation of that rule (as we shall see), the probability density will be localised at least close to the classically allowed $E > U$ part of the potential.
- The energies of bound states are *quantised* (or *discrete*).
- The lowest discrete energy level is called the *ground-state energy* of the system, and all the higher discrete energy levels are called *excited states*.
- The eigenfunctions of bound states vanish for $\mathbf{r} \rightarrow \infty$.

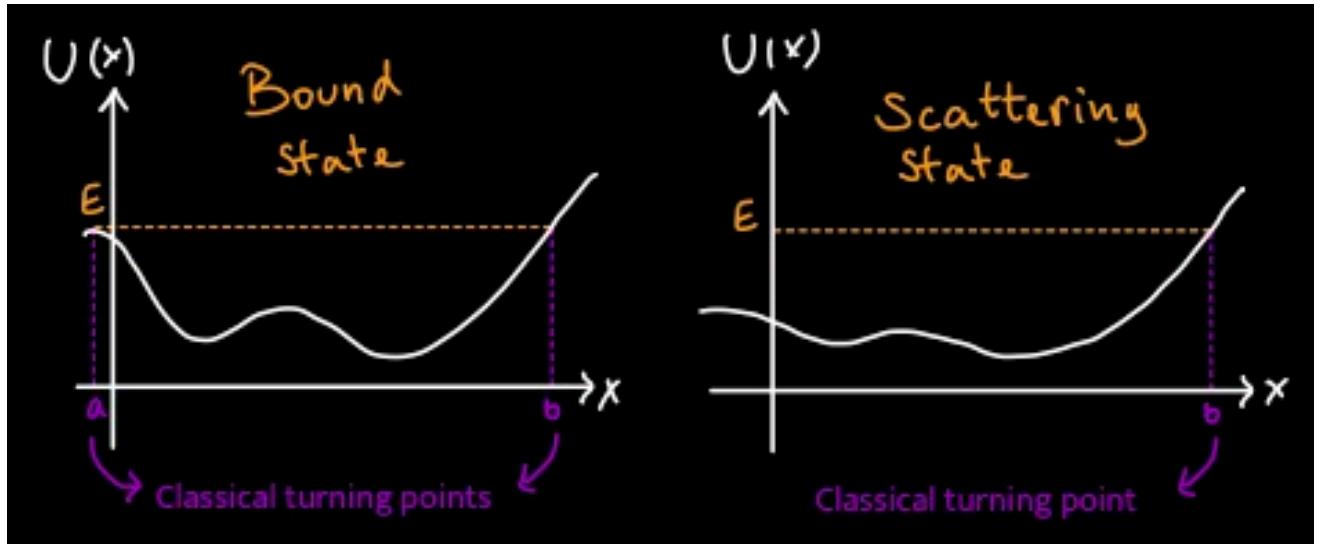


Figure 4.1: Bound and classical states in classical physics.

- The number of bound state is finite (including 0) or infinite, depending on the form of the potential U .

Furthermore, bound states are real valued function (with global phases), so their probability flux $\mathbf{J} \propto \psi^* \nabla \psi' - \psi \nabla \psi^*$ must be zero.

Scattering (or unbound) states

Scattering (or unbound) states are obtained if: $E > U_{\pm\infty}$.

- The energies of unbound states are such that the particle would be also classically unbound
- The energies of scattering states are *continuous* and can extend to $E = +\infty$
- Since the energies form a continuum, there is no energy quantisation for scattering states
- The eigenfunctions of unbound states are finite for $\mathbf{r} \rightarrow \infty$.

Figure 4.2 represents an example of potential with both bound and unbound states.

Now that we know what to expect to find for the energy eigenvalues and eigenfunctions in the solution of the TISE, let's go ahead and solve this for some potentials in one dimension.

In the following of this chapter we will be looking at *bound state solutions* for one-dimensional *potential wells*.

Well, well, well...

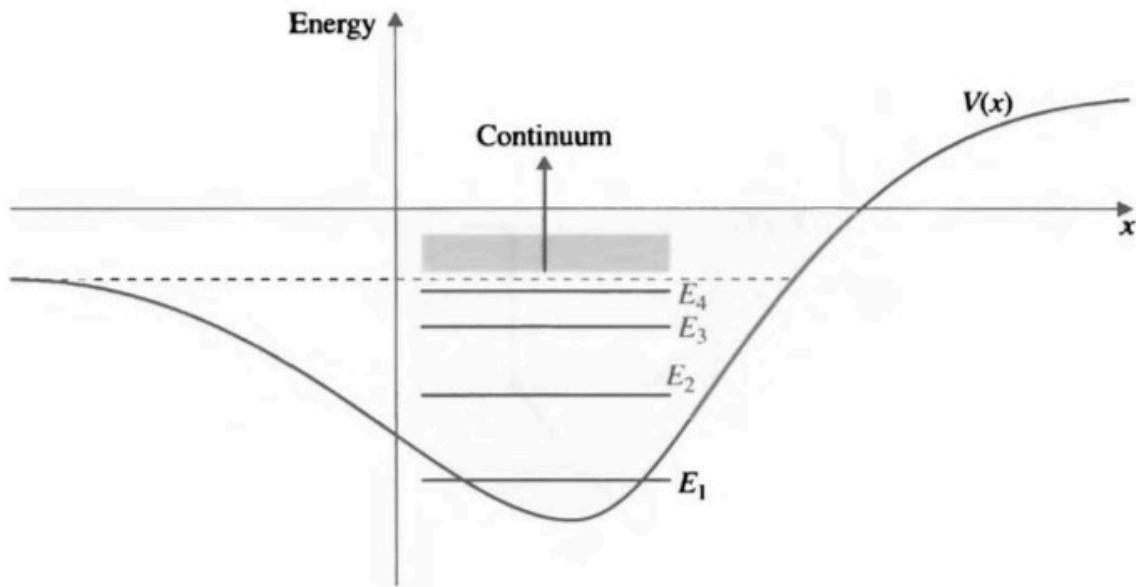


Figure 4.2: Example of an energy spectrum containing both bound states (discrete energy levels) and scattering states (continuum energies).



Figure 4.3: ...look what we have here.

4.3 The infinite well potential (particle in a box)

The potential for the infinite potential well has the following form:

$$V = \begin{cases} 0 & \text{for } -a \leq x \leq a, \\ \infty & \text{elsewhere,} \end{cases} \quad (4.2)$$

represented in Figure 4.4.

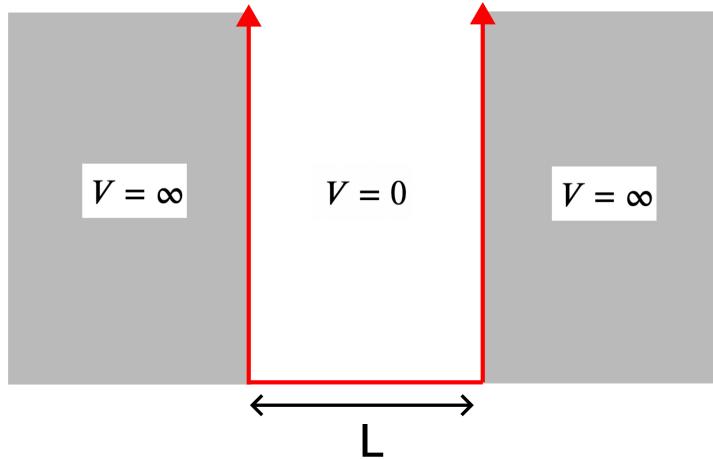


Figure 4.4: Infinite well potential. In this picture, $L = 2a$, if we refer to the definition in Equation 4.2, and $U(x) = V(x)$.

Outside the potential well the probability of finding the particle is zero, hence $\psi(x) = 0$ for $x < -a, x > a$.

This means that inside the well, the particle behaves like a free particle, since in Equation 3.23 the potential $U(x)$ is zero in the range $[-a, a]$:

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

We can rewrite this in the form

$$\frac{d^2\psi(x)}{dx^2} = -k^2\psi(x), \quad (4.3)$$

where we defined $k = \sqrt{2mE}/\hbar$, implicitly imposing $E > 0$.

Since the potential is flat within the well, we can show that plane waves (of any k) are good solutions to the Schrödinger's equation.

We could write these as a continuous spectrum of waves of given positive k ,

$$\psi_k(x) = a_x e^{ikx} + b_k e^{-ikx},$$

but we can just expand this and reparametrise it in terms of $\cos(kx)$ for $a_k = b_k$ and $\sin(kx)$ for $a_k = -b_k$, for convenience, given the boundaries of the potential.

So the easiest solution to the differential equation of Equation 4.3 is

$$\psi_k(x) = A_k \cos(kx) + B_k \sin(kx). \quad (4.4)$$

What next? We need to consider the *boundary conditions* of the function.

We know that the wavefunction must be *continuous* and we know that outside the well it is zero, so we need to impose that at the boundaries $x = -a$ and $x = +a$ the wavefunction must also be zero:

$$\psi(\pm a) = 0. \quad (4.5)$$

From Equation 4.4, using the symmetry of the functions

$$\begin{cases} \psi_k(-a) = A_k \cos(-ka) + B_k \sin(-ka) = 0 \implies A_k \cos(ka) - B_k \sin(ka) \\ \psi_k(a) = A_k \cos(ka) + B_k \sin(ka) = 0, \end{cases} \quad (4.6)$$

solving this system we obtain

$$\begin{aligned} A_k \cos(ak) &= 0 \implies k = \frac{n\pi}{2a}, \quad \text{for } n = 1, 3, 5\dots \\ B_k \sin(ak) &= 0 \implies k = \frac{n\pi}{2a}, \quad \text{for } n = 2, 4, 6\dots \end{aligned}$$

So replacing these values of k in the wavefunctions of Equation 4.6, we obtain even/odd pairs of eigenfunctions $\psi_n(x)$, which we define using the notation below:

$$\psi_n(x) = \begin{cases} u(x)_n^+ = A \cos \frac{n\pi x}{2a} & \text{for } n = 1, 3, 5\dots \\ u(x)_n^- = B \sin \frac{n\pi x}{2a} & \text{for } n = 2, 4, 6\dots \end{cases} \quad (4.7)$$

Note that now we have a discrete set of even and odd states! But wait, for the free particle we have a continuum of energy levels (unbound state), so what is different here, if we are assuming that within the well the particle can be considered a free particle, after all? The difference is... the boundaries! The loss of the continuous k spectrum and appearance of a discrete spectrum is due to the fact that we are in fact imposing that at the boundaries, and outside those, the wavefunction is zero!

This has another consequence: remember when we said that the free particle represented as a plane wave had a big normalization problem because fully delocalised? Well, it looks like we

do not have a fully delocalised particle after all, because we put infinite walls so we cheated the system, and now we can actually impose the normalization condition to find the constants A and B in Equation 4.7, by imposing $\int_{-\infty}^{\infty} \psi_n(x)dx = \int_{-a}^a \psi_n(x)dx = 1$.

For odd n :

$$\int_{-a}^a dx u(x)_n^+ = \int_{-a}^a A^2 \cos(n\pi x/2a) dx = A^2 a [1 + \sin(n\pi)/(n\pi)] = 1 \implies A = 1/\sqrt{a}.$$

For even n :

$$\int_{-a}^a dx u(x)_n^- = \int_{-a}^a B^2 \sin(n\pi x/2a) dx = A^2 a [1 - \sin(n\pi)/(n\pi)] = 1 \implies B = 1/\sqrt{a}.$$

Therefore the normalised eigenstates of the infinite well potential are as follows.

Eigenstates for the infinite square potential well

$$\psi_n(x) = \begin{cases} u(x)_n^+ = \frac{1}{\sqrt{a}} \cos \frac{n\pi x}{2a} & \text{for } n = 1, 3, 5, \dots \\ u(x)_n^- = B \sin \frac{n\pi x}{2a} & \text{for } n = 0, 2, 4, \dots \end{cases} \quad (4.8)$$

The eigenvalues of the energy, below, are derived from $E_k = \frac{\hbar^2 k^2}{2m}$, with $k = \frac{n\pi}{2a}$.

Eigenenergies for the infinite square potential well

$$E_n^\pm = \frac{\hbar^2 \pi^2}{8a^2 m} n^2 = \frac{\hbar^2 \pi^2}{2L^2 m} n^2, \quad \text{for } n = 1, 2, 3, \dots \quad (4.9)$$

where in the last passage we used $L = 2a$ for the length of the well.

Notice that $n = 0$, and therefore $E = 0$ is not allowed, as the particle can never be at rest. The lowest allowed energy is the *ground state energy*, or *zero-point energy*,

$$E_1 = \frac{\hbar^2 \pi^2}{8a^2 m},$$

and for $n > 1$ we get the *excited states*.

The existence of the zero-point energy is due to the fact that in this case the particle can't be *fully delocalised* as for the free particle case, because the wavefunction can't extend over the boundaries, so the momentum must have some uncertainty and can't be exactly zero.

But hang on, where did this " $\psi = 0$ if $U = \infty$ " rule come from? We will get a proper answer when we come to the delta-function well, but in short it is because of the infinite "energy penalty" on the particle should it penetrate the wall: the wavefunction will be forced to look like a decaying exponential with decay length $1/V = 0$. In the following section we will set

up the potentials differently, so the energy penalty (actually $V - E$) and hence the amplitude decay length are finite.

Notice also that the eigenenergies are proportional to n^2 , so they are not equally spaced: their distance increases with n .

A representation of the eigenstates and eigenenergies is shown in Figure 4.5.

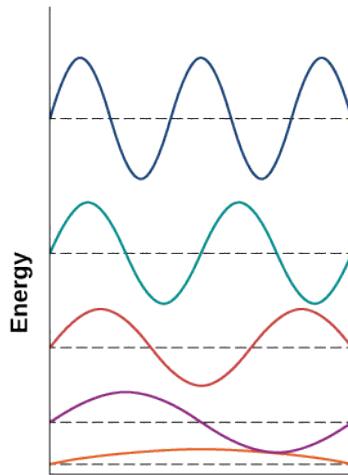


Figure 4.5: Eigenfunctions (solid lines) and eigenenergies (dotted lines) for the infinite well potential. Warning: these diagrams visually mix energy levels and wavefunctions. Each eigenfunction is drawn with its zero level at the height of the corresponding energy.

In the form we're using, with $x = 0$ at the centre of the well, the solutions alternate between $\cos(n\pi x/2a)$ and $\sin(n\pi x/2a)$ functions, with $\cos(\pi x/2a)$ as the even-symmetric ground state. That ground state wavefunctions have no zero-crossings (the $x = \pm a$ endpoints do not count) is a general feature of energy eigenstates, with each excited state adding one crossing.

We could have defined the wavefunction in the interval $[0, L]$ instead of $[-a, a]$. Note that in that case we would start with the same solution Equation 4.4, but the eigenstates would be different and have only “sine” terms, due to the different boundary conditions determining different values of A and B ! The reason why we defined these boundaries here is to use the symmetry property around the center of the well (we'll discuss this a bit further in the section on “parity”.)

4.3.1 Properties of infinite-square-well potential eigenfunctions

1. Normalization (we demonstrated this to derive A, B): $\int_{-\infty}^{\infty} dx u(x)_n^{\pm *} u(x)_n^{\pm} = 1$
2. Orthogonality: $\int_{-\infty}^{\infty} dx u(x)_n^{\pm *} u(x)_m^{\mp} = 0$; $\int_{-\infty}^{\infty} dx u(x)_n^{\pm *} u(x)_m^{\pm} = \delta_{mn}$

3. Eigenstates of \hat{H}, \hat{p}^2 , but not e.g. \hat{p}, \hat{x}
4. Expectation values: $\langle x \rangle = 0, \langle p \rangle = 0, \langle p^2 \rangle = 2mE_n^\pm$.

 Caution

Note that because of the orthogonality property there is no overlap between eigenstates, but there can be overlap between linear combinations of them. Changing $|a|$ will generally mix the eigenstates!

4.3.1.1 Exercise

Demonstrate the properties at points 3. and 4. above.

 Hints

Hint 1

Useful relations:

$$\begin{aligned}\sin^2(A) + \cos^2(A) &= 1 \\ \sin^2(A) &= \frac{1}{2}[1 - \cos(2A)] \\ \cos^2(A) &= \frac{1}{2}[1 + \cos(2A)] \\ \cos(A \pm B) &= \cos(A)\cos(B) \mp \sin(A)\sin(B) \\ \sin(A \pm B) &= \sin(A)\cos(B) \pm \cos(A)\sin(B)\end{aligned}$$

Hint 2

Use $\cos(A+B) \pm \cos(A-B)$, and $\sin(A+B) \pm \sin(A-B)$.

Solution

– No solution added (yet) –

A good exercise is to show that the stationary states do not have definite momentum or position, by application of the \hat{p} and \hat{x} operators: you can easily show that their expectation values are zero, consistent with the symmetry of the well. It may seem strange to not be in a definite-momentum state, especially since the Hamiltonian within the well contains only the kinetic energy term, $\hat{p}^2/2m$, which commutes with \hat{p} : this is because the sinusoidal solutions always contain equal amounts of positive and negative x , i.e. right-going and left-going momenta at the same time. They do, however, have non-zero $\langle x^2 \rangle$ and $\langle p^2 \rangle$ expectations.

As stationary states are eigenstates (of the Hamiltonian), they form an orthonormal basis (check this!) from which arbitrary states can be constructed by superposing different stationary states with different mixtures of (complex) coefficients (check again the interactive material

of Section 3.6) — subject to the overall probability-normalisation $\int |\psi|^2 dx = 1$ requirement, of course.

The overlap coefficients can be extracted via the $c_n = \int \psi^* \phi_n dx$ overlap integrals as usual, and the probabilities of a measurement collapsing into any particular eigenstate given by the mod-square of its coefficient, $P(n) = |c_n|^2$.

It is interesting to note that this (as do all such syntheses from stationary basis states) leads to wavefunctions of *indefinite energy*: this feels strange, and perhaps at odds with energy conservation, although we would find that that rule resumes nicely if we compute expected energy. This is a hint that a complete picture of relativistic energy and momentum conservation in QM is a bit more complex than we are ready to deal with!

4.3.2 Parity

Parity, i.e. the reflection symmetry $x \rightarrow -x$ is an useful symmetry, and by choosing to place the zero coordinate along x in the middle of the potential well, we obtained wavefunctions that respect the same symmetry, being even-symmetric under the parity transformation $x \rightarrow -x$.

The stationary states are also eigenstates of the parity operator \hat{P} , i.e. $\hat{P}\psi(x) = \psi(-x) = \pm 1$ (it has only two possible eigenvalues).

The spatial eigenstates $u_n^\pm(x)$ have definite **parity**, and we can see that by applying the parity operator \hat{P} on them

$$\hat{P}u_n^+(x) = u_n^+(x); \quad \hat{P}u_n^-(x) = -u_n^-(x) \quad (4.10)$$

The eigenstate alternate between parity $+1$ for the cos solutions, and parity -1 for the sin ones. Potentials with odd or indefinite symmetry, however, don't have to have solutions of definite parity at all (and odd potentials can't have bound states!)

Caution

The goodness of the parity quantum number is directly related to the definite parity of the potential, $U(x) = U(-x)$. **An asymmetric potential, $U(x) \neq U(-x)$, will not have eigenstates of definite parity.**

This is why earlier we justified this choice for defining the potential in the interval $[-a, a]$ instead of $[0, L]$.

You need to be careful about coordinates: a well in $[0 \dots L]$ has only “sine” solutions, and $-x$ is no longer useful, but the real symmetry is still there. Parity is defined about a coordinate origin: you need to choose an appropriate one.

4.4 The delta-function potential

What happens if we make the infinite well potential infinitely narrow, i.e. for $L \rightarrow 0$?

We obtain a delta well potential!

The infinite square well showed us a system where the wavefunction is non-trivial only in the classically allowed region with $E > U$, and we saw that it was a wave-like sinusoid in that region.

Let's change the “reference” zero potential (without loss of generality) so that $U = 0$ everywhere except at $x = 0$ where we place a *negative* delta-function well $U(x) = C - \alpha\delta(x)$. Without loss of generality, we can set $C = 0$, so

$$U(x) = \begin{cases} -\alpha\delta(x) & \text{at } x = 0 \\ 0 & \text{otherwise,} \end{cases}$$

where α is a positive constant determining its “strength”.

The delta-function potential is a neat way to show the behaviour of the complementary classically-disallowed region where $E < U$. We can then put both solution types together to build composite wavefunctions for potentials and with both allowed and finitely-disallowed regions in x .

The particle energy for a bound state has to be $E < 0$, so we expect to see it localised somewhere around the needle-like potential at $x = 0$. But there is no space for a non-trivial wavefunction within the well, and the particle kinetic energy is only the finite amount E short of being enough to escape the well, so this time we find a non-trivial wavefunction *outside* the well only.

Let's solve the TISE. For $x \neq 0$, $U(x) = 0$, so the TISE is

$$\frac{d^2\psi}{dx^2} = -\frac{2mE}{\hbar^2}\psi \equiv \kappa^2\psi \quad (4.11)$$

for $\kappa = \sqrt{-2mE}/\hbar$.

Why did we define this constant?

The same argument as for using plane wave solutions within the infinite well applies to the flat $U = 0$ potential, but this time with a relative minus sign on the “spare energy” $E - U$. We are looking at bound states first, so $E < 0$ and κ is real and positive. Were we to approach this naively to find the solution, we could assume the usual e^{ikx} form, compute that $k = \sqrt{2m|E|}/\hbar = i\sqrt{2m|E|}/\hbar$, and let the i factors cancel to turn the complex exponential into a real exponential $e^{\pm|k|x}$; being a little smarter we can anticipate this already and instead define the exponential constant as $\kappa = \sqrt{2m|E|}/\hbar$.

For real κ , i.e. $E < 0$, the general solution of Equation 4.11 is

$$\psi(x) = A_- e^{-\kappa x} + B_- e^{\kappa x},$$

but for $x \leq 0$ the first term blows up, so $A_- = 0$ for $x \leq 0$ and

$$\psi(x) = B_- e^{\kappa x}. \quad (4.12)$$

Vice versa, for $x \geq 0$ the second term of the general solution must be 0, so in that case we will have

$$\psi(x) = A_+ e^{-\kappa x} + B_+ e^{\kappa x},$$

with $B_+ = 0$, so for $x \geq 0$

$$\psi(x) = A_+ e^{-\kappa x}. \quad (4.13)$$

Now we need to impose the *standard boundary conditions*:

1. ψ must be *continuous*
2. $d\psi/dx$ must be continuous except at points at infinite potential

Continuity:

Equating Equation 4.12 and Equation 4.13 at $x = 0$: $\psi(0) = B_- = A_+ \equiv A$, so

$$\psi(\pm|x|) = Ae^{-\kappa|x|}.$$

The delta-function potential and the obtained solutions are represented in Figure 4.6.

For point 2 of the boundary conditions, the derivative of the wavefunction would be discontinuous at $x = 0$ due to the delta potential (infinite potential). To impose the continuity of the derivative everywhere else, integrate SE around 0, i.e. in an infinitesimal interval $[-\epsilon, \epsilon]$ and then take the limit $\epsilon \rightarrow 0$, obtaining a single allowed value of $\kappa = \frac{m\alpha}{\hbar^2}$.

We can find A by imposing the normalization condition with the mod-square as usual, and find

$$A = \sqrt{\kappa} = \frac{\sqrt{m\alpha}}{\hbar},$$

so the delta-function well, regardless of the strength α , has exactly *one (discrete) bound state* given by the following wavefunction and eigenvalue:

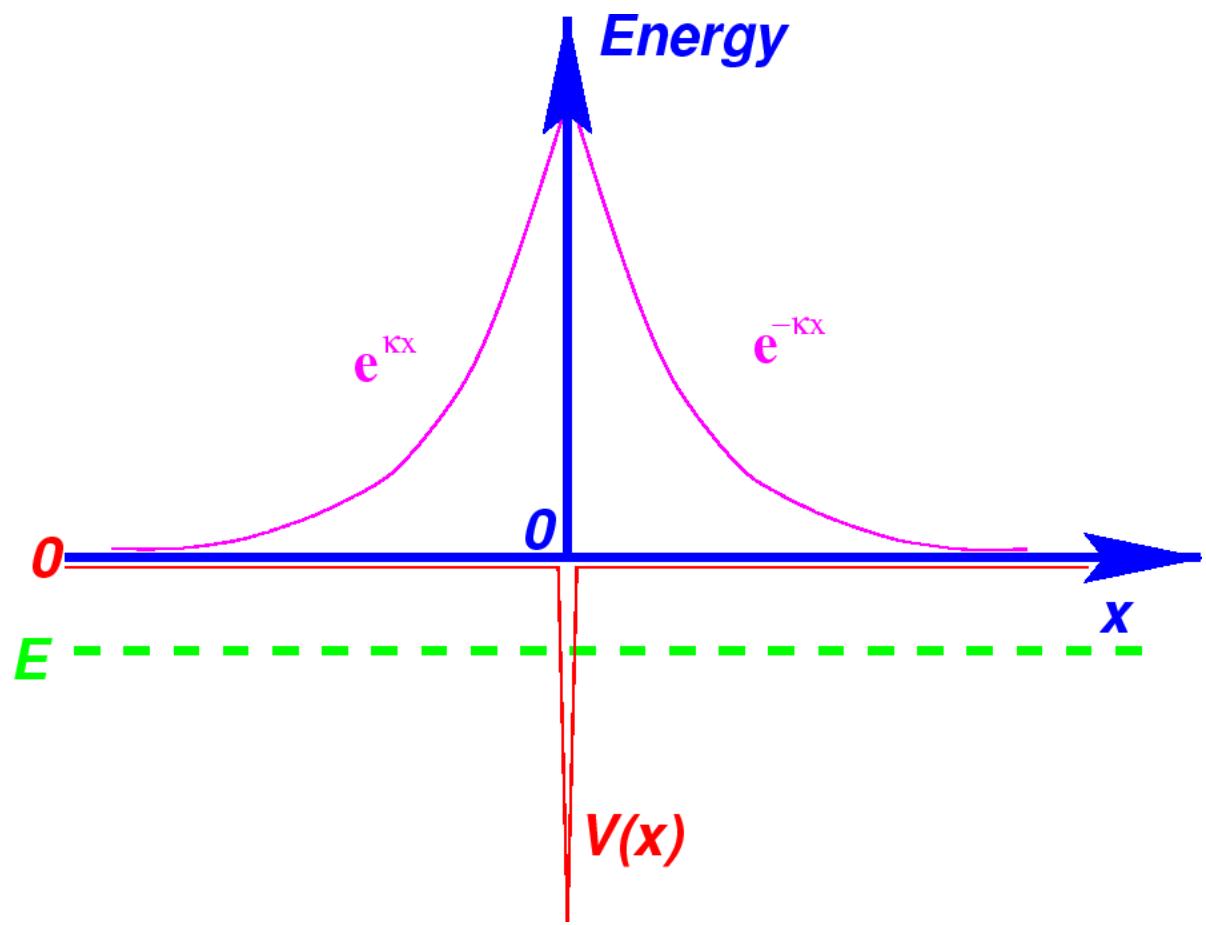


Figure 4.6: Delta-function potential and allowed solutions.

$$\psi(x) = \frac{\sqrt{m\alpha}}{\hbar} e^{-m\alpha|x|/\hbar^2}; \quad E = -\frac{m\alpha^2}{2\hbar^2}. \quad (4.14)$$

Demonstrate this as an exercise!

i Note

- The wavefunction has *tails* even outside the “well”, where we would not expect to find the particle. This is the *tunneling effect*.
- Note that there is *no quantization* of the allowed values for κ , and therefore for E : any given E has a valid stationary solution (valid for any α).

The typical length $\sim 1/\kappa$ to which the decaying exponentials “escape” the delta-function increases the smaller the energy deficit $|V - E|$ for escaping the well: $1/\kappa = \hbar/\sqrt{2m|E|}$. A finitely deep potential would break this continuous spectrum of energy eigenvalues by re-enabling the differentiability boundary condition... but of course there is no way to be both finitely deep and infinitely thin! Our next stop will, however, have aspects of both extreme solutions.

4.5 The finite square well

Let’s now imagine to start with the infinite well potential, but now lower the potential walls to a finite potential U_0 which will represent the depth of the well.

The finite square well is a well of both finite width and depth, from which both the infinite-square and delta-function wells are asymptotic limits. As its potential’s geometry is still composed of piecewise-assembled flat elements, it is the simplest case where we get to synthesise the distinct oscillatory and evanescent solution types (i.e. complex and decaying exponentials) together.

The potential is given by the function

$$U(x) = \begin{cases} -U_0 & \text{for } -a \leq x \leq a, \\ 0 & \text{for } |x| > a, \end{cases}$$

where U_0 is a positive constant.

The bound states will be found for $E < 0$, inside the well, where the general solution is oscillatory. Outside the well, for $E > 0$, the solution will be exponentially decaying and we have scattering states.

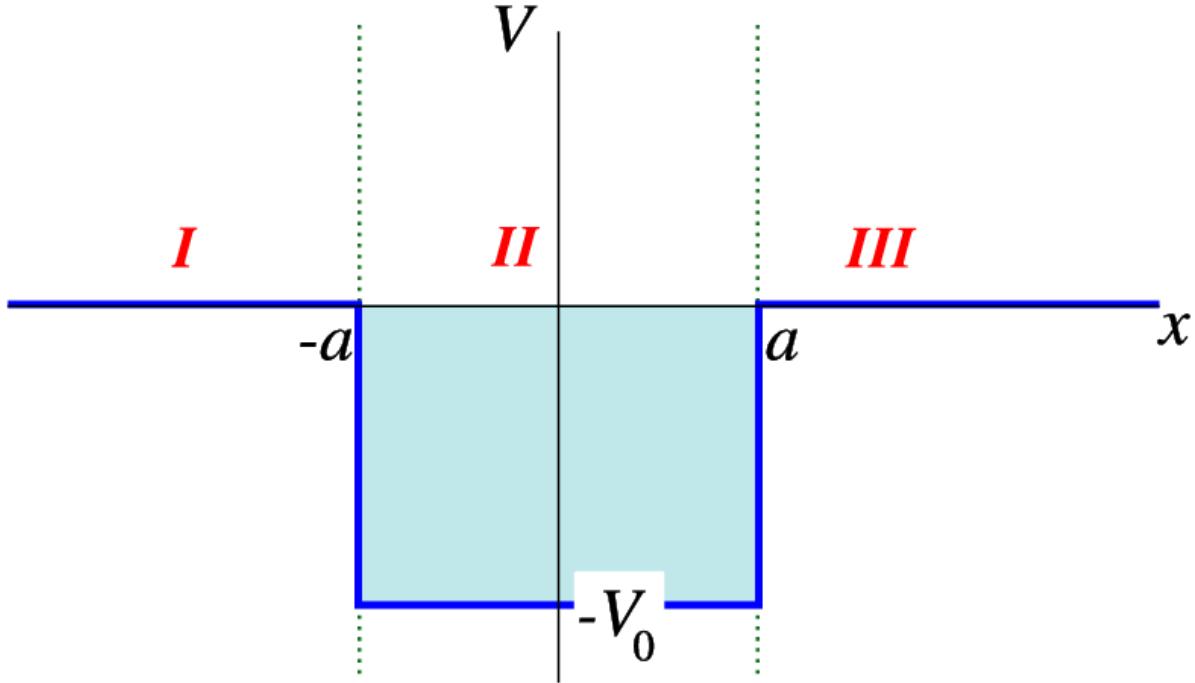


Figure 4.7: Square well potential. Here the potential is indicated as $V(x)$ instead of $U(x)$.

4.5.1 Bound states

Let's look at the bound states ($E < 0$).

Outside the well, in the region $x < -a$, the potential is 0, therefore the TISE can be rearranged in the form

$$\frac{d^2\psi}{dx^2} = \kappa^2\psi, \quad (4.15)$$

where $\kappa \equiv \sqrt{-2mE}/\hbar = \sqrt{2m|E|}/\hbar$ (real and positive).

Similarly to the case of the delta function, the general solution for this is $\psi(x) = A_- e^{\kappa x} + B_- e^{-\kappa x}$, but the second term would blow up for $x \rightarrow -\infty$, so just like for the delta-function:

$$\psi(x) = A_- e^{\kappa x} \quad \text{for } x < -a. \quad (4.16)$$

Due to the symmetry of the potential well, defined in the interval $[-a, a]$, we can find the solutions knowing that they have definite parity. It would in principle be sufficient to find the solutions for either positive or negative values of x and then use the parity condition.

For $x > a$, using the parity operator or following the same argument also seen for the previous case of the delta potential, the solution is

$$\psi(x) = A_+ e^{-\kappa x} \quad \text{for } x > a.$$

The coefficients B and F are determined by matching this wave smoothly by imposing the continuity conditions of the wavefunction and its derivative in space, inside the well.

The fact that $\psi(x)$ is nonzero at the walls increases the de Broglie wavelength in the well with respect to the case of a particle in the infinite well potential, and this lowers the energy and momentum of the particle (increasing $\Delta x \Rightarrow$ decreasing Δp).

Inside the well the energy E can't be lower than $-U_0$, so the energy must satisfy the condition $-U_0 \leq E < 0$, and the TISE for $|x| < a$ is

$$\frac{d^2\psi}{dx^2} + \lambda^2\psi = 0, \quad (4.17)$$

where $\lambda = \sqrt{2m(U_0 + E)}/\hbar = \sqrt{2m(U_0 - |E|)}/\hbar$ and $|E| = -E$ is the *binding energy* of the particle.

The solutions to Equation 4.17 include both even and odd oscillatory functions:

$$\psi(x) = B \cos(\lambda x) + C \sin(\lambda x) \quad \text{for } |x| > a$$

Putting these together, the solution for bound states, with $E < 0$, is

$$\psi(x) = \begin{cases} A_- e^{\kappa x} & \text{for } x < -a, \\ B \cos(\lambda x) + C \sin(\lambda x) & \text{for } |x| > a, \\ A_+ e^{-\kappa x} & \text{for } x > a. \end{cases}$$

We can use the symmetry however to find that $A_- = A_+ \equiv A$ because of the definite parity. This also implies that either $B = 0$ or $C = 0$.

We need to impose the BCs to find the coefficients, and specifically we need to impose the continuity of ψ and of $d\psi/dx$.

Continuity:

- Even function:

$$B \cos(\lambda a) = Ae^{-\kappa a}; \quad -\lambda B \sin(\lambda a) = -\kappa Ae^{-\kappa a},$$

and dividing the second equation above by the first:

$$\tan(\lambda a) = \kappa/\lambda. \quad (4.18)$$

- Odd function:

$$C \sin(\lambda a) = Ae^{-\kappa a}; \quad \lambda C \cos(\lambda a) = -\kappa Ae^{-\kappa a},$$

and again dividing the two equations:

$$-\cot(\lambda a) = \kappa/\lambda. \quad (4.19)$$

The energy levels are found by solving the *transcendental equations* Equation 4.18 and Equation 4.19, either numerically or graphically, cause we can't solve these analytically unfortunately.

It helps to visualise these functions and find graphically the solutions, corresponding to the points at the intersections of $f(\lambda) = \kappa/\lambda$ with the even and odd transcendental functions $g_{\pm}(\lambda) = \{\cos, \sin\}(\lambda a)$, as in Figure 4.8.

As expected for bound states, we obtain a set of *discrete* eigenvalues given by the quantum numbers in Figure 4.8, with **alternating even and odd** eigenfunctions - even for $n = 1, 3, \dots$ and odd for $n = 2, 4, \dots$

The spacing between the intersections is also increasing with n , just like in the infinite well potential, so *the eigenvalues are not equally spaced*.

It is also worth noting that in contrast with the infinite well potential, here we have a **finite number of eigenvalues** (e.g. 6 states represented in Figure 4.8 or 3 represented in Figure 4.9), but how many we have in general will depend on the strength of the potential. This is in contrast with the case of the infinite well potential where we had infinite discrete energy levels. In fact, there will be a finite n for which $E > 0$ and the state is no longer bound.

How does the number of eigenvalues change when we change the depth of the potential or the width?

You can normalize the wavefunction to find the values of the coefficients, but I will skip this here and consider instead what happens in two limit cases, where we make the well way deeper or where we make it very narrow.

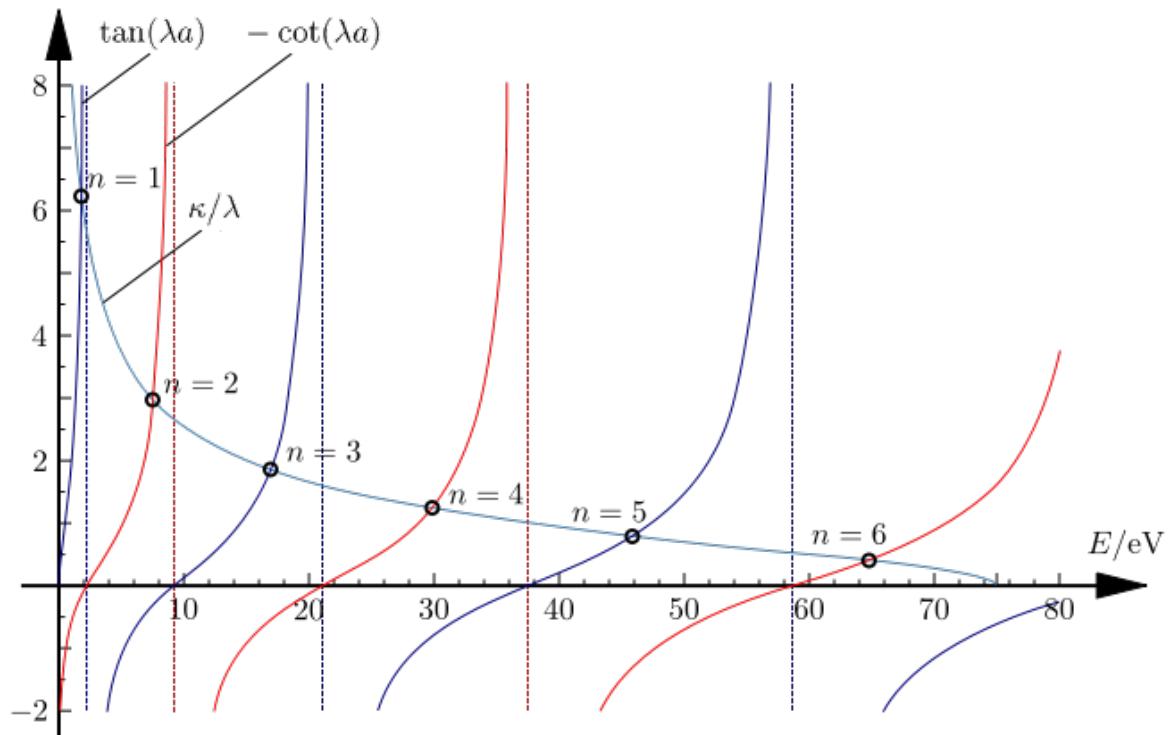


Figure 4.8: Graphic representation of functions in Equation 4.18 and Equation 4.19 obtained imposing the BCs. The intersections between both $-\cot(\lambda a)$ and $\tan(\lambda a)$ with κ/λ represent the solutions.

i Note

Note that we could have defined the square well potential with a *different zero of the energy*, having zero potential inside the well and a finite potential U_0 outside the well. The procedure and the eigenfunctions found are exactly the same, what would change in that case would be the energies in λ and κ to account for a shift of U_0 , i.e. we would have $\lambda = \sqrt{2mE}/\hbar$ and $\kappa = \sqrt{-2m(E - U_0)}/\hbar$.

Deep well limit

With the above consideration, let's conveniently shift the square well "up" so its outer walls are at a finite potential of U_0 and the inside of the well has zero potential. This is convenient just because earlier we considered this convention in the infinite well potential, and in fact if we keep increasing the potential U_0 , in the limit $U_0 \rightarrow \infty$, the infinite well potential is exactly what we should obtain.

So let's see what happens if we increase U_0 : κ will increase too with it, and will push the curve $f(\lambda)$ upward uniformly, with two effects:

- increase in the number of eigenvalues (intersections with the $g_{\pm}(\lambda)$ functions)
- the intersections move towards the asymptotic limits of the functions $g_{\pm}(\lambda)$ (see straight lines in Figure 4.8), i.e. $\lambda a = (n+1/2)\pi$ or $n\pi$ for the even and odd solutions respectively: you can verify that these correspond — as they should — to the half-wavelength solutions of the infinite square well.

Furthermore, the divergence of the $f(\lambda)$ function for small λ , while the $g_+(\lambda) = \tan(\lambda a)$ function heads to zero, means there is always guaranteed to be at least one bound state, of cos-type. The solutions hence have the same main features as seen in the infinite well: an even-symmetric ground state, excited states alternating in parity, and in the large- λ limit where $f(\lambda)$ is nearly flat, quadratically increasing state energy $E = (\hbar\lambda)^2/2m$.

Narrow potential limit

In the limit of infinitely narrow potential, i.e. $a \rightarrow 0$, we have the delta-function potential seen in the previous section.

The decaying exponentials in the walls also exhibit the delta-potential feature that they penetrate further for high-energy states, eventually achieving boundless range as $U_0 - E \rightarrow 0$, i.e. to freedom.

4.6 The harmonic well potential (quantum harmonic oscillator)

Many potentials could be approximated as harmonic oscillators by doing a Taylor expansion to the second order, around a minimum of the potentials. Quantum harmonic oscillators (QHO)

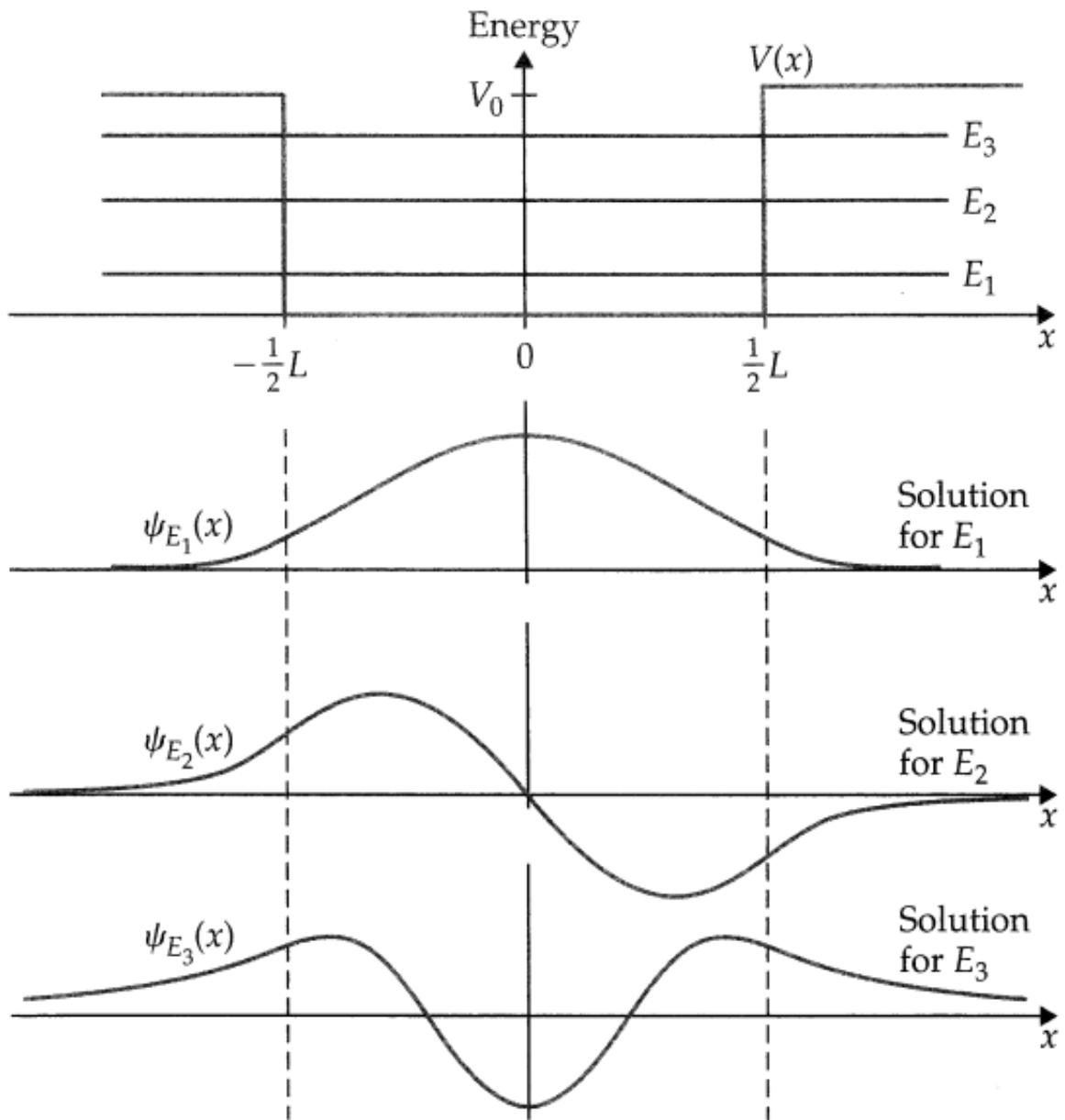


Figure 4.9: Finite square well potential and corresponding eigenvalues and eigenstates (3 possible states in this case).

are *extremely* useful in physics, and real vibrations can be indeed described with harmonic approximations.

The form of the potential for the harmonic oscillator is

$$U(x) = \frac{1}{2}Kx^2 = \frac{1}{2}m\omega^2x^2, \quad (4.20)$$

therefore the Hamiltonian and the corresponding TISE are, respectively:

$$\hat{H} = [\hat{p}^2 + (m\omega\hat{x})^2]/2m, \quad (4.21)$$

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \frac{1}{2}m\omega^2x^2\psi = E\psi. \quad (4.22)$$

Solving this TISE with a differential equation approach in the wave formalism is a bit of a nightmare and I do not expect you to do it, not using this method at least! We can take a guess and see that due to the symmetry of the system we would expect a ground state that is symmetric around the centre of the potential and has no nodes. The simplest form for this would be a Gaussian, and in fact the ground state of the QHO is given by Equation 4.24, and the eigenfunctions, with $n = 0, 1, 2, \dots$ are in general given by Equation 4.25, while the eigenvalues are Equation 4.23.

Eigenvalues for the QHO

$$E_n = (n + \frac{1}{2})\hbar\omega \quad (4.23)$$

- Note that the ground state energy is not zero! As for other bound states, we have a *zero-point*. For the QHO this is $E_0 = \frac{\hbar\omega}{2}$.
- Note that the eigenvalues are all equally spaced in the QHO, in contrast with the well potentials encountered so far. The energy spacing between consecutive energy levels is $\hbar\omega$.

This also seems an apt point to mention a general feature of all the bound-state wavefunctions we have studied: their ground state — $n = 1$ for the infinite well, $n = 0$ for the quantum (simple) harmonic oscillator (QHO) — always has a finite **zero-point energy** $E_{ground} > 0$.

The zero-point energy comes from the Heisenberg's uncertainty principle. In this case, imagine that your particle is exactly at the bottom of the well, so that the potential energy is 0. What would the implication be? If the position is very well defined, then the momentum must have a larger uncertainty, meaning that the kinetic energy would be higher and therefore the particle could not be *still* at the bottom of the harmonic potential! There is always going to be some energy even at the lowest possible state.

Ground state and eigenfunctions of the QHO

The ground state of the QHO is the Gaussian

$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left\{-\frac{m\omega}{2\hbar}x^2\right\}, \quad (4.24)$$

and the n -th excited state has the following eigenfunctions, built from the Hermite polynomials $H_n(\xi)$, where $\xi = \sqrt{m\omega/\hbar}x$,

$$\psi_n(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \frac{1}{\sqrt{2^n n!}} H_n(\xi) e^{-\xi^2/2}. \quad (4.25)$$

The Hermite polynomials are defined by repeated differentiation of a Gaussian:

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}. \quad (4.26)$$

The Hermite polynomials alternate in parity and add more nodes to the wavefunction with every increment of n .

It's neat to note how this works, since usually differentiation is thought of as reducing polynomial orders, but here it seems to increase them. The trick is in the Gaussian and the product rule: each time a derivative is taken, the Gaussian term produces a new multiplying factor of $-2x$, and hence raises the highest polynomial power by one. At the same time, the product rule requires that any already-existing polynomial terms be separately differentiated, which reduces their powers by one. The pre-multiplying inverse-Gaussian in Equation 4.26 is just there to remove the differentiation/multiplication-resilient Gaussian term; we then multiply the square-root of this back into place for our QHO solution.

The Hermite polynomials represent a “classic” orthogonal set, just like the Fourier terms that we have encountered in the finite square well. Another orthogonal set is given by the Legendre functions, which we will encounter when we deal with the angular momentum.

The eigenvalues and eigenfunctions are shown in Figure 4.10.

This is a good point to introduce the *matrix formalism*, i.e. the *bra-ket notation*, or *Dirac notation* that we mentioned at the beginning of the course, and solve the Schrödinger equation (Equation 4.22) using this. This formalism is also way more convenient to solve problems and exercises on the QHO (unless you really like doing integrals...), as we will see in an example at the end of the chapter. It'll also be very useful to deal with angular momentum later on!

I will skip all the formalism to start with, and we'll get to it in the following chapter.

Let's just build some useful rules with a new notation.

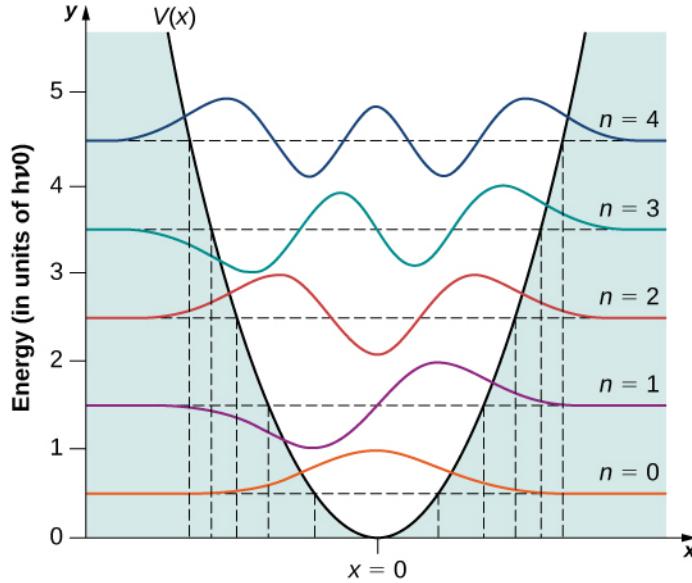


Figure 4.10: Eigenvalues and eigenstates (shifted to the levels of eigenvalues) of the QHO.

4.6.1 An introduction to the Dirac notation

Let's say that instead of using the notation $\psi_n(x)$ to describe the eigenfunctions (using the wave formalism), we replace these with the notation $|n\rangle$, where these are now vectors (called *kets*) indicating the states of Equation 4.25. Physically, we are not changing anything, only the way we decide to formalise them mathematically.

So now we have the ground state $|0\rangle$ and the excited states $|1\rangle$, $|2\rangle$, ... with energies given by Equation 4.23.

Let's call these *number states*.

The Hermitian conjugate of a state $|n\rangle$ is described by another vector (called *bra*) with the notation $\langle n|$, such that $\langle n| \equiv |n\rangle^\dagger$.

We can go up and down these states by providing or removing energy in multiples of $\hbar\omega$. We have like a “ladder” of states, equally spaced (see Figure 4.11).

How about we introduce some operators to move across this ladder of states? Enter the *ladder operators*! These are obtained from linear combinations of \hat{x} and \hat{p} , as follows.

! Ladder operators

The *ladder operators* are defined as:

$$\hat{a}_{\pm} \equiv \frac{1}{\sqrt{2\hbar m\omega}}(m\omega\hat{x} \mp i\hat{p}). \quad (4.27)$$

The position and momentum operator, in terms of the ladder operators, can be written as

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}(\hat{a}_+ + \hat{a}_-); \quad \hat{p} = i\sqrt{\frac{\hbar m\omega}{2}}(\hat{a}_+ - \hat{a}_-). \quad (4.28)$$

- Note that $\hat{a}_-^\dagger = \hat{a}_+$.
- The ladder operators do not commute: $[\hat{a}_-, \hat{a}_+] = 1$.
- The ladder operators are **not observables** so they are not Hermitian operators and do not correspond to any physical measurable quantity.
- Given a solution of the SE with energy E , application of \hat{a}_{\pm} return a solution with energy $E \pm \hbar\omega$, i.e. bring the state one level up or one down! However, since the lowest state is $|0\rangle$, $\hat{a}_-|0\rangle = 0$.

In general:

$$\hat{a}_-|n\rangle = \sqrt{n}|n-1\rangle; \quad \hat{a}_+|n\rangle = \sqrt{n+1}|n+1\rangle \quad (4.29)$$

🔥 Caution

Be careful when applying the operators on the bras! You need to consider the conjugate of both operator and state vector, so:

$$\langle n|\hat{a}_- = (\hat{a}_+|n\rangle)^\dagger = \sqrt{n+1}\langle n+1|; \quad \langle n|\hat{a}_+ = (\hat{a}_-|n\rangle)^\dagger = \sqrt{n}\langle n-1|. \quad (4.30)$$

Using the ladder operators, the Hamiltonian of the QHO (Equation 4.21) can be rewritten as

$$\hat{H} = \hbar\omega(\hat{a}_+\hat{a}_- + \frac{1}{2}), \quad \hat{H} = \hbar\omega(\hat{a}_-\hat{a}_+ - \frac{1}{2}) \quad (4.31)$$

It is therefore convenient to define the *number operator* \hat{n} .

ℹ️ Number operator and number states

The number operator

$$\hat{n} = \hat{a}_+\hat{a}_- \quad (4.32)$$

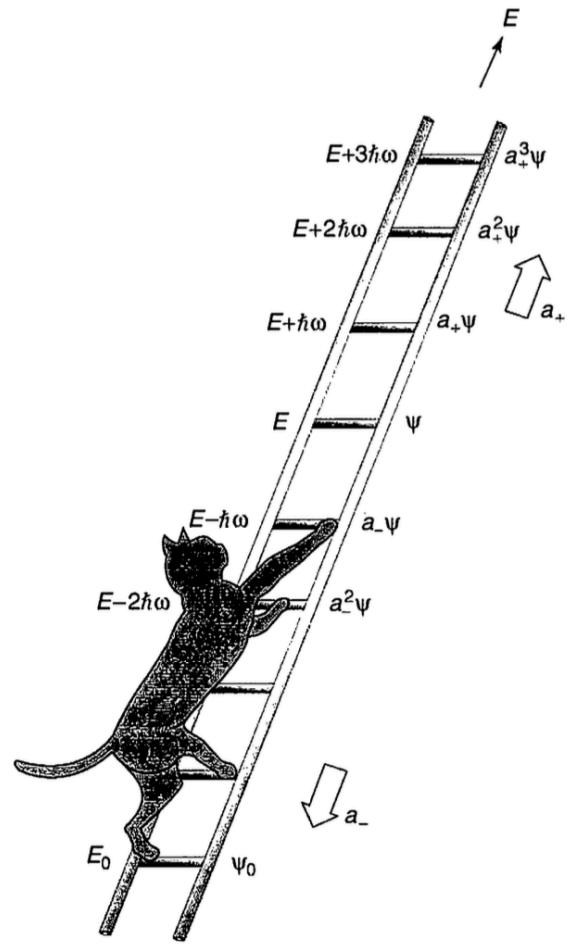


Figure 4.11: Graphic representation of the “ladder” of states of the QHO.

returns the *principal quantum number* n :

$$\hat{n}|n\rangle = n|n\rangle. \quad (4.33)$$

This is an eigenvalue equation, and the number states $|n\rangle$ constitute a *complete orthonormal basis*:

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

4.6.1.1 Exercise

Show Equation 4.31 starting from Equation 4.21.

Solution

Using Equation 4.27:

$$\begin{aligned}\hat{a}_-\hat{a}_+ &= \frac{1}{2\hbar m\omega}(i\hat{p} + m\omega\hat{x})(-i\hat{p} + m\omega\hat{x}) \\ &= \frac{1}{2\hbar m\omega}(\hat{p}^2 + (m\omega\hat{x})^2 - im\omega[\hat{x}, \hat{p}]) \\ &= \frac{1}{2\hbar m\omega}[\hat{p}^2 + (m\omega\hat{x})^2] + \frac{1}{2} \\ &= \frac{1}{\hbar\omega}\hat{H} + \frac{1}{2}\end{aligned}$$

Rearranging this: $\hat{H} = \hbar\omega(\hat{a}_-\hat{a}_+ - \frac{1}{2})$.

Similarly

$$\hat{a}_+\hat{a}_- = \frac{1}{\hbar\omega}\hat{H} - \frac{1}{2} \implies \hat{H} = \hbar\omega(\hat{a}_+\hat{a}_- + \frac{1}{2}).$$

4.6.2 Solving the harmonic oscillator

4.6.2.1 Ground state

We can find the ground state explicitly, knowing that $\hat{a}_-|0\rangle = 0$ and using the definition of \hat{a} in terms of \hat{p} and \hat{x} from Equation 4.27.

Moving back to the waves formalism:

$$\hat{a}_-\psi_0 \propto (i\hat{p}_m\omega\hat{x})\psi_0 = 0 \implies \frac{d\psi_0}{dx} = -\frac{m\omega}{\hbar}x\psi_0.$$

This has solution

$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left\{-\frac{m\omega}{2\hbar}x^2\right\},$$

which is a *normalised Gaussian*.

We can verify the ground state energy from $\hat{H}\psi_0$ with the Hamiltonian of Equation 4.31, using $\hat{a}_-\psi = 0$:

$$\hbar\omega(\hat{a}_+\hat{a}_- + \frac{1}{2})\psi_0 = E_0\psi_0 = \hbar\omega/2. \quad (4.34)$$

Excited states eigenenergies

In general, if we apply Equation 4.31 to the state $|n\rangle$, we find

$$\hbar\omega(\hat{n} + \frac{1}{2})|n\rangle = \hbar\omega(n + \frac{1}{2})|n\rangle = E_n|n\rangle \implies E_n = \hbar\omega(n + \frac{1}{2}). \quad (4.35)$$

4.6.2.2 Expectation values calculations

Typical exercises with QHO and other potentials in general require to calculate expectation values of various operators, e.g. of position, momentum, etc, on eigenstates or combinations thereof.

We saw that in the waves formalism the expectation value of an operator \hat{A} on a state $\psi_n(x)$ is

$$\langle \hat{A} \rangle = \int \psi_n^*(x)\hat{A}\psi_n(x)dx. \quad (4.36)$$

How do we calculate the equivalent with the Dirac notation? We use the Hermitian conjugate of the *ket* $|n\rangle$, called a *bra*, $\langle n| \equiv |n\rangle^\dagger$.

So when we want to calculate an expectation value, we still have a sandwich in the *bra-ket* (now you know why they have these names), as follows:

$$\langle \hat{A} \rangle = \langle n|\hat{A}|n\rangle. \quad (4.37)$$

Let's consider a typical example with the QHO, both with the wave notation and the Dirac notation.

At the end of the chapter we'll see more on the Dirac notation.

4.6.2.3 Exercise

Calculate $\langle \hat{p} \rangle$ on the ground state of the QHO, using two methods:

1. Using ψ_0 as in Equation 4.24 and the wave formalism
2. Using the Dirac notation with the state $|0\rangle$.

Solution

1. Let's use Equation 4.24 and the formula for the expectation values Equation 4.36, with $\hat{p} = -i\hbar \frac{d}{dx}$:

$$\begin{aligned}\langle \hat{p} \rangle &= \int_{-\infty}^{\infty} dx \psi_0^* \hat{p} \psi_0 \\ &= \sqrt{\frac{m\omega}{\pi\hbar}} \int_{-\infty}^{\infty} dx i\hbar \frac{m\omega}{\hbar} x \exp\left\{-\frac{m\omega}{\hbar}x^2\right\} = 0\end{aligned},$$

where we can say that it is 0 because of the parity of the function: it is odd, so the integrals in the positive and negative coordinates cancel each other.

This method requires more familiarity with evaluation of derivatives and integrals. Let's look at the second method.

2. Let's use \hat{p} as defined in Equation 4.28, evaluated on the state $|0\rangle$:

$$\begin{aligned}\langle \hat{p} \rangle &\propto \langle 0 | (\hat{a}_+ - \hat{a}_-) | 0 \rangle = \langle 0 | \hat{a}_+ | 0 \rangle - \langle 0 | \hat{a}_- | 0 \rangle \\ &\propto \langle 1 | 0 \rangle - 0 = 0.\end{aligned}$$

This method requires to know (check formula sheet) how the ladder operators act on number states, and to use the fact that these are orthonormal. Easy peasy!

4.7 Rules for wavefunctions in general potentials

The nature of QM is that once we go beyond these simple potentials it rapidly becomes impossible to calculate analytic solutions: we already saw this for even a case as simple as the finite square well. It's a nice example of Rutherford's comment that "All of physics is either impossible or trivial.". Rutherford went on to say that "It is impossible until you understand it, and then it becomes trivial.", so in that spirit this section is dedicated to showing you that even if a solution cannot be calculated in the nice closed form we're used to, and we have to resort to numerical computation instead if we want accurate numbers, it is very possible — trivial, Rutherford would say — to understand why they look the way they do.

The arguments we will use are based on our existing set of boundary and integrability conditions, plus features given by the TISE (Equation 3.23):

1. A stationary wavefunction $\psi(x)$ must always be continuous, i.e. no sudden jumps in amplitude across differentially small distances dx .
2. It must also be differentiable, i.e. smooth, except at infinite-sized potential steps: for those a kink is permitted.
3. All wavefunctions must go to zero at infinite distance, for integrability: as bound states are defined to have $E < V$ at infinity, this means that the wavefunction must go to zero in classically forbidden regions.
4. Any general potential can be built by concatenating flat-potential wavefunctions for each flat region, and matching them via the boundary conditions. You can think of general potentials as a stack of thin square potentials: same rules apply, i.e. continuous and smooth wavefunctions. *Every wavefunction is locally a real or complex exponential, connected to the next piece via BCs.*

This is the essence of the **WKB approximation** technique - but for sketching purposes you can imagine a roughly equivalent piecewise-flat potential and fudge the smoothing of the steps.

5. The TISE can be written in the form: $\frac{d^2\psi}{dx^2}/\psi = -\frac{2m}{\hbar^2}(E - U) \Rightarrow$ curvature is toward the x -axis, i.e. toward $\psi = 0$, in classically allowed regions $E - U > 0$, away from it in forbidden ones, i.e. away from the $\psi = 0$ axis in disallowed regions.

For a fixed $E - U$, the degree of curvature is proportional to the value of ψ : a large amplitude will be highly curved, e.g. the curvature of the cos function has its maximum value where the cosine is largest.

6. The wavefunction is oscillatory where net positive kinetic energy $E - U$; evanescent/exponential where negative.

In a region of x with a flat potential, $U(x) = \text{const}$, the solution will be wavelike and oscillatory if $E > U$, i.e. it is classically allowed, and an exponential in classically forbidden regions with $E < U$. Rule 3 forces these exponentials to decay at the edges of the potential toward $\pm\infty$.

7. The typical inverse scale of both the oscillating and decaying solutions — respectively the wavenumber and the decay constant — becomes larger as the magnitude of excess or deficient energy with respect to the potential increases.

Curvature/frequency scale with energy difference: $-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} x^2 \rightarrow \frac{\hbar^2 k^2}{2m} = (E - U)$, with $\sim 1/k$ as more kinetic energy $E - U$ in allowed regions leads to larger wavenumbers and hence smaller wavelengths; a greater energy deficit $U - E$ in forbidden regions gives a more rapidly decaying exponential.

8. *Amplitude is smaller where KE is larger (more velocity, less likely to be found there)*
Counterintuitively, in allowed regions the envelope of the wavefunction will be smaller where the kinetic energy $E - U$ is larger: intuitively, where the particle has more velocity, it spends less of its time and is less likely to be found.
9. *Principal quantum number $n = 1, \dots \implies n$ bumps or $n - 1$ nodes.* Solutions always come in a perhaps finite stack indexed by the *principle quantum number n* . The ground state, with lowest n , always has zero nodes ($\psi = 0$ points); each higher state adds one node. Alternatively, the ground state only has one “bump”; the higher states add a bump for each unit increase in n .
10. *Remember definite parity for symmetric potentials.*

If the potential has even symmetry, the stationary solutions will have definite parity, i.e. they will be exactly odd or even functions. If the potential doesn't have even symmetry, the definite-energy wavefunctions will generally be asymmetric.

With these rules in hand, you should be able to render an acceptable guess of how the n th energy level of stationary wavefunction should look in any potential: some examples are given in Figure 4.13 and Figure 4.14.

Below is (embedded) a document that goes more in detail and can be useful for sketching wavefunctions. A summary is reported below anyway. (If you are reading this from a non-html format, you can find the file on Moodle or at [this link](#).)

4.7.1 Examples of general potentials

4.7.2 Periodic boundary conditions

In some cases, for instance for particles in a ring, it is possible to not have *any* boundary conditions that fix wavefunction value at a certain place.

In such cases require equal value and derivative in $\varphi \rightarrow \varphi + 2\pi$ round the ring.

For a flat $V(\varphi)$, the solution is a plane wave $\psi(\varphi) \sim e^{im\varphi} = e^{im(\varphi+2\pi)}$.

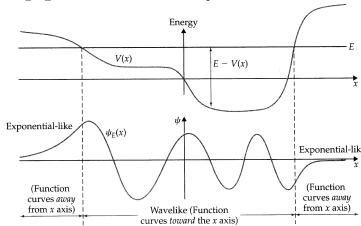
Require $e^{2\pi im} = 1 \implies$ integer m . This means that we must fit the whole wavelengths around the ring (as in Figure 4.15), and this is what we have seen for the Bohr atom and what we will see for the spherical harmonics.

The BC is now geometry rather than potential.

Rules for Sketching Wavefunctions

(Adapted from "Particles Behave Like Waves" by Thomas A. Moore.)

1. Solutions to the Schrödinger equation curve *toward* the x axis in classically allowed regions (where $E - V(x) > 0$) and *away* from the x axis in classically forbidden regions (where $E - V(x) < 0$).
 - Implication: Solutions are *wavelike* in classically allowed regions and *exponential-like* in forbidden regions.
 - Follows from: $-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + V(x)\psi(x) = E\psi(x) \implies \frac{d^2\psi}{dx^2} = -\frac{2m}{\hbar^2}[E - V(x)]\psi(x)$
 - The classical turning points, where $E - V(x) = 0$, are *inflection points*.
2. The curvature of solutions increases with $|E - V(x)|$. Note that $|E - V(x)|$ corresponds to the distance between the energy line and the potential energy curve on a potential energy graph. Implications:
 - In classically allowed regions, the *local wavelength* of the wavelike part of the solution gets shorter (or local wavenumber k gets larger) as $|E - V(x)|$ increases.
 - Recall that $-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} \rightarrow \frac{\hbar^2 k^2}{2m}$ and $k = \frac{2\pi}{\lambda}$.
 - In classically forbidden regions, the *length* of any exponential-like tail gets shorter as $|E - V(x)|$ increases.
3. The local amplitude of the wavelike part of a solution *decreases* as the value of $|E - V(x)|$ increases.
 - You are more likely to find the particle in regions where its momentum (therefore velocity) is lower.
 - The following figure illustrates the above points:



1

Figure 4.12: Rules for sketching wavefunctions. Unsure who the author of this is.

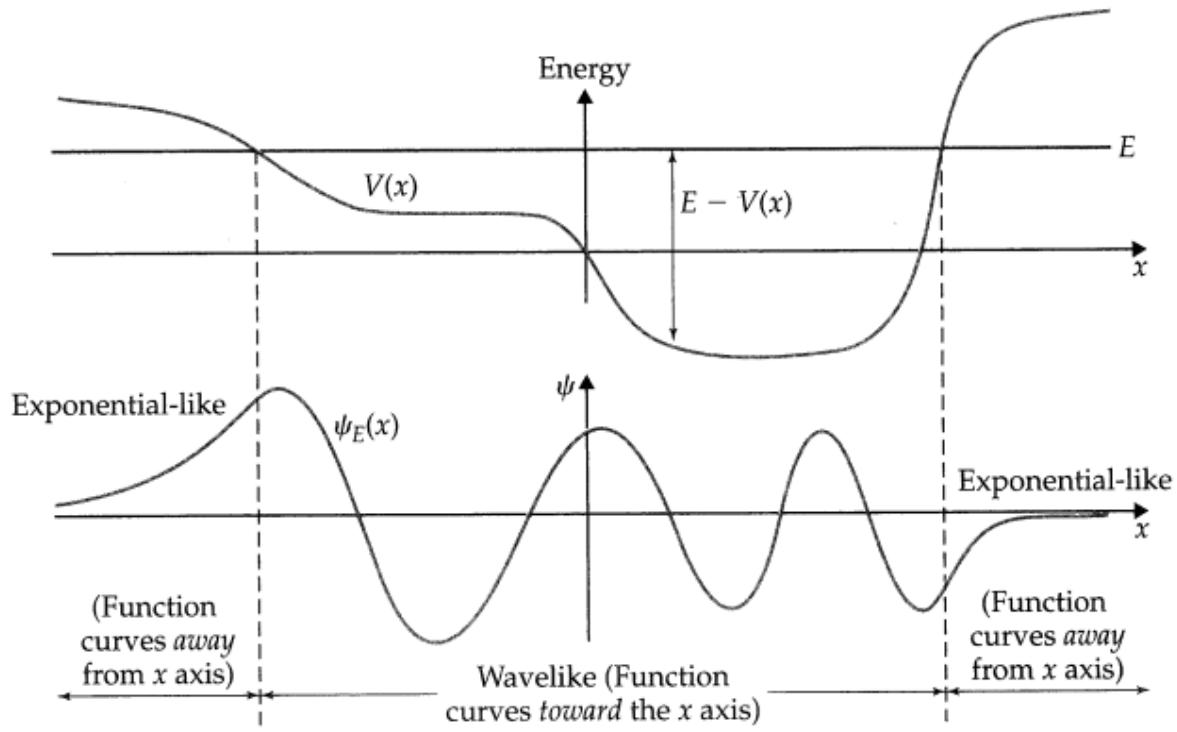


Figure 4.13: Example of wavefunction associated to a potential. The turning points mark the limit between oscillatory and exponential solutions. In the oscillatory part, the amplitude is larger and the frequency smaller where the KE is smaller (left side).

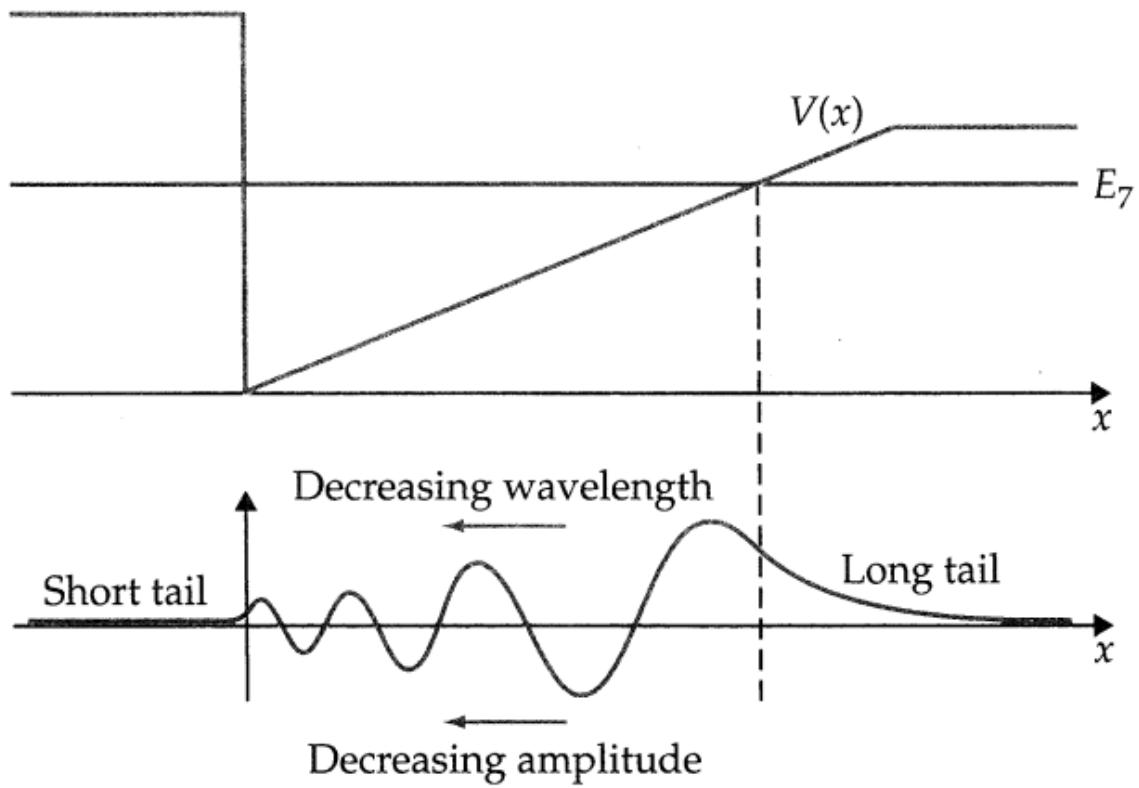
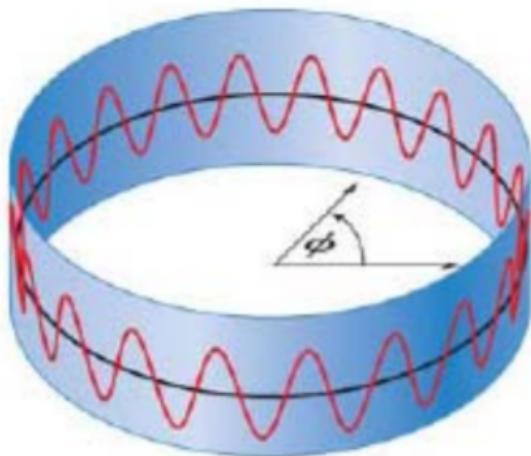
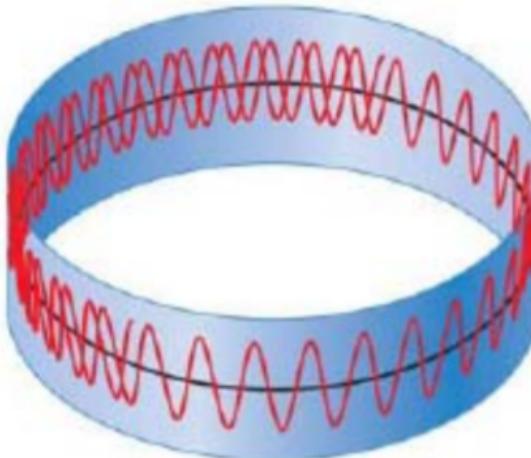


Figure 4.14: Example of a “linear” potential and associated wavefunction.



$$m_l = \pm \text{integer}$$



$$m_l \neq \pm \text{integer}$$

Figure 4.15: Example of a wavefunction in a periodic (ring) potential.

4.8 More on the Dirac notation

At the beginning of the course we saw that there are two equivalent formalisms to describe quantum mechanics: wave formalism and matrix formalism.

Until the QHO, we have used the former, which uses wavefunctions dependent on spatial coordinates (or coordinates in momentum space), and where operators are determined by functions acting on the wavefunctions.

With the matrix formalism however, we have states described by *row and column vectors* (respectively “bra” and “ket”) and operators described by *matrices* acting on the states.

In the Dirac notation, every wavefunction can be written in the form of a **state vector** $|\Psi\rangle$ in a given basis (e.g. position or momentum).

4.8.1 Bra, ket and operators

Kets are column vectors:

$$|x\rangle = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \end{pmatrix},$$

where x_i are the components of the vector.

Bras are row vectors and the Hermitian conjugate of ket-vectors (i.e. transpose the vector and take the complex conjugates of its elements):

$$\langle x| = (x_1^* \ x_2^* \ x_3^* \ \dots) = |x\rangle^\dagger.$$

The dagger (\dagger) indicates the Hermitian conjugate.

Operators are represented by matrices, of the form

$$\hat{A} = \begin{pmatrix} A_{11} & A_{12} & A_{13} & \dots \\ A_{21} & A_{22} & A_{23} & \dots \\ A_{31} & A_{32} & A_{33} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}$$

So now you understand why earlier on we said that an operator \hat{A}^\dagger is defined as the *conjugate transpose* of \hat{A} : you need to consider the complex conjugate and transpose the matrix.

Eigenvalue equation with states

The **eigenvalue equation** for an operator \hat{A} is

$$\hat{A}|\phi_i\rangle = a_i|\phi_i\rangle, \quad (4.38)$$

where a_i are the eigenvalues corresponding to the possible outcomes of the measurement, and $|\phi_i\rangle$ are the eigenstates.

4.8.2 Inner and outer product

The *inner product* between ket and bra vectors returns a *scalar* (i.e. a number):

$$\langle\psi|\phi\rangle = (\psi_1^* \ \psi_2^* \ \psi_3^* \ ...) \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \dots \end{pmatrix} = \sum_i \psi_i \phi_i^* \quad (4.39)$$

The *outer product* gives a matrix:

$$|\psi\rangle\langle\phi| = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \dots \end{pmatrix} (\phi_1^* \ \phi_2^* \ \phi_3^* \ ...) = \begin{pmatrix} \psi_1\phi_1^* & \psi_1\phi_2^* & \psi_1\phi_3^* & \dots \\ \psi_2\phi_1^* & \psi_2\phi_2^* & \psi_2\phi_3^* & \dots \\ \psi_3\phi_1^* & \psi_3\phi_2^* & \psi_3\phi_3^* & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \quad (4.40)$$

4.8.3 Mapping to waves formalism

We can use these ingredients to see how we can go from waves to matrix formalism and viceversa.

4.8.3.1 Eigenstates as a basis

The linear superposition of wavefunctions seen earlier can be applied also to the vector notation with kets or bras:

$$|\Psi\rangle = \sum_i c_i |\phi_i\rangle, \quad (4.41)$$

where $\{\phi_i\}$ are eigenstates of a Hermitian operator, forming a *complete, orthonormal basis*, and c_i are the amplitudes.

Note

A composite state no longer satisfies the eigenvalue equation Equation 4.38:

$$\begin{aligned}\hat{A}|\Psi\rangle &= \sum_i c_i \hat{A}|\phi_i\rangle = \sum c_i a_i |\phi_i\rangle \\ &\neq a \sum_i c_i |\phi_i\rangle = a|\Psi\rangle.\end{aligned}$$

States built from non-degenerate eigenstates do not have *definite quantum numbers!*

When we consider a wavefunction $\psi(x)$, this is defined in the position space, so we have to use the x basis to write it. How do we give this information in the Dirac notation? We need to use the basis $\{|x\rangle\}$, i.e. a set of quantum states, one for each x , satisfying the **identity operator** in Equation 4.42. The same could be done for the p (momentum) basis.

Identity operators

For position and momentum basis, the identity operators can be defined as

$$\hat{I} = \int dx |x\rangle \langle x| = \int dp |p\rangle \langle p|. \quad (4.42)$$

In general, for any *complete orthonormal basis* with vectors $\{|\phi_i\rangle\}$ (with i integer numbers),

$$\hat{I} = \sum_i |\phi_i\rangle \langle \phi_i|.$$

In the basis of *position eigenfunctions* we can then define

$$|\psi\rangle = \hat{I}|\psi\rangle = \int dx |x\rangle \langle x|\psi\rangle = \int dx \psi(x) |x\rangle,$$

where

$$\psi(x) = \langle x|\psi\rangle. \quad (4.43)$$

4.8.3.2 Orthogonality of eigenstates

The *overlap* between two states is given by the inner product

$$\langle\psi|\phi\rangle = \int dx \psi^*(x) \phi(x). \quad (4.44)$$

States of an orthonormal basis $\{|\phi_i\rangle\}$ satisfy the following properties:

$$\langle \phi_n | \phi_k \rangle = \int \phi_n^* \phi_k dx = \delta_{nk},$$

so the inner product of a wavefunction with itself is its normalization:

$$\langle \psi | \psi \rangle = \int |\psi(x)|^2 dx = 1$$

4.8.3.3 Eigenstates probabilities

Given a state $|\Psi\rangle$ written as superposition of eigenstates (Equation 4.41), the inner product with itself is

$$\begin{aligned} \langle \Psi | \Psi \rangle &= \left\langle \sum_i c_i \phi_i \right| \sum_j c_j \phi_j \rangle = \sum_{i,j} c_i^* c_j \langle \phi_i | \phi_j \rangle = \sum_{i,j} c_i^* c_j \delta_{i,j} \\ &\quad \sum_i |c_i|^2 = 1. \end{aligned} \tag{4.45}$$

This follows from the normalization condition. The modulus square of the amplitude c_i indicates the probability $P_i = |c_i|^2$ of finding the composite state in the i eigenstate component.

4.8.3.4 Expectation values

Remember the sandwich!

$$\begin{aligned} \langle \hat{A} \rangle &= \int \Psi^*(x) A \Psi(x) dx \\ &= \langle \Psi | \hat{A} | \Psi \rangle = \sum_i |c_i|^2 a_i = \sum_i P_i a_i, \end{aligned} \tag{4.46}$$

where we used Equation 4.45 in the last passages.

What does this result mean physically?

The expectation value is the probability-weighted average of eigenvalues. This is in line with the correspondence principle!

4.8.3.5 Individual measurements

What happens in repeated measurements of a single system?

Let's suppose that I have a *single system* with wavefunction Ψ and do a measurement of \hat{A} on it.

The measurement collapses the wavefunction to one of the eigenstates, i.e. $\Psi \rightarrow \phi_i$ with probability $P_i = |\phi_i|^2$, returning the quantum number a_i .

What happens if I repeat the measurement on the same system after the first one?

The system will keep being $\Psi = \phi_i$ with $P_i = 1$, so any consecutive measurement will return a_i .

To repeat the process we can apply a different operator with a different eigen-basis $\{\varphi_i\}$ repeats the process, via the decomposition $\phi_i = \sum_j \langle \phi_i | \varphi_j \rangle | \varphi_j \rangle$.

The “projected probability amplitude” applies between *any* two wavefunctions ψ and ψ' , e.g. due to sudden changes of potential function: $c = \sqrt{P} = \langle \psi | \psi' \rangle$.

4.8.4 Exercise

From Equation 4.41, show that the amplitudes c_i can be written as $c_i = \langle \phi_i | \Psi \rangle$.



Hint

Use the orthonormality of the eigenstates ψ_i .

Solution

Starting from $\Psi = \sum_j c_j \phi_j$, let's calculate the overlap with ψ_i :

$$\langle \phi_i | \Psi \rangle = \sum_j c_j \langle \phi_i | \phi_j \rangle = c_i,$$

where we used the orthonormality condition $\langle \phi_i | \phi_j \rangle = \delta_{ij}$.

4.9 Summary

- Potentials deeper than the energy of a state, E , localize it spatially to various degrees: a bound state
- Bound states can be described by the TISE, i.e. invariant spatial eigenfunctions with stationary-state time-evolution
- The form of the potential determines the form of the quantization through boundary conditions: eigenstates from continuous (free particle) \rightarrow discrete for any non-trivial potential
- General solution is (as always) a linear combination of orthogonal eigenstates... *even energy doesn't have to be definite!*
- For finite potentials (and delta...), wavefunction tunnels into classically forbidden regions... *but zero flux.*
- The Dirac notation uses the matrix formalism of QM and can be useful to avoid derivatives and integrals, i.e. in the QHO calculations.

5 Solutions of Schrödinger Equations in 1D - Scattering states

5.1 Scattering states

We now move on to scattering states, i.e. those for which $E > U(\pm\infty)$. We will specifically look at sharp, finite potential steps — scattering off infinite steps is rather boring: the incoming wavepacket just bounces back — and just like with bound states, this means the solutions are plane waves or decaying exponentials through each distinct $U(x) = \text{const}$ region.

Unlike the bound states, though, the boundary conditions no longer force the solutions to be real-valued (modulo a trivial phase), meaning that now the probability flux $\mathbf{J} \propto \psi^* \nabla \psi - \psi \nabla \psi^*$ will be non-zero and the probability “flows” through space.

Generally we saw that, like in the case of free particles, we should describe scattering states using wavepackets. However, this is algebraically very messy, so we’ll work with 1D spatial plane waves, which are a good approximation far from the potential anyway, but the phenomena are the same. Scattering states are not localised in space, so the wavefunctions are not normalisable.

In 1D, right-moving and left-moving plane waves (with t-dependence hidden) have respectively the following forms:

$$\psi_R \sim e^{ikx} \quad \psi_L \sim e^{-ikx}. \quad (5.1)$$

Let’s see then why a plane wave gives a non-zero flux.

Let’s suppose we have a right-propagating plane wave $\psi(x) = Ae^{ikx}$. The probability flux is

$$J = -\frac{i\hbar}{2m} \left(\psi^* \frac{d\psi}{dx} - \psi \frac{d\psi^*}{dx} \right) = \frac{\hbar k}{m} |A|^2 = v|A|^2. \quad (5.2)$$

This result applies to any region with constant k .

We will deal with the typical problem of plane waves moving between $\pm\infty$, encountering step potentials/finite barriers and scattering off them. Note that this is in the infinite-time, steady-state limit! Steps can be higher or lower than E , but not infinite, because in QM, due to

its probabilistic nature, there is no region inaccessible to our particle, regardless of its energy, since the matter wave associated with the particle is nonzero everywhere. This is due to tunneling, which we have already encountered in the delta-function potential and in the finite square well.

5.2 Scattering from a step potential

Let's start with the simple case of a step with a finite potential V , as in Figure 5.1.

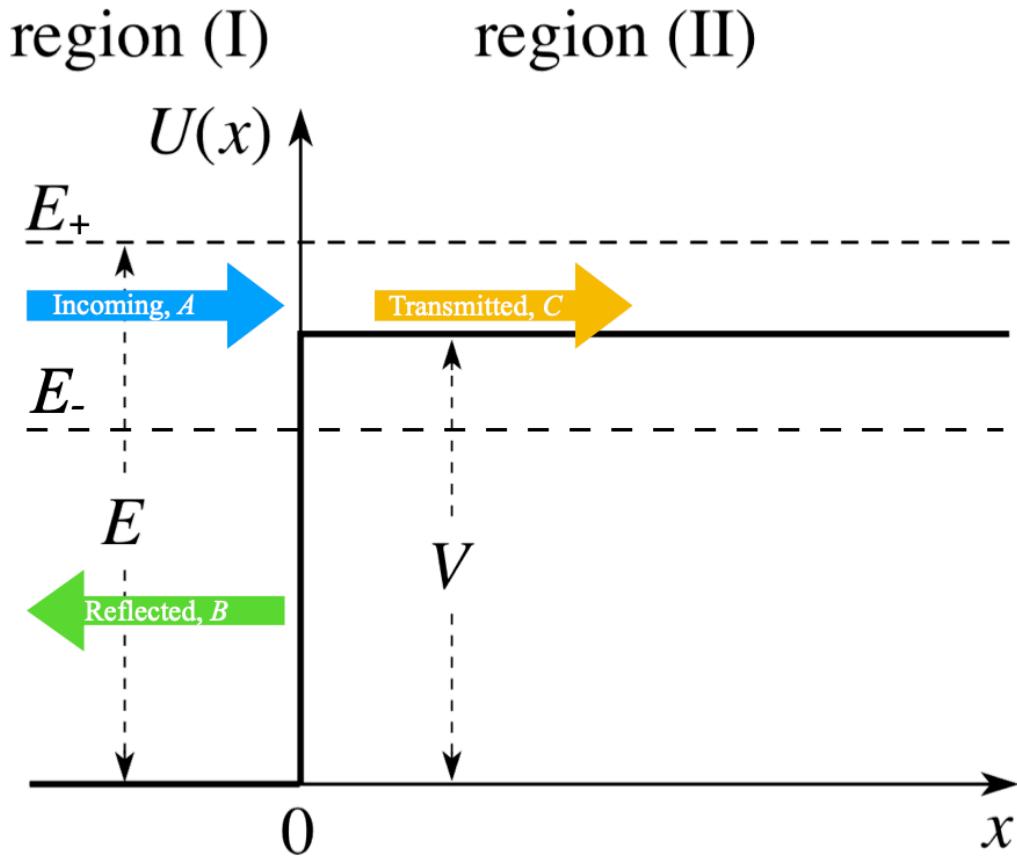


Figure 5.1: Step potential of height V , with two possible cases for incoming waves having energy $E_+ > V$ or $E_- < V$.

In Equation 5.2 we saw that there is a non-zero probability flux for plane waves. This applies to incoming wave, but also to the reflected and transmitted ones in Figure 5.1.

Incident, reflected and transmitted waves have respectively the following forms:

$$\psi_I(x) = Ae^{ik_1x}, \quad \psi_R(x) = Be^{-ik_1x}, \quad \psi_T = Ce^{ik_2x}, \quad (5.3)$$

and corresponding probability fluxes

$$J_I \propto k_1|A|^2, \quad J_R \propto k_1|B|^2, \quad J_T \propto k_2|C|^2. \quad (5.4)$$

The probability fluxes must satisfy the **flux conservation**:

$$J_I = J_R + J_T, \quad (5.5)$$

and this can be normalised to J_I to give the condition $R + T = 1$, where R and T are the flux reflection and transmission factors, i.e.

$$R = \frac{J_R}{J_I} = \frac{|B|^2}{|A|^2}, \quad T = \frac{J_T}{J_I} = \frac{k_2|C|^2}{k_1|A|^2}. \quad (5.6)$$

The flux conservation therefore implies:

$$R + T = 1 \implies k_2|C|^2 = k_1(|A|^2 - |B|^2). \quad (5.7)$$

5.2.1 Scattering from “allowed” potential step ($E > V$)

Let's consider the case of the incoming plane wave having energy $E = E_+ > V$, as in Figure 5.2.

Let's split the wavefunction in two parts, for regions I and II on the two sides of the step. These are respectively:

$$\begin{aligned} \psi_1(x) &= Ae^{ik_1x} + Be^{-ik_1x}, \quad \text{with } k_1 = \sqrt{2mE/\hbar^2}, \\ \psi_2(x) &= Ce^{ik_2x}, \quad \text{with } k_2 = \sqrt{2m(E-V)/\hbar^2}. \end{aligned}$$

As usual, let's impose the continuity condition for the wavefunction and its derivative at the interface between the two regions, i.e. $x = 0$:

$$\begin{aligned} \psi_1(0) &= \psi_2(0) \implies A + B = C, \\ \psi'_1(0) &= \psi'_2(0) \implies ik_1A - ik_1B = ik_2C \implies k_1(A - B) = k_2C. \end{aligned}$$

Rearranging these, we get:

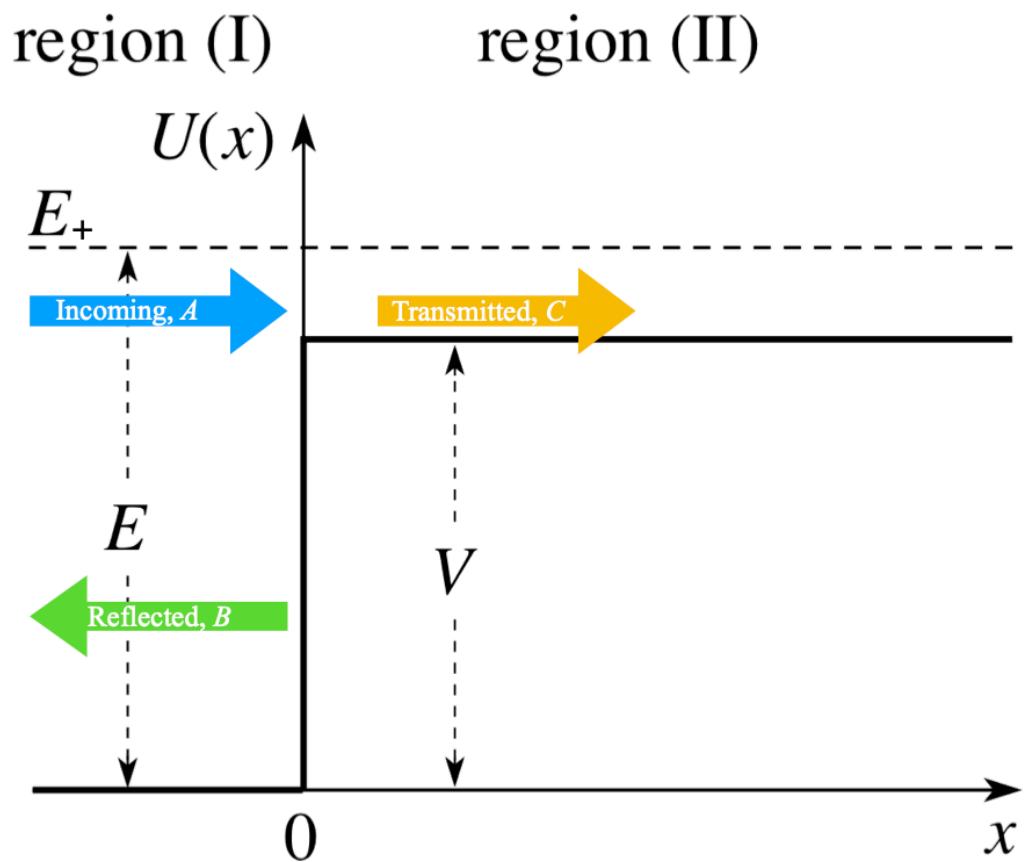


Figure 5.2: Step potential of height V hit by an incident wave of energy E_+ , which gets transmitted and reflected.

$$R = \frac{|B|^2}{|A|^2} = \frac{(k_1 - k_2)^2}{(k_1 + k_2)^2},$$

$$T = \frac{k_2 |C|^2}{k_1 |A|^2} = 1 - R = \frac{4k_1 k_2}{(k_1 + k_2)^2}.$$

5.2.1.1 Exercise

Consider $k_2 = k_1/2$ with $A = 1$: find wavefunctions and probability distributions. What are the max/min incoming probability densities?

5.2.2 Scattering from “forbidden” potential step ($E < V$)

Let's now consider the case of scattering if $E = E_- < V$, as in Figure 5.3.

Let's repeat the same procedure even in this case, writing the wavefunctions for the two regions on the sides of the step. The difference now is that in the not-allowed region (II) there can't be a traveling wave, but there will be an evanescent (exponentially decaying wave) due to tunneling. Therefore:

$$\psi_1(x) = A e^{ik_1 x} + B e^{-ik_1 x}, \quad \text{with } k_1 = \sqrt{2mE/\hbar^2},$$

$$\psi_2(x) = D e^{-k_2 x}, \quad \text{with } k_2 = \sqrt{2m(V-E)/\hbar^2}.$$

Imposing the continuity condition for the wavefunction and its derivative at the interface between the two regions, i.e. $x = 0$:

$$\psi_1(0) = \psi_2(0) \implies A + B = D,$$

$$\psi'_1(0) = \psi'_2(0) \implies ik_1 A - ik_1 B = -k_2 D \implies k_1(A - B) = -k_2 D.$$

Rearranging these, we get:

$$R = \frac{|B|^2}{|A|^2} = \frac{|ik_1 + k_2|^2}{|ik_1 - k_2|^2} = \frac{|z|^2}{|-z^*|^2} = 1,$$

$$T = 1 - R = 0,$$

as expected. There is no probability flux beyond the step, in region II, i.e. $T = 0 \implies J_T = 0$. However, this does not mean that there is no probability density in that “forbidden region”! As for the bound states, the probability distribution penetrates the wall with typical length scale, or **skin depth**,

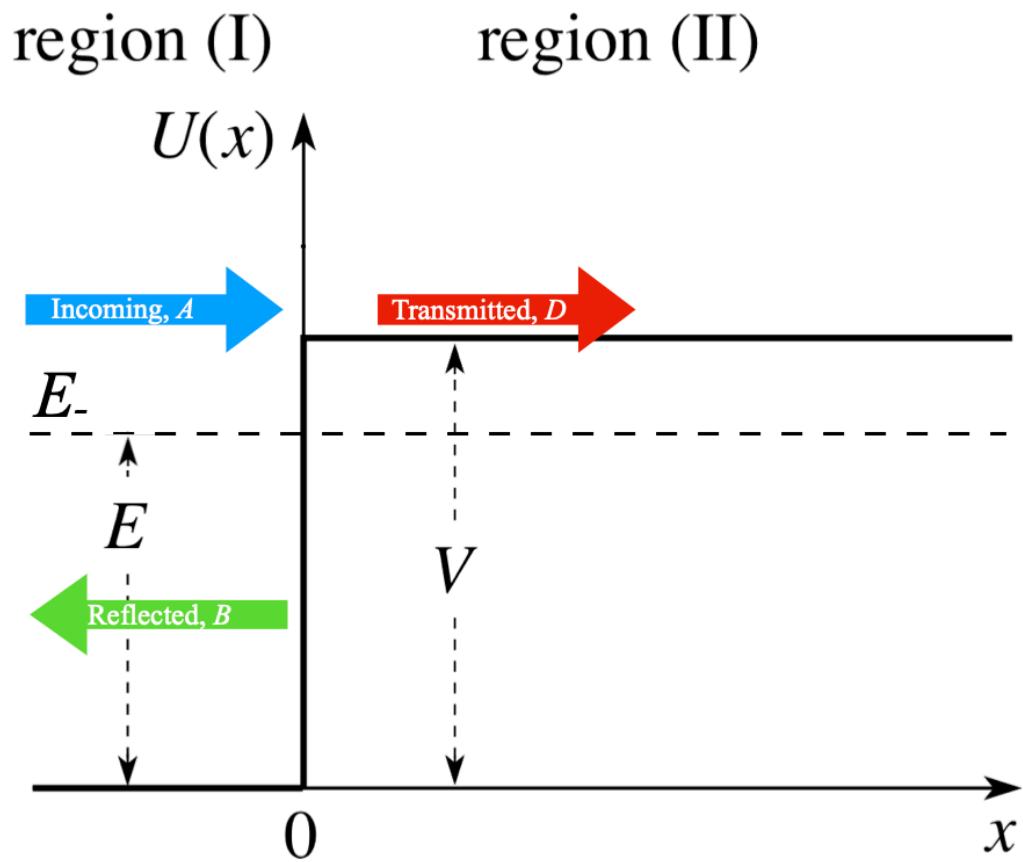


Figure 5.3: Step potential of height V hit by an incident wave of energy E_+ , which gets reflected. The transmitted component is 0 in the sense that there is not a probability flux in region 2, but there is still a probability density in region II due to tunneling.

$$\delta = \frac{1}{2k_2} = \frac{\hbar}{\sqrt{8m(V - E)}}, \quad (5.8)$$

so there is a *static* probability density in region II - there is *no probability current* there.

5.3 Scattering from potential well (double step)

Let's now consider a potential well (or double step), represented in Figure 5.4.

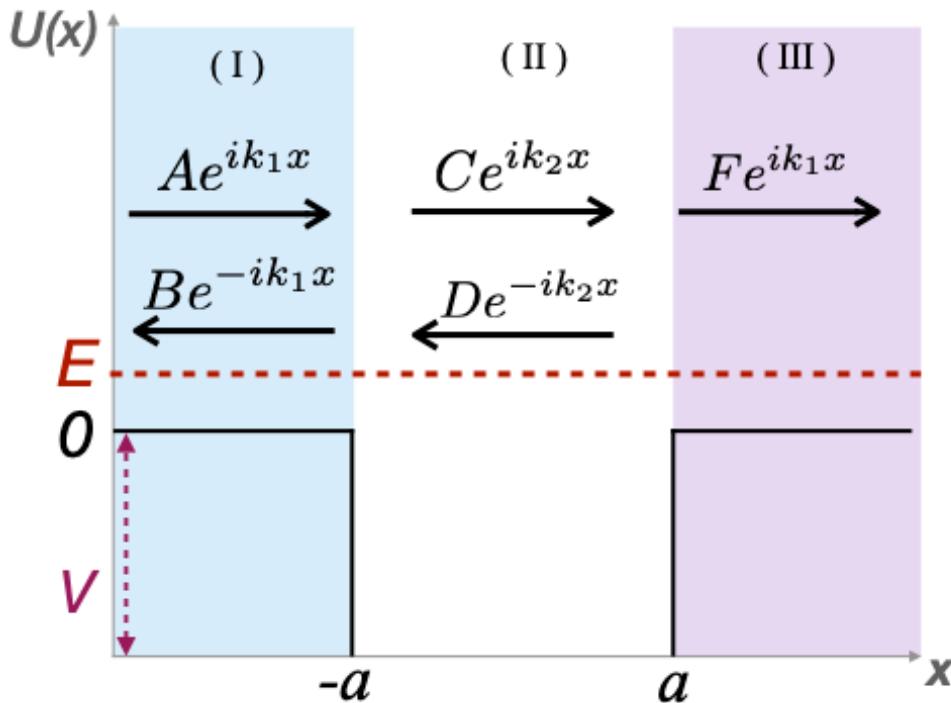


Figure 5.4: Scattering from a finite square well.

This is not very different from the previous case: the principles are the same, just now it's a pain from an algebraic point of view cause we have 5 wave components, i.e. 5 amplitudes to deal with! The wavefunctions in the three regions are:

$$\begin{aligned}\psi_1(x) &= Ae^{ik_1 x} + Be^{-ik_1 x}, \quad \text{with } k_1 = \sqrt{2mE/\hbar^2}, \\ \psi_2(x) &= Ce^{ik_2 x} + De^{-ik_2 x}, \quad \text{with } k_2 = \sqrt{2m(V+E)/\hbar^2}, \\ \psi_3(x) &= Fe^{ik_1 x}.\end{aligned}$$

Note that now k_2 is different from the cases we considered earlier, since the plane wave travels from region 1 at higher potential, i.e. zero, to region 2 at lower (negative) potential, i.e. $-V$. In region 2 you may also use a sum of sin and cos function instead of the exponential: it doesn't matter, the treatment is equivalent. Also note that in region 3 the wavenumber is k_1 as in region 1, since the potential is the same.

As always, we need to set up the BCs in two points: $x = -a$ and $x = a$.

The business for finding the different amplitudes of reflected and transmitted waves is again performed by applying the continuity and smoothness boundary conditions across each potential step.

The algebra gets messy... I'll just go straight to the final key result for A vs F :

$$4k_1 k_2 A = [(k_1 + k_2)^2 e^{-2ik_2 a} - (k_1 - k_2)^2 e^{2ik_2 a}] F e^{2ik_1 a}. \quad (5.9)$$

Such double-steps are officially an algebraic pain in the a...mplitude to solve: not rocket science, just fiddly and heavy on book-keeping, as there are now 5 amplitudes to solve for - check [1] for the derivation if you are really curious - and 4 equations given by the BCs, meaning that there is a free parameter. We can then use the normalisation condition - we can't normalise plane waves, the normalisation comes from $R + T = 1$.

The most interesting take-aways are that the interference of incoming and reflected waves, with the same wavenumber, gives rise to interference patterns in the form of standing waves on the “incident” side of the step. This interference is then seen again within the region between two steps.

5.3.1 Resonant scattering from potential well - Ramsauer effect

Let's consider 2 possible limits: $E \gg |V|$ and $E \ll |V|$ for Equation 5.9 and see how these affect the transmission.

Limit $E \gg |V|$

$E \gg |V| \implies k_1 \approx k_2 \equiv k$, therefore Equation 5.9 returns

$$4k^2 A \approx (2k)^2 e^{-2ika} F e^{2ika},$$

giving the transmission factor

$$T = |e^{2ika} e^{-2ika}|^2 = 1,$$

meaning that if the finite well is “barely visible”, the transmission is perfect.

Limit $E \ll |V|$

Let's consider the limit of Equation 5.9 for $E \ll |V| \implies k_1 \ll k_2$:

$$\begin{aligned} 4k_1 k_2 A &\approx k_2^2 (e^{-2ik_2 a} - e^{2ik_2 a}) F e^{2ik_1 a} \\ &= -2iFk_2^2 \sin(2k_2 a) e^{2ik_1 a}. \end{aligned}$$

Therefore the transmission factor is

$$T = \frac{|F|^2}{|A|^2} = \frac{4k_1^2}{k_2^2 \sin^2(2k_2 a)}. \quad (5.10)$$

Note that the argument in the sin depends on the amplitude of the well, as well as on the depth of the potential, coming from k_2 .

We could plot T vs $E/|V|$ and see that it oscillates with oscillations damped from below for higher energies, always with maxima at 1, like in Figure 5.5.

This is called the **Ramsauer-Townsend effect** and it does not have a classical explanation, it is purely a quantum effect due to the wavelength associated to the wavefunction. *If the incident wavelength is a multiple of the width of the well, then there is perfect transmission,* i.e.

$$\sin(2ka) \rightarrow 0 \implies 2ka = n\pi \implies T = 1. \quad (5.11)$$

The resonant transition corresponds to the infinite-well solutions where half-wavelengths fit exactly into the well, but now with the BCs switched so the amplitudes are *maximal* at the well edges.

5.4 Potential barriers and tunneling ($E < V$)

Let's now consider the case of a classically forbidden barrier with $E < V$, as in Figure 5.6.

This case is similar to the previous one when setting up the wavefunctions, but now in region II we have evanescent waves:

$$\begin{aligned} \psi_1(x) &= Ae^{ik_1 x} + Be^{-ik_1 x}, \quad \text{with } k_1 = \sqrt{2mE/\hbar^2}, \\ \psi_2(x) &= Ce^{k_2 x} + De^{-k_2 x}, \quad \text{with } k_2 = \sqrt{2m(V-E)/\hbar^2}, \\ \psi_3(x) &= Fe^{1k_1 x}. \end{aligned}$$

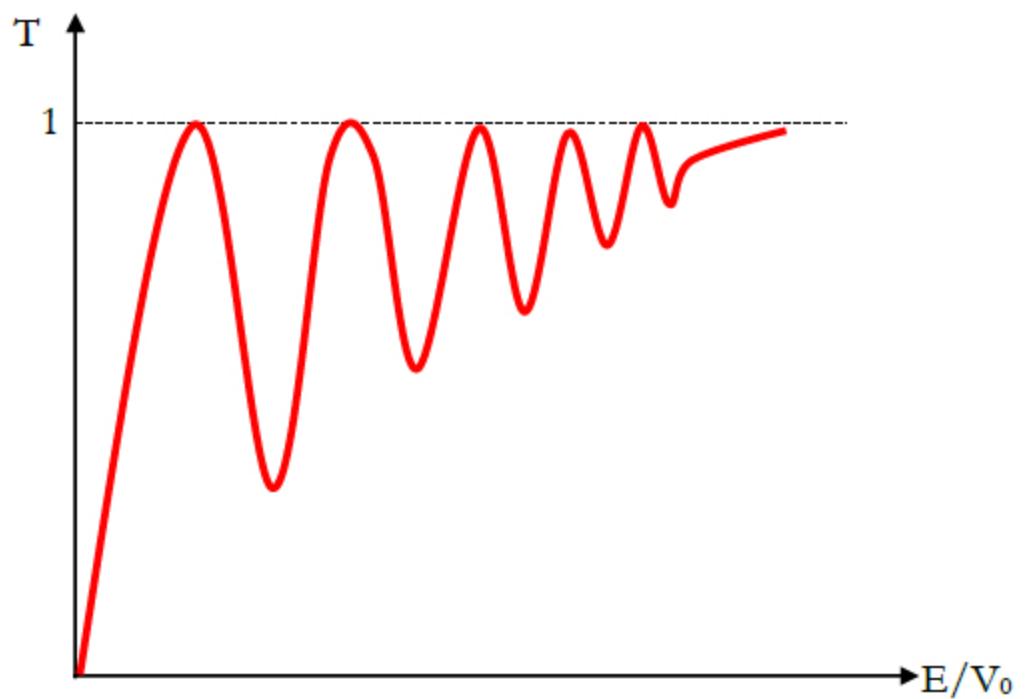


Figure 5.5: Ramsauer-Townsend effect

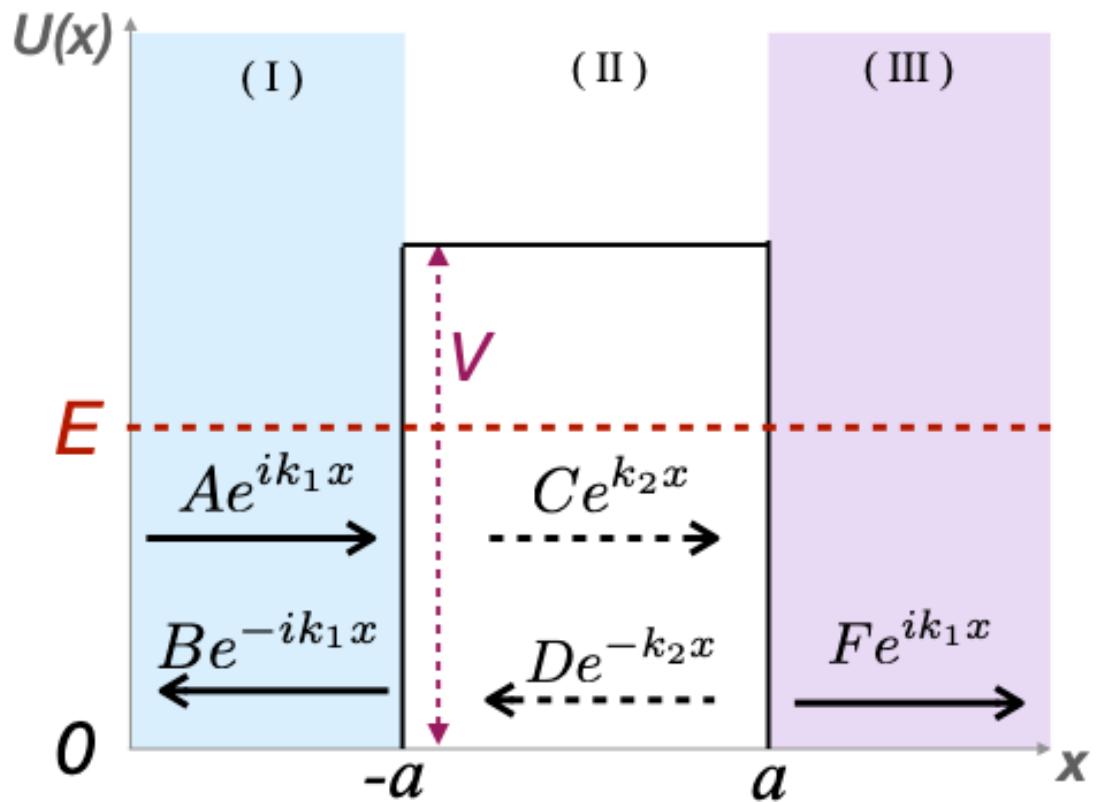


Figure 5.6: “Forbidden” square barrier potential, with $E < V$.

In the case of a classically forbidden barrier, the reflected wave within the barrier also plays a key role, even though they are evanescent rather than propagating waves in that region. The headline result from scattering from a finite-height barrier is that the particle can perform **quantum tunnelling** in which a finite, albeit exponentially suppressed, wavefunction manages to “leak through” the classically forbidden region and resume its oscillatory behaviour on reaching the classically allowed region on the other side.

We might actually have guessed that such a thing was possible from our previous study of bound states: there is no problem with having two oscillatory regions connected by a classically forbidden region, such as in a double-well. But a non-zero wavefunction does not mean zero flux, and we already saw that a single step with decaying exponential $\psi(x) = Ce^{-k_2 x}$ in the forbidden region carries *no flux*, and the reflection is perfect: $R = 1$, $T = 0$.

The procedure to find the amplitudes is still the same: impose continuity conditions for wavefunction and its derivative.

We obtain, for A vs F :

$$4ik_1 k_2 A = [(k_2 + ik_1)^2 e^{-2k_2 a} - (k_2 - ik_1)^2 e^{2k_2 a}] F e^{2ik_1 a},$$

which gives the transmission and reflection factors:

$$\begin{aligned} T &= \left[1 + \frac{(k_2^2 + k_1^2)^2 \sinh^2(2k_2 a)}{4k_2^2 k_1^2} \right]^{-1} \\ R &= 1 - T = \left[1 + \frac{4k_1^2 k_2^2}{(k_1^2 + k_2^2)^2 \sinh^2(2k_2 a)} \right] \end{aligned} \quad (5.12)$$

How can we have flux on the left- and right-hand sides of the barrier, but no flux within?

The answer is that we can't: despite containing “real-valued”, exponentially decaying solutions within the barrier, the reflected (left-going) evanescent wave $Ce^{k_2 x}$ has a different overall complex phase (hidden in C) than its right-going partner $De^{-k_2 x}$. It is hence the phase-interaction of the two decaying waves within the barrier that creates a non-zero flux within the barrier, and the sums of those two evanescent waves on the left and right sides of the barrier that both reduce the reflected amplitude B (so that $R < 1$), and increase the overall transmitted amplitude C above zero ($T = 1$). Reflections are surprisingly subtle things!

The tunnelling phenomenon is of phenomenal importance, both in human affairs via quantum tunnelling technologies such as scanning tunnelling microscopes, and in deeply important natural processes such as nuclear alpha decay and the inverse process of alpha capture in nuclear fusion and stellar nucleosynthesis.

6 Angular momentum theory

6.1 Schrödinger's model of the atom and electron orbitals

In Chapter 1 we introduced briefly the Schrödinger's model of the atom, which introduced the idea of electron orbitals and *quantum numbers* related to those, the angular momentum and the spin. Let's see them again.

6.1.1 Electron orbitals

Electron orbitals introduced by Schrödinger are characterised by quantum numbers. Schrödinger introduced the concept of electron orbitals and *quantum numbers* that describe them:

- **Principal Quantum Number (n):** Indicates the energy level and relative size of the orbital. It can take positive integer values (1, 2, 3, ...).
- **Angular Momentum Quantum Number (l):** Defines the shape of the orbital and can take values from 0 to $n - 1$ for each value of n . Each value of l corresponds to a specific type of orbital (s, p, d, f...).
- **Magnetic Quantum Number (m_l):** Describes the orientation of the orbital in space and can take integer values from $-l$ to $+l$, including 0.
- **Spin Quantum Number (m_s):** Specifies the electron's spin direction, which can be either $+1/2$ or $-1/2$.

Now it's finally time to see more formally what these quantities are and how to deal with them. We will focus on angular momentum, which is the key to describing atomic, molecular and nuclear spectra, and on spin - you will need both concepts for atomic physics next year.

6.2 Orbital angular momentum (OAM)

Angular momentum is the last major topic in this course, and the only one for which working in more than one dimension is essential. This is because angular momentum requires extra dimensions into which a rotation can take place — and for it to be interesting there has to be more than one possible rotational axis. We are, of course, blessed in having three spatial dimensions in reality, so that we can use the usual vector product and define the classical angular momentum:

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ x & y & z \\ p_x & p_y & p_z \end{vmatrix} \Rightarrow \begin{cases} L_x = yp_z - zp_y \\ L_y = zp_x - xp_z \\ L_z = xp_y - yp_x \end{cases} \Rightarrow L_i = \sum_{j,k} \epsilon_{ijk} r_j p_k, \quad (6.1)$$

where ϵ_{ijk} is the **Levi-Civita tensor**, which has cyclic symmetry and it is:

- 0 if any of its ijk indices is repeated;
- 1 if ijk are in the $xyz = 123$ (cyclic) order;
- -1 if ijk is in the $zyx = 321$ or any other anticyclic order.

Figure 6.1 can help you remember this.

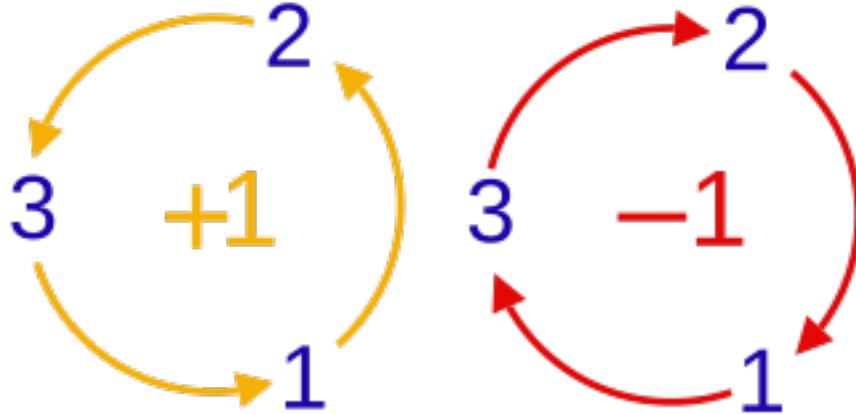


Figure 6.1: A scheme to remember the value of the Levi-Civita tensor.

Replacing $\mathbf{r} \rightarrow (\hat{x}, \hat{y}, \hat{z})$ and $\mathbf{p} = -i\hbar\nabla \rightarrow -i\hbar(\partial_x, \partial_y, \partial_z)$, with $\partial_i = \partial/\partial i$, the angular momentum components of Equation 6.1 can be casted in the form:

$$\hat{L}_x = -i\hbar(y\partial_z - z\partial_y), \quad \hat{L}_y = -i\hbar(z\partial_x - x\partial_z), \quad \hat{L}_z = -i\hbar(x\partial_y - y\partial_x), \quad (6.2)$$

or, using the Levi-Civita tensor:

$$\hat{L}_i = \sum_{jk} \epsilon_{ijk} \hat{r}_j \hat{p}_k = -i\hbar \sum_{jk} \epsilon_{ijk} r_j \partial_k. \quad (6.3)$$

The operators L_i are Hermitian, since components are all of the form $\hat{r}_i \hat{p}_j$, with $i \neq j$, and the canonical commutator is

$$[\hat{r}_i, \hat{p}_j] = i\hbar \delta_{ij}. \quad (6.4)$$

6.2.1 Commutation relations for the orbital angular momentum

Using the canonical commutator of Equation 6.4, we can show the following commutation relations:

$$[\hat{r}_i, \hat{L}_j] = i\hbar \epsilon_{ijk} \hat{r}_k, \quad [\hat{p}_i, \hat{L}_j] = i\hbar \epsilon_{ijk} \hat{p}_k, \quad (6.5)$$

where we implied a sum over k , following the same structure of Equation 6.3.

However, the most important commutation relations are the following.

i Commutation relations for components of angular momentum

$$[\hat{L}_i, \hat{L}_j] = i\hbar \epsilon_{ijk} \hat{L}_k, \quad (6.6)$$

Different components of the angular momentum are *incompatible* eigenstates: they can't have *simultaneously* well-defined L_i projections in different directions.

The ordering sensitivity between angular momentum in any two axial directions leads to a difference that points in the third.

This means that we can't have states that are simultaneously eigenstates of two different components of the angular momentum, e.g. L_x and L_y . The components satisfy an **uncertainty relation**, according to the generalised uncertainty principle (Equation 3.47):

$$\Delta L_i \Delta L_j \geq \frac{\hbar}{2} |\langle L_k \rangle|, \quad \text{with } i \neq j \neq k, \quad (6.7)$$

where the indices indicate the components x, y, z .

The square of the angular momentum however commutes with each component:

$$[\hat{L}_i, \hat{L}^2] = 0 \quad \text{for } \hat{L}^2 = \hat{\mathbf{L}} \cdot \hat{\mathbf{L}} = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2. \quad (6.8)$$

The components of the angular momentum and its square, L_i and L^2 , are *compatible*: we can have *simultaneous* eigenstates of full and projected angular momentum.

This is the same commutativity structure as rotations in 3D. It is no accident that the algebraic structure of angular momentum in 3D and rotations in 3D should coincide - in the exact same way that the Hamiltonian/energy is the generator of time-translations, the linear momenta are generators of spatial translations, and the invariance of physics under those translations leads to conservation of energy and momentum — angular momenta are the generators of angular rotations, and are again conserved as a result in most physical systems.

6.2.1.1 Exercise

Demonstrate Equation 6.8 using the commutation properties.

6.2.2 Angular momentum in spherical polar coordinates

In this section we will derive the description of angular momentum using the wave formalism to solve SEs in 3D. The principles we use are still the same of the solutions of SEs in 1D, but we will write the components in spherical polar coordinates instead of cartesian coordinates, as they better suit the symmetry of the system.

Let's rewrite the operators in polar coordinates:

$$\hat{r} = r\mathbf{u}_r, \quad \nabla = \mathbf{u}_r \frac{\partial}{\partial r} + \mathbf{u}_\theta \frac{1}{r} \frac{\partial}{\partial \theta} + \mathbf{u}_\phi \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi}, \quad (6.9)$$

where \mathbf{u}_i are the unit vectors.

From Equation 6.1, we can write the angular momentum in polar coordinates as:

$$\hat{\mathbf{L}} = -i\hbar \mathbf{r} \times = -i\hbar \left(-\mathbf{u}_\theta \frac{1}{\sin \theta} \frac{\partial}{\partial \phi} + \mathbf{u}_\phi \frac{\partial}{\partial \theta} \right). \quad (6.10)$$

i Note

Note that Equation 6.10 doesn't have radial dependence! L only depends on the angular part of the wavefunction.

6.2.3 Solution of the 3D TISE in spherical polar coordinates

We will now show that angular momentum appears naturally through quantization of the solutions of the TISE in 3D, with spherical symmetry. So let's start with the 3D TISE:

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}). \quad (6.11)$$

Let's replace the Laplacian in spherical coordinates (derived from Equation 6.9) and let's assume an angle-independent potential $V(r)$, i.e. central potential. We obtain the TISE in 3D, in polar coordinates:

$$-\frac{\hbar^2}{2m} \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \Psi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Psi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \Psi}{\partial \phi^2} \right] + V(r)\Psi = E\Psi. \quad (6.12)$$

Let's assume that we can separate the wavefunction in a part for the radial component and one for the angular components: $\Psi(r, \theta, \phi) = R(r)Y(\theta, \phi)$.

Replacing this in Equation 6.12, and multiplying by $-2mr^2/\hbar^2$, we obtain

$$\left[\frac{1}{R} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) - \frac{2mr^2}{\hbar^2} [V(r) - E] \right] = \frac{-1}{Y} \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 Y}{\partial \phi^2} \right] \quad (6.13)$$

We can see that the LHS of the above equation is only dependent on r , while the RHS depends only on the angular coordinates, therefore the two sides have to be equal to the same constant, which we call λ for now.

We saw that \hat{L} depends on the angular components only, so let's focus on the RHS of Equation 6.13 first, and rearranging it we can rewrite it in the form:

$$\sin \theta \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y}{\partial \theta} \right) + \frac{\partial^2 Y}{\partial \phi^2} = -\lambda \sin^2 \theta Y. \quad (6.14)$$

Again, let's try with the separation of variables, $Y(\theta, \phi) \equiv \Theta(\theta)\Phi(\phi)$, so the above equation returns

$$\left[\frac{1}{\Theta} \sin \theta \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Theta}{\partial \theta} \right) + \lambda \sin^2 \theta \right] = -\frac{1}{\Phi} \frac{\partial^2 \Phi}{\partial \phi^2}, \quad (6.15)$$

and we have the same situation seen earlier: the LHS depends only on θ while the RHS depends on ϕ , so they have to be equal to the same constant! Let's call this m^2 , for reasons that will

get more clear later, and let's also impose $\lambda = \ell(\ell + 1)$. Again, the reason for this will be clarified soon.

Let's start with solving the RHS of Equation 6.15, because it's easy:

$$\frac{d^2\Phi}{d\phi^2} = -m^2\Phi \implies \Phi(\phi) = e^{im\phi}, \quad (6.16)$$

where m could be negative because both $e^{im\phi}$ and $e^{-im\phi}$ are good solutions. This is a plane wave solution, and we also observe that $\Phi(\phi) = \Phi(\phi + 2\pi)$, so this means that m must be an integer, i.e. $m = 0, \pm 1, \pm 2, \dots$ is a quantum number.

If you remember, we have seen that geometric periodic boundary conditions give rise to quantization with Bohr's atomic model, and we emphasised this again in Chapter 4, talking about periodic boundary conditions. Now we are seeing something similar here: once again, the boundary conditions give rise to the quantization.

Let's now solve the LHS of Equation 6.15, but as anticipated, let's now replace $\lambda = \ell(\ell + 1)$:

$$\frac{1}{\Theta} \sin \theta \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Theta}{\partial \theta} \right) + [l(l+1) \sin^2 \theta - m^2] = 0 \implies \Theta(\theta) = AP_l^m(\cos \theta), \quad (6.17)$$

where

$$P_l^m(x) = (-1)^m (1-x^2)^{m/2} (d/dx)^m P_\ell(x) \quad (6.18)$$

is the *associated Legendre function* (ALF), and

$$P_\ell(x) \equiv \frac{1}{2^\ell \ell!} \left(\frac{d}{dx} \right)^\ell (x^2 - 1)^\ell \quad (6.19)$$

are the *Legendre polynomials*, which are defined for ℓ integer and positive, i.e. $\ell = 0, 1, 2, \dots$ is the **orbital angular momentum quantum number**. The first six Legendre polynomials are shown in Figure 6.2.

ALFs (Equation 6.18) are not polynomials, but m th derivatives of the Legendre functions, so in order to be non-zero, the condition $|m| < \ell$ must be satisfied, so $m = -\ell, \dots, \ell$ has $2\ell + 1$ possible values.

Spherical harmonics and corresponding quantum numbers

The full angular solution, with normalization factors, are the **spherical harmonics**

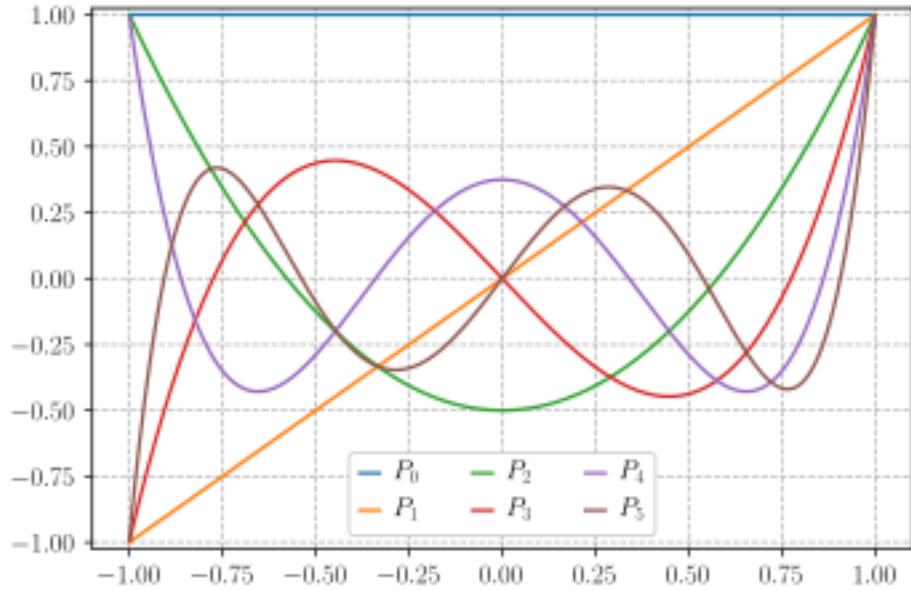


Figure 6.2: Plot of the first six Legendre polynomials.

$$Y_\ell^m(\theta, \phi) = \sqrt{\frac{2\ell+1}{4\pi} \frac{(\ell-m)!}{(\ell+m)!}} e^{im\phi} P_\ell^m(\cos \theta), \quad (6.20)$$

which are the solid-angle equivalent of Fourier series terms, based on the ALF of Equation 6.18 and constitute an orthonormal function basis, i.e.

$$\langle Y_\ell^m | Y_{\ell'}^{m'} \rangle = \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta (Y_\ell^m)^* Y_{\ell'}^{m'} = \delta_{\ell\ell'} \delta_{mm'}. \quad (6.21)$$

The integer ℓ is the **orbital angular momentum quantum number**: its values are $\ell = 0, 1, 2, \dots$, corresponding to the named s, p, d, f, \dots atomic orbitals.

The integer m is the **magnetic quantum number**, representing the quantized z -projections of the angular momentum, and has $2\ell + 1$ possible values: $m = -\ell, -\ell + 1, \dots, \ell - 1, \ell$.

The spherical harmonics are shown in Figure 6.3 and Figure 6.4; ℓ determines the polynomial order, while m effectively controls the orientation.

6.2.4 Action of parity, L and L^2 on spherical harmonics

Let's see what's the action of some operators on the spherical harmonics.

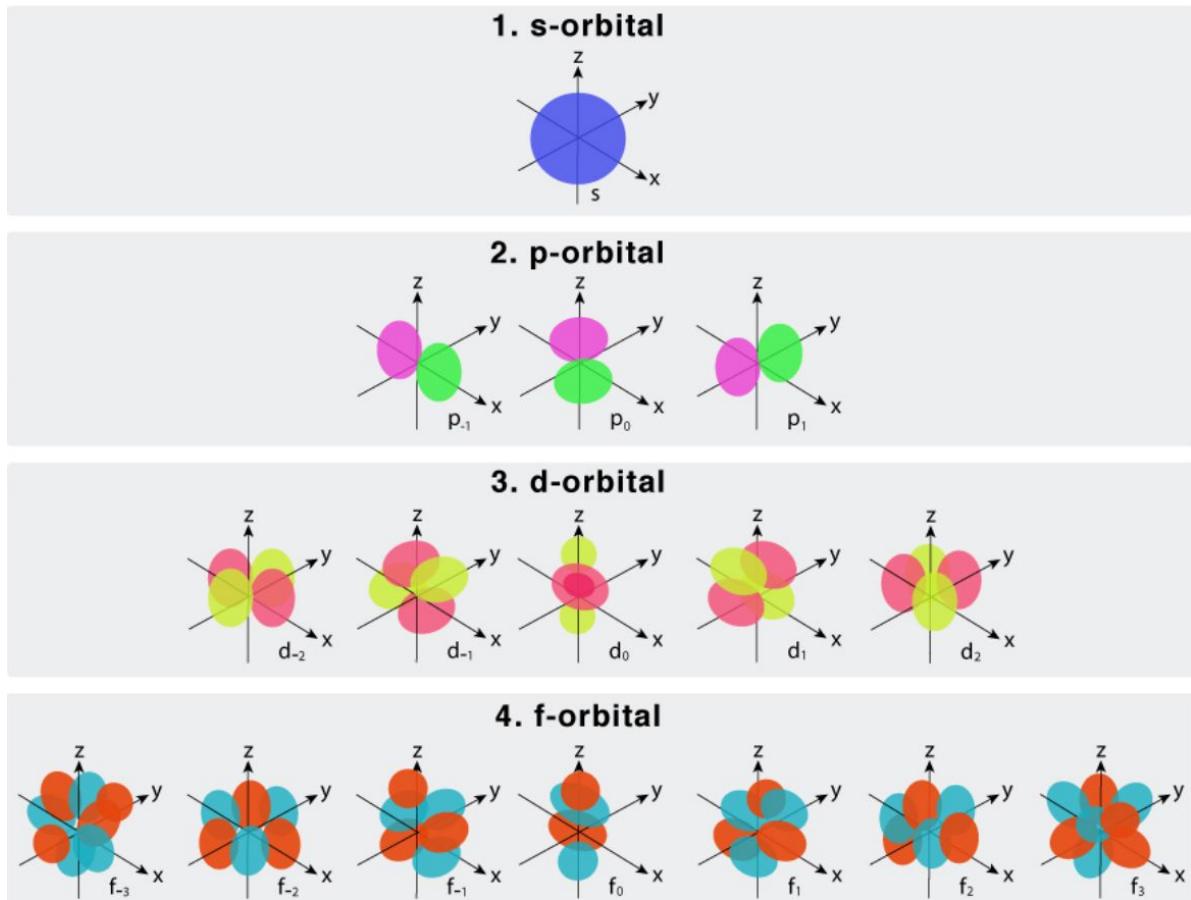


Figure 6.3: Spherical harmonics - atomic orbitals examples.

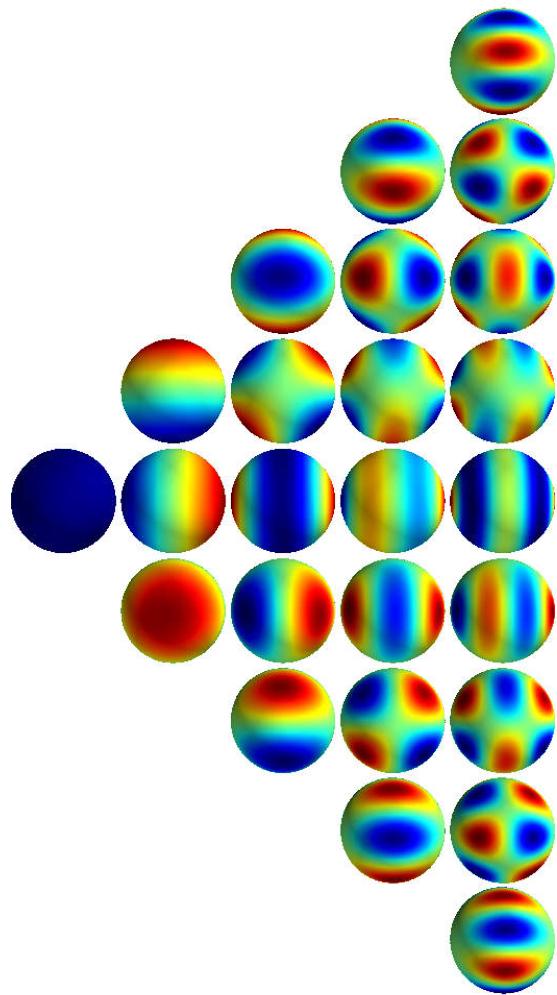


Figure 6.4: Spherical harmonics mapped to the surface of a sphere - m starting from +4 from the top right, going to -4 at the bottom right.

Parity

Under parity transform $\theta \rightarrow \pi - \theta$, $\phi \rightarrow \phi + \pi$:

$$e^{im\phi} \rightarrow e^{im\pi} e^{im\phi} = (-1)^m e^{im\phi}$$

$$P_\ell^m(\cos \theta) \rightarrow P_\ell^m(\cos(\theta + \pi)) = (-1)^{\ell-m} P_\ell^m(\cos \theta)$$

From these, we can find that the total parity is:

$$\hat{P}Y_\ell^m = (-1)^\ell Y_\ell^m. \quad (6.22)$$

z-projection of angular momentum

Under L_z :

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \phi} \implies \hat{L}_z Y_\ell^m = m\hbar Y_\ell^m \quad (6.23)$$

Squared angular momentum

The squared angular momentum is

$$\hat{L}^2 = -\hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right], \quad (6.24)$$

therefore

$$\hat{L}^2 Y_\ell^m = \ell(\ell+1)\hbar^2 Y_\ell^m \quad [\hat{L}^2, \hat{H}] = 0. \quad (6.25)$$

It is important to note that the above are all eigenvalue equations, with Equation 6.23 and Equation 6.25 being particularly important! The eigenvalues are discretised: ℓ and m are quantum numbers, and the fact that $L_z^{\max} = \ell\hbar$ is less than $L^{\max} = \sqrt{\ell(\ell+1)}\hbar$ is related to the uncertainty principle.

Both the total orbital angular momentum and its projection along z (conventionally we selected this direction) are quantised, as shown in Figure 6.5. For a given L_z eigenstate (blue cone in Figure 6.5), the $L_{x,y}$ values live somewhere indeterminate on the rim of the cone. The radius of the sphere, $\sqrt{\ell(\ell+1)}\hbar$ is the total angular momentum and cannot be obtained by the top/bottom projection cones. Note there is nothing magic about the z -axis: quantization of angular momentum applies to whatever axis - x, y, z , or anywhere in-between - but different projection axes are incompatible.

[Here](#) is a link to a nice description of spherical harmonics via the Cosmic Microwave Background power spectrum.

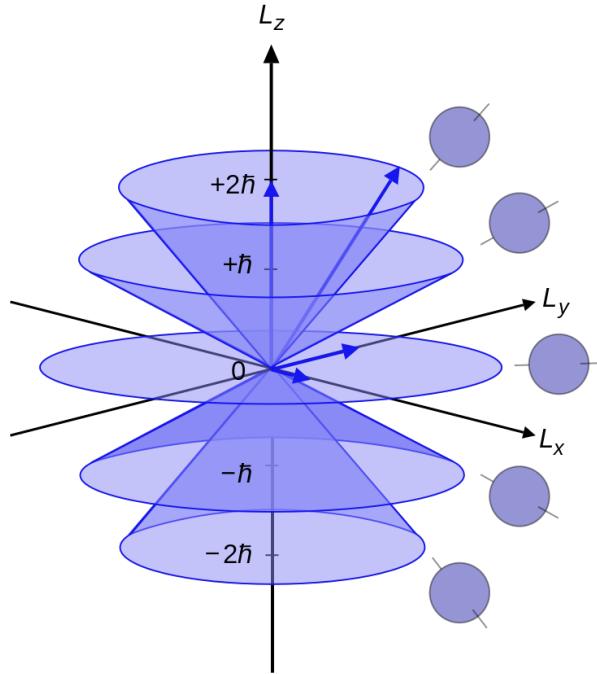


Figure 6.5: Quantization of angular momentum projections.

6.2.5 Orbital angular momentum and spherical harmonics in Dirac notation

So far we focused on L_z and L^2 , but what if we want to determine L_x and L_y ?

It is *very* convenient at this point to swap to the Dirac notation by defining a ket state with the two quantum numbers ℓ, m .

Eigenvalue equations for L_z and L^2

In the Dirac notation, defining $Y_\ell^m \equiv |\ell, m\rangle$, we can rewrite Equation 6.23 and Equation 6.25 as:

$$L_z|\ell, m\rangle = m\hbar|\ell, m\rangle \quad (6.26)$$

$$L^2|\ell, m\rangle = \ell(\ell+1)\hbar^2|\ell, m\rangle \quad (6.27)$$

6.2.6 Angular momentum ladder operators

Just like with QHO energy eigenstates, we can define ladder operators to move between (z) angular momentum eigenstates: these are the *angular momentum ladder operators* \hat{L}_\pm , defined

as follows.

Angular momentum ladder operators

The angular momentum ladder operators are defined as

$$\hat{L}_\pm \equiv \hat{L}_x \pm i\hat{L}_y, \quad (6.28)$$

and have the function of changing m by one unit, without affecting ℓ :

$$\hat{L}_\pm |\ell, m\rangle = \sqrt{\ell(\ell+1) - m(m \pm 1)}\hbar |\ell, m \pm 1\rangle. \quad (6.29)$$

As with the QHO, moves off the ladder (beyond the limits $m = \pm\ell$) return zero:

$$\hat{L}_\pm |\ell, \pm\ell\rangle = 0. \quad (6.30)$$

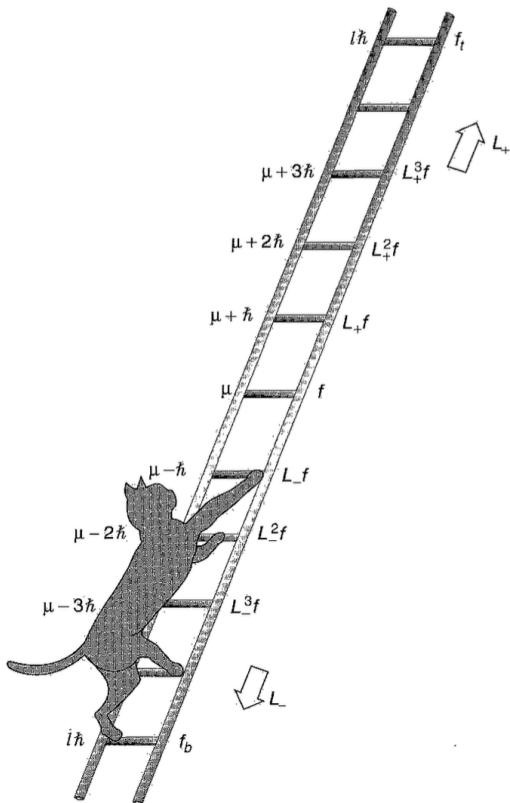


Figure 6.6: Angular momentum ladder operators.

Ladder operators can be very helpful: they allow expressions of $\hat{L}_x, \hat{L}_y, \hat{L}_z^2$ etc. in terms of \hat{L}_z and \hat{L}^2 (cf. ℓ and m eigenstates from Equation 6.26 and Equation 6.27).

But we started this by talking about the z axis and how to measure e.g. L_x — what relevance do operators for changing the z -projection eigenstates have to measuring/extracting the x and y ones? Simply that we can “invert” the ladder operator definitions to give *algebraic* (rather than differential (!), thanks to Dirac) expressions for $L_{x,y}$:

We can rearrange Equation 6.28 to obtain:

$$\hat{L}_x = (\hat{L}_+ + \hat{L}_-)/2, \quad \hat{L}_y = (\hat{L}_+ - \hat{L}_-)/2i, \quad (6.31)$$

From this we can derive the commutation relations

$$[\hat{L}_z, \hat{L}_\pm] = \pm \hbar \hat{L}_\pm, \quad [\hat{L}_+, \hat{L}_-] = 2\hbar \hat{L}_z. \quad (6.32)$$

From Equation 6.31 it is easy to see that:

1. spherical harmonics are not eigenstates of L_x or L_y , as they are composed of single ladder operators which change the state away from the original Y_ℓ^m ;
2. their expectation values must be zero, as the change of ket state forces the overlap integrals to zero.

This latter result also makes physical sense: if we are in a definite- L_z state, the orientation of the x - and y -axes around $\hat{\mathbf{z}}$ is arbitrary and can't pick out one direction as more special than the other: as the projected values can be both positive and negative, with 0 in the middle, by symmetry $\langle L_x \rangle = \langle L_y \rangle = 0$.

This weaponisation of ladder operators and orthogonality to calculate differentially awkward properties can be extended to e.g. the algebraic form of the $\hat{L}_x^2 \equiv \hat{L}_x \hat{L}_x$ operator in terms of the trivially evaluated operators $\hat{L}_z, \hat{L}_z^2, \hat{L}^2$ and \hat{L}_\pm , for which the spherical harmonics are either eigenstates or the targets of Equation 6.29:

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 \implies \hat{L}_x^2 = \hat{L}^2 - \hat{L}_z^2 - \hat{L}_y^2,$$

and using Equation 6.31, this gives

$$\begin{aligned} \hat{L}_x^2 &= \hat{L}^2 - \hat{L}_z^2 - \frac{1}{(2i)^2}(\hat{L}_+ - \hat{L}_-)^2 \\ &= \hat{L}^2 - \hat{L}_z^2 + \frac{1}{4}(\hat{L}_+ \hat{L}_+ + \hat{L}_- \hat{L}_- + \hat{L}_- \hat{L}_+ + \hat{L}_+ \hat{L}_-). \end{aligned} \quad (6.33)$$

The evaluation of this is essentially “reading off” the Y_ℓ^m eigenvalues, with just a little care about the $Y_\ell^m \rightarrow Y_\ell^{m\pm 1}$ state change between the first and second ladder operations in the terms $\hat{L}_- \hat{L}_+$ and $\hat{L}_+ \hat{L}_-$. The $\hat{L}_+ \hat{L}_+$ and $\hat{L}_- \hat{L}_-$ terms contribute nothing to the expectation

value as their overlap-integral terms will be $\langle \ell, m | \ell, m \pm 2 \rangle = 0$, but they do ensure that \hat{L}_z^2 does not share eigenbasis with $\hat{L}_z, \hat{L}_z^2, \hat{L}^2$ - nor with \hat{L}_y etc., for that matter.

6.2.7 Computing angular momentum observables

In Chapter 4, with 1D wavefunctions, we saw that a measurement with an operator \hat{O} randomly collapses the wavefunction into eigenfunctions of \hat{O} . Expectation values $\langle \hat{O} \rangle$ are the mean eigenvalues, averaged over wavefunction collapses.

Like in 1D, this averaging, this averaging is still directly obtained from an integral over spatial configurations — but now we are in 3D, hence need an angular integral over solid angle $d\Omega(\theta, \phi)$:

$$\langle \hat{O} \rangle = \langle \psi | \hat{O} | \psi \rangle = \int d\Omega \psi^* \hat{O} \psi = \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta \psi(\theta, \phi)^* \hat{O} \psi(\theta, \phi). \quad (6.34)$$

6.2.7.1 Some tips

1. Given a general angular wavefunction, first see if you can spot its decomposition into spherical harmonic terms:
 - the c_i coefficients and m, ℓ values are then easily read off
 - expectation calculations simplify to $\langle \hat{O} \rangle = \sum_i |c_i|^2 \lambda_i$ for eigenvalues λ_i
2. Try to decompose general operators into ladder operators, and hence \hat{L}_z and $\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$.
 - These are much easier to work with, because the spherical harmonics are eigenfunctions of these operators.
 - The rewriting is done by expanding e.g. $\hat{L}_{x,y}$ in terms of “backward” linear combinations of \hat{L}_+ , and looking for opportunities to substitute standard commutators (which generate new \hat{L}_i operators), on new axes.
 - Practice! The maths is not so bad, with a bit of practice.

6.2.7.2 Exercise

Evaluate $\langle L_x^2 \rangle$ in terms of ladder operators, and $\langle L_z \rangle$ and $\langle L^2 \rangle$ using the differential operators on spherical harmonics and combinations, e.g. $\psi = \sqrt{5/4\pi} \cos^2 \theta$.

6.3 Spin: intrinsic angular momentum

Besides the orbital angular momentum, which can be (roughly) related to the rotational motion of the electron in its orbit, there is another *intrinsic* angular momentum: the **spin** S , which can't be explained in classical terms, it is an irreducible property of the electron.

The spin operator \hat{S} , its components and its square have the exact same properties and commutator algebra of the orbital angular momentum:

$$[\hat{S}_i, \hat{S}_j] = i\hbar\epsilon_{ijk}\hat{S}_k, \quad [\hat{S}_i, \hat{S}^2] = 0. \quad (6.35)$$

In analogy with the angular momentum states $|\ell, m\rangle$, the spin states are described by the **spin quantum number** s and the **spin magnetic quantum number** m_s , i.e. $|s, m_s\rangle$. To avoid confusion with the quantum number m introduced for the orbital angular momentum, we will call this m_ℓ in the following, i.e. $|\ell, m\rangle \rightarrow |\ell, m_\ell\rangle$.

The physical processes underlying orbital and intrinsic angular momentum are entirely separate, therefore any of the components of one commutes with the components of the other, i.e.

$$[\hat{S}_i, \hat{L}_j] = 0. \quad (6.36)$$

Eigenvalue equations for S_z and S^2 , and ladder operators

Similarly to Equation 6.26 and Equation 6.27:

$$\hat{S}_z|s, m_s\rangle = m_s\hbar|s, m_s\rangle \quad (6.37)$$

$$\hat{S}^2|s, m_s\rangle = s(s+1)\hbar^2|s, m_s\rangle \quad (6.38)$$

We can define ladder operators also for the spin, exactly like for the orbital angular momentum:

$$\hat{S}_\pm = \hat{S}_x \pm i\hat{S}_y, \quad (6.39)$$

which return

$$\hat{S}_\pm|s, m_s\rangle = \hbar\sqrt{s(s+1) - m_s(m_s \pm 1)}|s, m_s \pm 1\rangle. \quad (6.40)$$

There is one important difference between the spin eigenstates and the OAM ones: the spin eigenstates are not spherical harmonics, i.e. they are not functions of θ and ϕ , and there is no a priori reason to exclude half-integer values of s (and therefore of m_s). In fact:

$$s = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots; \quad m_s = -s, -s+1, \dots, s-1, s.$$

Every elementary particle has a specific value of spin, s . Here are some examples:

- pi mesons: $s = 0$;
- electrons: $s = 1/2$;
- photons: $s = 1$;
- delta baryons: $s = 3/2$.

6.4 Total angular momentum

We can define the **total angular momentum**, $\hat{\mathbf{J}}$ (see Figure 6.7), as the sum of the orbital and spin angular momenta:

$$\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}, \quad (6.41)$$

with resulting **total angular momentum quantum number**

$$j = |\ell - s|, \dots, \ell + s,$$

i.e. j takes all the integer values allowed between completely antiparallel and completely aligned sources of angular momentum. Each j then has the usual set of $2j+1$ **quantized projection states**, $m_j = m_\ell + m_s \in \{-j, \dots, j\}$.

The components of the total angular momentum satisfy the commutation relation

$$[\hat{J}_i, \hat{J}_j] = i\hbar\epsilon_{ijk}\hat{J}_k. \quad (6.42)$$

Since $[L, S] = 0$, \hat{L}^2 and \hat{S}^2 are both compatible with \hat{J}^2 :

$$\hat{J}^2 = (\hat{\mathbf{L}} + \hat{\mathbf{S}})^2 = \hat{L}^2 + \hat{S}^2 + 2\hat{\mathbf{L}} \cdot \hat{\mathbf{S}},$$

therefore

$$[\hat{J}^2, \hat{L}^2] = [\hat{J}^2, \hat{S}^2] = 0.$$

However, \hat{L}_i and \hat{S}_i are incompatible with \hat{J}^2 :

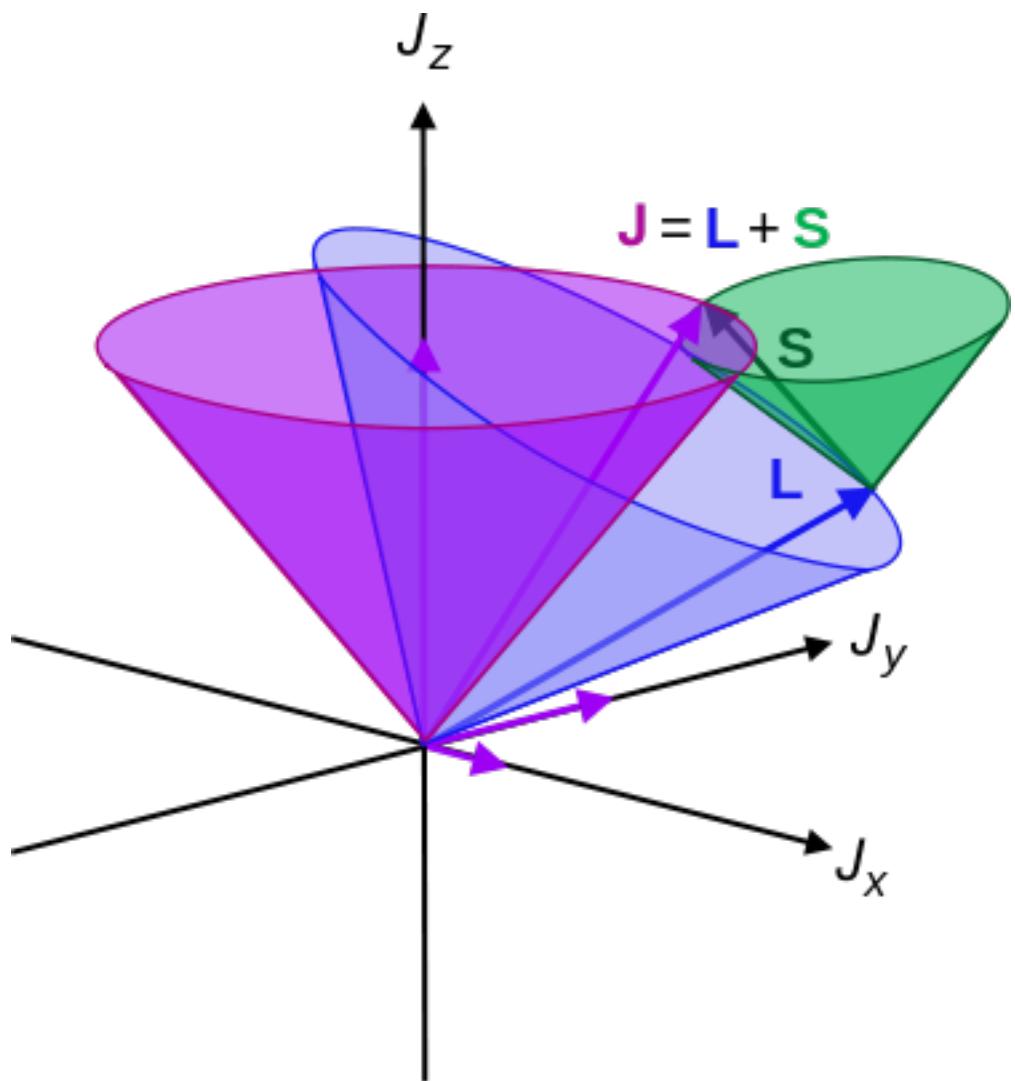


Figure 6.7: Total angular momentum.

$$[\hat{J}^2, \hat{L}_z] = 2[\hat{L}_x, \hat{L}_z]\hat{S}_x + 2[\hat{L}_y, \hat{L}_z]\hat{S}_y \neq 0.$$

The eigenvalue equation for the total angular momentum is

$$\hat{J}^2|j\rangle = j(j+1)\hbar^2|j\rangle, \quad (6.43)$$

where $|j\rangle$ is the state having $|j = \ell + s\rangle$.

This topic can get much broader, and we could be talking about how to measure the spin and angular momentum, but this is something that you will see in the next year's courses, especially in Atomic Physics, so this seems like a good stopping point.

6.5 Summary

- Angular momentum requires 3D wavefunction treatment. Wavefunction in a central potential is separable into r, θ , and ϕ dependences: angular part given by orthogonal spherical harmonics.
- L^2 quantum number is $\ell(\ell+1)$; \hat{L}_z quantum number is m_ℓ .
- Commutation relations: can't simultaneously measure $L_{x,y,z}$, but can measure one projection plus L^2 . Projection always smaller than full to preserve uncertainty.
- Ladder operators also exist for angular momentum: raise and lower m_ℓ for fixed ℓ . Can use with orthogonality in expectation value calculations
- The spin S is an intrinsic angular momentum independent of L , but obeying the same commutation relations and properties, except for the fact that its quantum number s can be *half-integer*.
- Combine L and S into total momentum J .

7 Summary of exercises

This will be updated as we progress with the lectures.

7.1 Chapter 1

7.2 Chapter 2

7.3 Chapter 3

7.4 Chapter 4

7.5 Chapter 5

8 FAQs

Bibliography

- [1] Brian Harold Bransden and Charles Jean Joachain. “Introduction to quantum mechanics”. In: (1989).
- [2] Richard Phillips Feynman, Robert Benjamin Leighton, and Matthew Sands. *The Feynman lectures on physics; New millennium ed.* Originally published 1963-1965. New York, NY: Basic Books, 2010. URL: <https://cds.cern.ch/record/1494701>.
- [3] David J Griffiths and Darrell F Schroeter. *Introduction to quantum mechanics*. Cambridge university press, 2019.
- [4] A.E. Haas. “Über die elektrodynamische Bedeutung des Planck’schen Strahlungsgesetzes und über eine neue Bestimmung des elektrischen Elementarquants und der Dimension des Wasserstoffatoms”. In: Sitzungsberichte der kaiserlichen Akademie der Wissenschaften in Wien. 2a, 119 pp. 119-144 (1910).
- [5] F.K. Richtmyer, E.H. Kennard, and J.N. Cooper. *Introduction to Modern Physics*. International series in pure and applied physics. McGraw-Hill, 1969. URL: <https://books.google.co.uk/books?id=e9JGswEACAAJ>.