5CCM234A/6CCM332A Introductory Quantum Theory Lecture Notes

P. Kravchuk, N. Lambert, G. Watts, P.P. Cook and others.

Lectured by Dr. Paul P. Cook

12th May 2025

"Those who are not shocked when they first come across quantum theory cannot possibly have understood it"

Niels Bohr

"I don't like it, and I'm sorry I ever had anything to do with it."

Erwin Schrödinger

"I think I can safely say that nobody understands quantum mechanics."

Richard Feynman

0. Contents

1 The Old Quantum Theory **7**

1.1 Planck's formula 7
1.2 The photoelectric effect 11
1.3 Compton scattering 12
1.4 So what is light? 13
1.5 Rutherford's and Bohr's atoms 14
1.6 The correspondence principle 16
1.7 de Broglie waves 17

2 The Schrödinger Equation and Wave-Functions **19**

2.1 An Equation for Matter Waves 19
2.2 Interpreting the wave-function 21
2.3 The Free Particle 24
2.4 Uncertainty and wave-function spreading 26
2.5 Operators and Expectation Values 27
2.6 The Hamiltonian Operator and the Schrödinger equation 29
2.7 Conservation of Probability 30
2.8 Evolution of expectation values and the Ehrenfest theorem 32

3 Simple Quantum-Mechanical Systems **37**

3.1 Separation of Variables and Energy Levels 37
3.2 The Infinite Potential Well 38
3.3 The Free Particle Again 42
3.4 Bound States and Scattering States 47
3.5 The Delta-Function Potential Well 48
3.5.1 Bound states in the delta-function well 49
3.5.2 Scattering states of the delta-function well 51
3.6 The Finite Square Potential Well 55

3.6.1 Bound states in a finite square potential well . . . 56
3.6.2 Scattering States in a finite square potential well . . . 60
3.7 The Potential Barrier . . . 66

4 A Mathematical Framework for Quantum Mechanics 73
4.1 Hilbert spaces . . . 74
4.2 Observables . . . 79
4.2.1 Hermitian operators and the adjoint operator . . . 80
4.2.2 Eigenfunction of Hermitian Operators . . . 81
4.3 Dirac's Bra-Ket Notation . . . 85
4.3.1 Matrix elements in bra-ket notation . . . 88

5 The Statistical Interpretation of Quantum Mechanics 93
5.1 Examples of measurement for continuous spectra . . . 95
5.1.1 Measuring position . . . 95
5.1.2 Measuring momentum . . . 97
5.2 Simultaneous Measurement and Commutativity . . . 98
5.3 The Generalised Uncertainty Principle . . . 99
5.3.1 The Energy Time Uncertainty Relation . . . 102
5.4 The Time Evolution Operator . . . 104
5.5 The Heisenberg Picture . . . 108
5.5.1 Equations of Motion In The Heisenberg Picture . . . 110

6 The Harmonic Oscillator 113
6.1 The Classical Solution . . . 113
6.2 The Quantum Solution . . . 114
6.2.1 The Algebraic Approach . . . 114
6.2.2 The Analytical Approach . . . 120

7 Further Ideas 125
7.1 Tensor Products . . . 125
7.2 Spin States . . . 128
7.2.1 Tensor Products of Spin States . . . 129
7.3 Quantum Teleportation . . . 130
7.4 The No-Cloning Theorem . . . 133
7.5 The EPR Thought Experiment and Bell's Inequality . . . 134

Appendices

A Waves and Fourier Transforms **143**
A.1 Travelling Waves and Wave Equations 143
A.2 The Fourier Transform 146

B Tutorial Exercises **151**

Acknowledgements

These lecture notes were based on the book 'Lectures on Quantum Mechanics' by Steven Weinberg. I inherited the set of notes written by Petr Kravchuk, which themselves (to the best of my knowledge) were based on the notes of Neil Lambert. Gérard Watts updated and edited the notes and I have made minor edits as well. The problem sets have been written and passed down by the authors mentioned above. All typographical errors and errors of any other type are, of course, my fault and I will be grateful for all comments on any kind of mistake. In that regard, I am very grateful to Sangavai Punnyaseelan for telling me about a typo in the notes, which has been corrected in this version of the lecture notes.

Paul P. Cook, 28$^{\text{th}}$ October, 2024.

1. The Old Quantum Theory

1.1 Planck's formula

In the late 19$^{\text{th}}$ century the basis of the theoretical understanding of our world was formed by the "classical theory": Newton's laws (or more sophisticated formulations such as Lagrangian or Hamiltonian mechanics) and Maxwell's theory of the electromagnetic field. However, by 1900 and shortly after that there was enough experimental evidence that made it clear that the classical theory does not completely describe the nature.

Historically the first problems appeared when physicists tried to describe the spectrum of radiation emitted by heated objects. It was observed that any physical body at non-zero temperature[1] emits electromagnetic radiation with a continuous spectrum of wavelengths/frequencies.

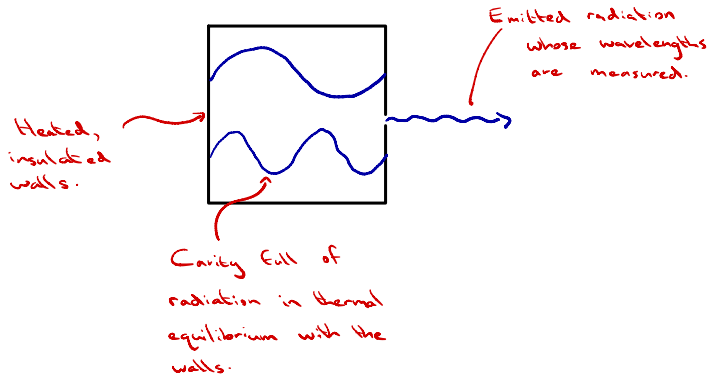


Figure 1.1: A sketch of a blackbody: a model for producing electromagnetic waves which are in thermal equilibrium with a body at temperature T. The frequencies in the emitted radiation can then be investigated.

[1]Non-zero here means "above the absolute zero", which is at zero degrees Kelvin.

Not all frequencies of radiation are produced equally, instead certain frequencies dominate. This is the same phenomenon that makes iron objects glow red when they are heated. We are interested in the answer to the following question: Consider some physical object at a temperature T, it loses heat through electromagnetic radiation (vibrating electrons lead to electromagnetic waves), into which frequencies is the energy it emits distributed? The function we are seeking is called the spectral energy density[2], and denoted $u(\nu, T)$ and it tells us about the energy contained in frequencies between ν and $\nu + d\nu$. In fact

$$u(\nu, T) = N(\nu)\langle E\rangle$$

where $N(\nu) = \frac{8\pi\nu^2}{c^3}$ is the number of standing waves in a cavity with frequencies between ν and $\nu + d\nu$, and $\langle E\rangle$ is the average energy of the oscillators in the wall of the cavity at temperature T. There exists a simple argument which implies that if our object is a black body, i.e. it absorbs all incident radiation, then u is independent of the physical nature of the body and only depends on thermodynamic properties of electromagnetic radiation. Indeed, if an object is in equilibrium with radiation, it should emit exactly as much radiation as it absorbs, and should be at the same temperature T as the electromagnetic field. But a black body absorbs all radiation that reaches it, whose properties are completely determined by the electromagnetic field.

Lord Rayleigh and James Jeans had used statistical mechanics[3] to determine $\langle E\rangle$, the ideas that worked so successfully for thermodynamics said that energy was distributed according to the Boltzmann distribution, so that the expectation of the energy in an oscillating wave was

$$\langle E\rangle = \frac{\int_0^\infty E e^{(-E/k_B T)} dE}{\int_0^\infty e^{-E/k_B T} dE} = k_B T$$

where $k_B = 1.380649 \times 10^{-23}$ Joules per Kelvin is called the Boltzmann constant. The classical computation of Jeans and Rayleigh stated that the expected energy was independent of the frequency of the electromagnetic waves carrying the energy and gave

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

which became known as the Rayleigh-Jeans law. Unfortunately the Rayleigh-Jeans law only agreed with experimental results for small frequencies. In figure 1.2 the experimental results

[2] An equivalent quantity is called the radiance $B(\nu, T)$ and it is simply related to the spectral energy density by $B(\nu, T) = \frac{c}{4\pi} u(\nu, T)$.

[3] The details of the classical computation uses Maxwell's equations and ideas from statistical physics for a black body, which we have only outlined here. For more information see, for example, Lectures on Quantum Mechanics by Steven Weinberg or Quantum Mechanics - Concepts and Applications by Nouredine Zettili, but even so a firm appreciation of statistical physics is recommended to appreciate the full computation.

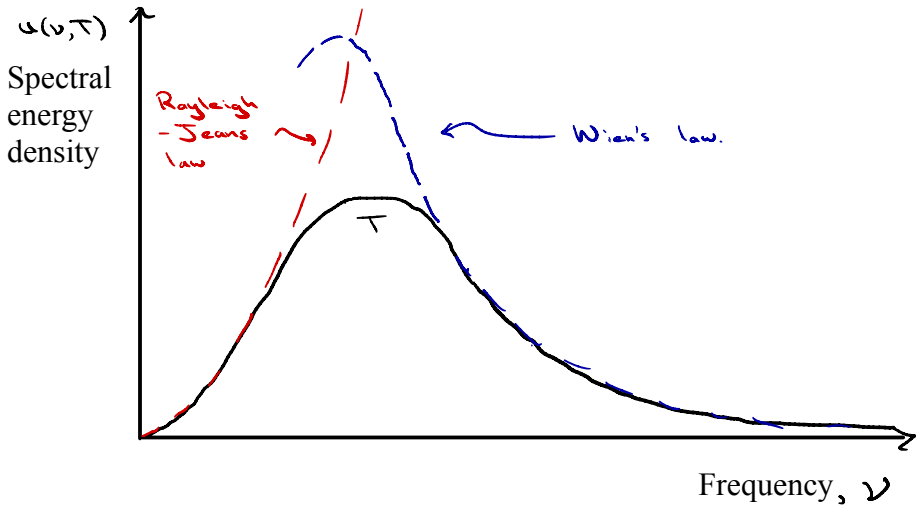


Figure 1.2: Comparison of Planck, Rayleigh-Jeans and the Wien laws for the blackbody radiation at a fixed temperature T. The solid curve shows the results from experiment, which agrees with the prediction of Planck.

are sketched with a solid line, and the Rayleigh-Jeans law is shown in the dashed line to the left of the sketch. There was a second, different formula for $u(\nu, T)$ due to Wien, who derived his equation from the Stefan-Boltzmann law that power radiated is proportional to T^4. This was also on solid ground in classical physics, and Wien's law was

$$u(\nu, T) = A\nu^3 e^{-\beta\nu/T}$$

where A and β were parameters that could be adjusted to fit the experimental data. Wien's law agreed with the experimental results only for large ν (Wien's law is sketched on figure 1.2 in the dashed line mostly to the right of the figure). So there were two well-founded classical predictions that failed to give the results found in experiments. The Rayleigh-Jeans law in fact predicted that infinite energy is radiated from a blackbody at finite temperature T (found by integrating $u(\nu, T)$ over all the frequencies to find the total energy). This was called the ultraviolet catastrophe (because u diverged for large ν).

One feature that was found in the experiments is that for a given temperature T the spectral energy density $u(\nu, T)$ has a maximum at some ν. In other words, there is some characteristic frequency $\nu(T)$ or equivalently a wavelength $\lambda(T)$ at which most of the energy is radiated. If we were to apply the classic argument using statistical physics and Maxwell's equations, we would expect to find a theoretical prediction for $\lambda(T)$ in terms of T, the speed of light c, and the Boltzmann constant k_B. But no such expression exists: it is impossible to form an expression

with units of length using these constants! Consider

$$\lambda(T) \propto T^{\alpha} c^{\beta} k_B^{\gamma}. \tag{1.1}$$

Note that T has units K (degrees Kelvin), c has units m/s, and k_B has units of energy divided by temperature, $kg \cdot m^2 \cdot s^{-2} \cdot K^{-1}$. Altogether, the right-hand side above has the units of $K^{\alpha}(m/s)^{\beta}(kg \cdot m^2 \cdot s^{-2} \cdot K^{-1})^{\gamma}$. The units of λ should be metres m. These do not match for any choice of α, β, γ. (Verify this!)

In the problem set you will show, using the same idea of dimensional analysis, that the classical argument must give the Rayleigh-Jeans law:

$$u(\nu, T) = \frac{N\nu^2}{c^3} k_B T \tag{1.2}$$

for some numerical constant N. The complete argument shows that N turns out to be 8π.

The breakthrough came in 1900 when Max Planck proposed a new formula for $u(\nu, T)$ that described the experimental data rather well. How did he manage that? Well, one thing that is obvious from the above discussion is that we need to add some new physical constants to our arsenal if we are to get correct answers that have correct units. This is how Planck's constant h came to be. Introducing it breaks the dimensional analysis argument that we made above, and infinitely many new formulas can be written. Planck's argument rested upon the idea that the energy exchanged between the vibrating particles in the wall of the blackbody and the electromagnetic radiation were not exchanged continuously, instead he said that at a frequency ν only the energies

$$E = nh\nu$$

for $n \in \{0, 1, 2, 3, \ldots\}$ could be exchanged. Then the computation of $\langle E \rangle$ from statistical mechanics becomes

$$\langle E \rangle = \frac{\sum_{n=0}^{\infty} nh\nu e^{-nh\nu/k_B T}}{\sum_{n=0}^{\infty} e^{-nh\nu/k_B T}} \tag{1.3}$$

$$= \frac{h\nu}{e^{h\nu/k_B T} - 1} \tag{1.4}$$

and so Planck predicted the formula (using $u = N(\nu)\langle E \rangle$)

$$u(\nu, T) = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{(h\nu/k_B T)} - 1}. \tag{1.5}$$

Planck's formula matched the experimental data, and for large ν it becomes $\frac{8\pi h\nu^3}{c^3} e^{-\frac{h\nu}{k_B T}}$ i.e. Wien's law, while for small ν it becomes $\frac{8\pi\nu^2}{c^3}$ i.e. the Rayleigh-Jeans law (use $e^{\frac{h\nu}{k_B T}} - 1 \approx \frac{h\nu}{kT}$).

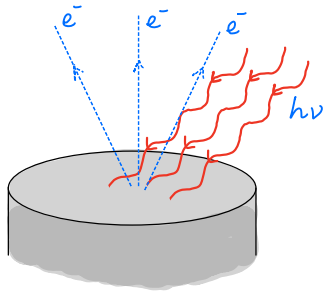


Figure 1.3: A cartoon of electrons being emitted from a metal sample irradiated by light.

The reason that the Rayleigh-Jeans formula works so well for small ν is that for low frequency energy is exchanged in small lumps, which approaches the continuum distribution of E.

Planck was motivated by finding a formula to fit the experimental data, but key to explaining his idea was that electromagnetic energy could only be absorbed or emitted in portions, which he called quanta (singular "a quantum"), of energy $E = h\nu$ where $\nu = c/\lambda$ is the frequency of radiation. In other words, we only care that the energy is exchanged in discrete portions (quanta) if these portions are noticeably large. Large compared to what? The only energy scale we have is the scale of thermal fluctuations $k_B T$. So if $h\nu \ll k_B T$ then the size of the portions should be irrelevant and Planck's formula should reduce back to the classical one. Indeed, it does.

This discovery showed that there are some things that physicists definitely did not understand about electromagnetic radiation. As we will see, the quantum $E = h\nu$ and the Planck's constant h ignited the development of the quantum theory.

1.2 The photoelectric effect

The next development came with Einstein's explanation of the photoelectric effect. It was observed that when a piece of metal is illuminated by light with the right properties, its charge steadily changes. This was interpreted as electrons being emitted from the metal by the electromagnetic forces (figure 1.3).

Recall that according to the classical theory light is an electromagnetic wave, i.e. a wave of electric and magnetic fields. An electric field produces a force on charged particles and can transmit energy to them. It is then natural that electrons can be accelerated by light and eventually ejected from the metal. This phenomenon was first investigated by Hertz in 1887. The amplitude squared of the electric field is proportional to the intensity of the light source I,

and so the classical theory predicts that the kinetic energy K of the emitted electrons should grow with I.

Instead, it was found that K does not depend on the intensity I and depends instead on the frequency of light ν. In fact, for ν below some threshold value ν_0, K goes to zero and there is no electron emission no matter the intensity. When the emission does happen, only the number of emitted electrons depended on intensity. Even for very low intensity light if the frequency of the light satisfied $\nu > \nu_0$ then electrons were emitted. If $\nu > \nu_0$ then the number of electrons ejected increased with the intensity but that number did not depend on ν.

In order to explain these results, in 1905 Einstein proposed that a stronger form of Planck's idea holds: it is not just that the electromagnetic field exchanges energy with matter in discrete portions, but light in fact consists of particles, photons, carrying the energy $h\nu$. When a photon is absorbed by an electron, the latter receives energy up to $h\nu$.[4] Some fixed work W is spent on removing the electron from its binding to the metal, and the maximal kinetic energy of emitted electrons becomes

$$K = h\nu - W. \tag{1.6}$$

Of course, if the frequency ν is too small and $h\nu < W$, then there is not enough energy to remove the electron from the metal, and no effect is observed. This predicts $\nu_0 = W/h$. The intensity I of the light source is proportional to the number of photons, which explains why the number of emitted electrons grows with I, but not their energies. The formula (1.6) was later confirmed experimentally by Millikan. Einstein's Nobel Prize was awarded for his work on photoelectric effect.

1.3 Compton scattering

An effect that is somewhat similar in nature to photoelectric effect is Compton scattering. Consider a free electron that interacts with an incoming electromagnetic wave of frequency ν. According to the classical theory, the electron will be accelerated by the wave, and will start to move in some manner. The motion will happen at the same characteristic frequency ν, and the electron would thus produce electromagnetic radiation with frequency $\nu' = \nu$. There could be corrections to ν' due to the velocity of the electron, but these would depend on the amplitude of the wave.

Instead Arthur Compton observed in an experiment in 1923 that the wavelength $\lambda' = c/\nu'$ of the radiation emitted by the electrons was related to that of the incident radiation by

$$\lambda' - \lambda = \lambda_e(1 - \cos\theta), \tag{1.7}$$

[4]Possibly less because the electron can re-emit a photon, or interact with the metal.

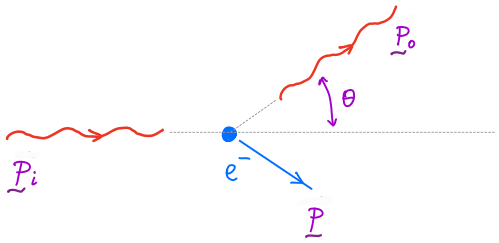


Figure 1.4: Kinematics of the Compton effect. The incoming radiation is coming from the left, with the wavelength λ, and the radiation scattered at angle θ has the wavelength λ'.

where θ is the angle at which the radiation is measured (see figure 1.4) and λ_e is the Compton wavelength of the electron,

$$\lambda_e = \frac{h}{m_e c}. \tag{1.8}$$

As you will see in the problem set, this formula is easily explained by Einstein's hypothesis that light consists of photons. The key idea is that a photon can scatter off an electron, transferring a part of the energy $h\nu$ to it. The scattered photon will then have a lower energy $h\nu'$, with $\nu' < \nu$ and thus $\lambda' > \lambda$. Working out the conservation of energy and momentum using relativistic formulas leads to (1.7).

1.4 So what is light?

The above examples seem to prove conclusively that light consists of photons carrying energies $E = h\nu$. But, of course, previously physicists came to the conclusion that light is an electromagnetic wave, and they had equally convincing evidence based on interference and diffraction. The question then arises: what, in the end, is light? Is it a particle or a wave?

The answer is of course that it is neither: as we have seen, either concept fails in explaining one or another experiment. So we need a new concept that would have the properties of both particles and waves, and a mathematical formalism to describe this concept. The new concept is that of a quantum particle. The mathematical formalism is going to be the subject of this module.

However, I have to make the disclaimer that we will not develop the mathematical formalism that describes photons. The reason for this is that we will study *non-relativistic Quantum Mechanics*. It is not applicable to photons, which move at the speed of light and are inherently relativistic. The relativistic version of Quantum Mechanics is Quantum Field Theory, which is far beyond the scope of this module.

Quantum Mechanics is however applicable to non-relativistic electrons and other particles,

and these will serve as the experimental context for our lectures. So let's now discuss the other developments that grew out of the ideas discussed in the previous sections and led the physicists to reconsider the nature of electrons, nuclei, and other particles.

1.5 Rutherford's and Bohr's atoms

In 1911-1913 Ernest Rutherford developed a model of the atom, based on experiments performed under his direction by Hans Geiger and Ernest Marsden. The experiments were based on scattering of alpha-particles (helium nuclei) off a thin metal foil, and showed that sometimes the alpha-particles scatter at large angles. This was inconsistent with the previous models of the atom, and supported Rutherford's idea of a small yet heavy positively-charged nucleus (which could significantly affect the trajectory of the heavy-ish alpha-particles) surrounded by relatively light negatively-charged electrons.

Despite being supported by experimental evidence, Rutherford's atom was not theoretically-self consistent. Consider for example the Hydrogen atom, which has one electron. According to the classical theory, the electron would orbit the atom in circular or elliptical orbits, just as planets orbit the sun. However, circular orbits mean that the electron moves with acceleration, and accelerating charges emit electromagnetic radiation, which carries away energy. The electron would then move to orbits with smaller radius and higher orbital frequency, and the radiation wavelength would shorten. Eventually the electron would fall onto the nucleus.

This means that in the classical theory, Rutherford's atom is inherently unstable which is of course not what is found experimentally. Electromagnetic radiation by atoms has been observed, but it behaves completely differently from what Rutherford's model predicts.

Specifically, it was observed that the light is emitted by atoms only at a discrete set of wavelengths. For Hydrogen, generalising earlier Balmer's formula, Johannes Rydberg found that these wavelengths can be empirically described by the formula

$$\frac{1}{\lambda_{n_1,n_2}} = R\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right) \tag{1.9}$$

where R is the Rydberg constant, and $n_1 > n_2$ are any pair of positive integers.

Niels Bohr found a way to explain this spectrum and compute R in terms of fundamental constants. His idea was that classically, when an electron orbits the nucleus, the frequency at which it radiates is determined by the period T of its orbit, $\nu = 1/T$. Since we know from other experiments that radiation can only be emitted in quanta with energies $E = h\nu$, this suggests that only a discrete set of frequencies should be allowed. Bohr effectively postulated that the angular momentum L of the electron is quantised. For a circular orbit,

$$L = m_e r v = \hbar n, \tag{1.10}$$

where $\hbar$ is some new constant, m_e is electron's mass, r is the radius of the orbit, v is electron's velocity and n is a positive integer. On the other hand, we have that the Coulomb's force must give the correct centripetal acceleration

$$\frac{e^2}{4\pi\epsilon_0 r^2} = \frac{m_e v^2}{r}. \tag{1.11}$$

We can find from (1.10) and (1.11)

$$v = \frac{1}{4\pi\epsilon_0}\frac{e^2}{n\hbar}, \tag{1.12}$$

$$r = 4\pi\epsilon_0 \frac{n^2\hbar^2}{m_e e^2}, \tag{1.13}$$

and the electron's total energy

$$E_n = \frac{m_e v^2}{2} - \frac{1}{4\pi\epsilon_0}\frac{e^2}{r} = -\frac{1}{2}\left(\frac{1}{4\pi\epsilon_0}\right)^2 \frac{m_e e^4}{n^2\hbar^2}. \tag{1.14}$$

This formula then explains the spectrum of wavelengths (1.9): if the electron is excited to the orbit with $n = n_1$, it can then emit a photon with energy $E_{n_1,n_2} = E_{n_1} - E_{n_2}$ if it transitions to a lower-energy orbit with $n = n_2 < n_1$. This photon will have the wavelength λ_{n_1,n_2} determined by (as follows from $E_{n_1,n_2} = h\nu_{n_1,n_2}$ and $\nu = c/\lambda$)

$$\frac{1}{\lambda_{n_1,n_2}} = \frac{E_{n_1,n_2}}{hc} = \frac{1}{2}\left(\frac{1}{4\pi\epsilon_0}\right)^2 \frac{m_e e^4}{\hbar^2 hc}\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right), \tag{1.15}$$

which agrees with Rydberg's formula (1.9) if

$$R = \frac{1}{2}\left(\frac{1}{4\pi\epsilon_0}\right)^2 \frac{m_e e^4}{\hbar^2 hc}. \tag{1.16}$$

To determine $\hbar$, Bohr reasoned that for large n (and thus large r), the behaviour should be similar to the classical one, and so the frequency of radiation should be given by the frequency ν with which the electron revolves around the nucleus,

$$\nu = \frac{v}{2\pi r} = \frac{1}{2\pi}\left(\frac{1}{4\pi\epsilon_0}\right)^2 \frac{m_e e^4}{n^3\hbar^3}. \tag{1.17}$$

On the other hand, from the quantum point of view, it should correspond to transitions between neighbouring states E_n and E_{n-1}, which produce photons with frequency ν given by

$$h\nu = \frac{m_e e^4}{2\hbar^2}\left(\frac{1}{4\pi\epsilon_0}\right)^2 \left(\frac{1}{(n-1)^2} - \frac{1}{n^2}\right) = \left(\frac{1}{4\pi\epsilon_0}\right)^2 \frac{m_e e^4}{\hbar^2 n^3} + O(n^{-4}). \tag{1.18}$$

Comparing (1.17) and (1.18) we find

$$\frac{hm_e e^4}{2\pi\hbar^3 n^3} = \frac{m_e e^4}{\hbar^2 n^3}, \tag{1.19}$$

which implies

$$\hbar = \frac{h}{2\pi}. \tag{1.20}$$

The constant $\hbar$ is called the reduced Planck's constant. Plugging this into (1.16) we find

$$R = \frac{m_e e^4}{8\epsilon_0^2 ch^3}, \tag{1.21}$$

which matched the experimental value.

This set of postulates, namely

- the electron can stay in the stable orbits (energy levels) determined by the quantisation condition (1.10) without producing electromagnetic radiation,
- the electron can transition between different energy levels E_{n_1} and E_{n_2} by emitting a photon with energy $E_{n_1} - E_{n_2}$,

are known as the Bohr's model of the atom. While it successfully explains the values of the energy levels and the spectrum of hydrogen, it does not explain where the postulates are coming from. For example, how do we compute spectra of atoms with more electrons, such as helium which has two electrons? New ideas were needed to explain the origin of Bohr's model.

1.6 The correspondence principle

In the discussion of Bohr's atom, we have used the idea that in some limit, the *quantum calculation* should reproduce the classical one. The reason for that is, of course, that the classical theory was developed to match the experiments of the time, and it is accurate in most situations. If we assume that the quantum theory should be more general, extending the classical theory to a wider domain of experiments and physical phenomena, it must be the case that the quantum theory should reproduce the classical theory in an appropriate limit. This is Bohr's correspondence principle.

In which situations should we recover classical theory? So far, the key element of quantum phenomena was that various quantities were quantised in multiples of Planck's constant h, e.g. $E = nh\nu$, and $L = n\hbar$. Therefore, we should expect the classical theory to be valid in situations when h can be assumed to be negligibly small (so there is effectively no quantisation), or, when relevant, in circumstances where the quantum numbers are large, i.e. for large n. However

it is not always as simple as taking a limit in which n becomes large, the key requirement is that in a regime in which classical experimental results become valid can be obtained then the quantum description should match the classical case. In the example of the hydrogen atom, when the angular momentum ($L = n\hbar$) is large (i.e. for large n) compared to $\hbar\nu$, and so also the radius and energy of the electron is large, the classical approximation should become valid. It is when n is large that the difference in energy levels of the hydrogen atom become very close together, approaching a continuum in the limit, and it is the continuous release of energy which is apparent in classical physics. In the example of the black-body radiation each photon has energy $h\nu$ and in the setting where this number is very small, i.e. for temperatures T such that $k_B T \gg h\nu$ then we recover the classical results (the exchange of energy lumps approaches the limit of continuous exchange of energy and the Rayleigh-Jeans law emerges from Planck's expression for the spectral energy density).

1.7 de Broglie waves

In 1923 Louis de Broglie suggested that since light apparently can behave both as a particle and as a wave, it could be true that other particles, such as electrons, have this property as well. In other words, there could exists situations in which the electron behaves as a wave.

If the electron behaves as a wave, then what is the frequency ν and the wavelength λ of this wave? de Broglie conjectured the following relations,

$$E = h\nu, \tag{1.22}$$

$$p = \frac{h}{\lambda}, \tag{1.23}$$

where E is the energy of the electron and $p = mv$ is its momentum. de Broglie motivated these relations using special relativity. A simple way to see that they make sense is by noting that the first relation is just the same relation as was found for photons. It turns out that the second one is too: the momentum of a photon is related to its energy by $E = pc$, which implies $p = h/\lambda$. de Broglie conjectured that these relations hold also for electrons.

The first piece of evidence supporting de Broglie's theory is that it explains Bohr's quantisation condition (1.10). Indeed, if an electron is on a stable stationary orbit, it is natural to expect that there is an integer number of wavelengths λ that fit on this orbit. If the length of the orbit is ℓ we then expect $\ell = n\lambda$ for a positive integer n. The length of the orbit is $\ell = 2\pi r$ and (1.23) implies $\lambda = \frac{h}{m_e v}$, so from $\ell = n\lambda$ we find

$$2\pi r = \frac{nh}{m_e v}, \tag{1.24}$$

or

$$m_e r v = \frac{nh}{2\pi} = \hbar n, \tag{1.25}$$

precisely the Bohr's condition (1.10).

Another expectation that we might have if electron can behave as a wave is that it should be able to exhibit the phenomena of diffraction and interference, just like light. Indeed, in 1927 George Thomson and also Clinton Davisson and Lester Germer were able to experimentally observe diffraction of electrons, confirming de Broglie's theory.

The theoretical developments described in this chapter and their extensions formed what is now known as the "old quantum theory". While it did explain some experiments, it did so in a rather unsystematic and heuristic way. For example, while de Broglie's hypothesis postulates that electrons have wave-like properties, the equations governing these new waves were unknown. In 1925 Schrödinger proposed a new wave equation to describe de Broglie's waves and Heisenberg developed "matrix mechanics". This formed the basis of the modern quantum theory that we will study in the following chapters.

2. The Schrödinger Equation and Wave-Functions

2.1 An Equation for Matter Waves

Let us try to find a wave equation for the waves of the kind that de Broglie conjectured should describe electrons. A wave that describes an electron should have the following properties:

- it should have the frequency ν related to electron's energy by $E = h\nu$,
- it should have the wavelength λ related to electron's momentum by $p = \frac{h}{\lambda}$.

Suppose that the wave is described by a single function $\Psi(\boldsymbol{x}, t)$, where t is time and $\boldsymbol{x} \in \mathbb{R}^3$ given by $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ is the position in space.[1] Then a simple plane wave with the above properties is[2]

$$\Psi(\boldsymbol{x}, t) = e^{i(\boldsymbol{p}\cdot\boldsymbol{x} - Et)/\hbar}, \tag{2.1}$$

where, as before, $\hbar = h/2\pi$ and $\boldsymbol{p} = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}$. Of course, there is no reason to expect that an electron in something like a hydrogen atom would be described by such a simple wave. But it

[1] One can question this assumption and ask whether we should instead use a vector-valued function, like in the case of electromagnetic waves. This objection is perfectly valid and in fact (non-relativistic) electrons are described by two-component wave-functions and this is related to the notion of spin, which is outside of the scope of this module. In the absence of magnetic fields one can however get around with just one component.

[2] The overall sign in the exponential is a convention, and we choose it to match the standard one. (This choice dictates the signs in front of i throughout all of quantum physics.) The relative sign between Et and $\boldsymbol{p}\cdot\boldsymbol{x}$ is important and is chosen so that the wave travels in the direction of $\boldsymbol{p}$, not $-\boldsymbol{p}$. We will see this more importantly below.

is reasonable to assume that a free electron in empty space would. For such an electron the energy is $E = \frac{p^2}{2m}$, where m is its mass[3]. We would like our wave equation to allow (2.1) as a solution for all $\boldsymbol{p}$ and $E = \frac{p^2}{2m}$. The simplest equation that does the job is

$$i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{x},t) = -\frac{\hbar^2}{2m}\nabla^2\Psi(\boldsymbol{x},t), \tag{2.2}$$

where

$$\nabla^2 = \Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}. \tag{2.3}$$

Indeed, the left-hand side of (2.2) is simply $E\Psi(\boldsymbol{x},t)$ and the right-hand side is $\frac{\boldsymbol{p}^2}{2m}\Psi(\boldsymbol{x},t)$.

So, equation (2.2) describes the free plane waves (2.1). This is so far not very exciting (although it is already non-trivial, as we will see in the next section). In order to understand something like the Hydrogen atom, we need an equation that would describe the electron's motion under the influence of some external forces. So suppose that our electron moves not freely, but in some potential $V(\boldsymbol{x})$, so that its energy is now

$$E = \frac{p^2}{2m} + V(\boldsymbol{x}). \tag{2.4}$$

Schrödinger guessed that we should modify the wave equation by including $V(\boldsymbol{x})$ in (2.2)

$$\boxed{i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{x},t) = -\frac{\hbar^2}{2m}\nabla^2\Psi(\boldsymbol{x},t) + V(\boldsymbol{x})\Psi(\boldsymbol{x},t).} \tag{2.5}$$

This makes sense because the first term on the right-hand side of (2.5) in the case of plane waves (2.1) was equal to $\frac{p^2}{2m}\Psi(\boldsymbol{x},t)$, so if we want to find an analogue of (2.4), it is natural to just include $V(\boldsymbol{x})\Psi(\boldsymbol{x},t)$ in (2.5).

The equation (2.2) is known as the (time-dependent) Schrödinger equation. Of course, at this point (2.5) is just a guess. How do we know it is correct? Historically, the first piece of evidence was that Schrödinger solved this equation for the Hydrogen atom, where $V(\boldsymbol{x}) = -\frac{1}{4\pi\epsilon_0}\frac{e^2}{r}$. He obtained the energy levels which were conjectured by Bohr, and which were in a perfect agreement with the experiment. Since then the Schrödinger equation and its generalisations have been solved (often numerically) in lots of examples, and no disagreements with the experiment are known.

Comment(s). *1. The time-dependent Schrödinger equation is first-order in time. This means that if we know $\Psi(\boldsymbol{x},t_0)$ for all $\boldsymbol{x} \in \mathbb{R}^3$ at a fixed time $t = t_0$ then the right-hand-side of (2.5) allows us to find $\frac{\partial\Psi}{\partial t}$, which, in turn, can be used to find $\Psi(\boldsymbol{x},t)$ at other times.*

[3]We will only study non-relativistic physics, where this formula for the kinetic energy is correct. We could also add the $+mc^2$ term, but in the end this would drop out of all equations because it does not depend on $\boldsymbol{p}$.

2. *The time-dependent Schrödinger equation is a linear equation and so we can superpose (add) two solutions to find a third.*

We arrived at this equation motivated by describing the behaviour of the electron, but the same equation applies to all (non-relativistic) particles. In the rest of this section we will discuss the physics of the free Schrödinger equation and the meaning of the "wave-function" Ψ.

2.2 Interpreting the wave-function

Let us study the solutions of the free (i.e. $V(\boldsymbol{x}) = 0$) Schrödinger equation (2.2) in more detail. Let's first look at the plane wave (2.1),

$$\Psi(\boldsymbol{x}, t) = e^{i(\boldsymbol{p}\cdot\boldsymbol{x} - Et)/\hbar}. \tag{2.6}$$

At a moment of time t, where is the particle described by this wave? It seems natural to expect that $\Psi(\boldsymbol{x}, t)$ should vanish, or at least be very small, far away from the particle. However, the solution (2.6) is non-zero everywhere, and its "size" $|\Psi(\boldsymbol{x}, t)| = 1$ for all $\boldsymbol{x}$ and t.

So it seems impossible to tell where the particle is for the wave solution (2.6). We will understand why this is the case a bit later. Now (2.6) is a wave solution with the required properties of de Broglie's matter wave, but as the time-dependent Schrödinger in (2.5) is a linear equation we may find alternative wave solutions by taking sums, and even integrals, of waves having the same form as (2.6), i.e.

$$\Psi(\boldsymbol{x}, t) = \int d^3p \; A(\boldsymbol{p}) e^{i(\boldsymbol{p}\cdot\boldsymbol{x} - Et)/\hbar} \tag{2.7}$$

(where $\int d^3p = \int\int\int dp_1 dp_2 dp_3$) also satisfies $i\hbar\frac{\partial}{\partial t}\Psi = E\Psi$, $-\frac{\hbar^2}{2m}\nabla^2\Psi = \frac{p^2}{2m}\Psi$. The expression in (2.7) can be used to build a wavepacket, a travelling wave bounded by an enveloping wave-function and itself made up of simple harmonic waves. The overall shape of the travelling wave comes from the coefficients $A(\boldsymbol{p})$, we will look more closely at wavepackets in chapter 3.

We will now consider an alternative wave-function at $t = 0$, we will revisit it later and see it as a wavepacket, and we will show that it solves the Schrödinger equation in the problem sets. The function at $t = 0$ is

$$\Psi(\boldsymbol{x}, 0) = Ae^{-\frac{(\boldsymbol{x}-\boldsymbol{x}_0)^2}{4\Delta x^2} + i\boldsymbol{p}\cdot\boldsymbol{x}/\hbar}. \tag{2.8}$$

Here A is some real constant that we will determine momentarily, while Δx is a positive constant, $\boldsymbol{x}_0$ is a constant vector (used to inform of us about the location of a particle at $t = 0$) (we emphasise that Δx is a constant and not a function of $\boldsymbol{x}$). We are motivated to consider this wave-function as it, unlike (2.6) at $t = 0$, is very small unless $\boldsymbol{x}$ satisfies $|\boldsymbol{x} - \boldsymbol{x}_0|$ is less than

or of the order of Δx. But just as the wave-function in (2.6) it also contains the $i\boldsymbol{p}\cdot\boldsymbol{x}/\hbar$ in the exponential, which makes it at least look like the particle has momentum $\boldsymbol{p}$ located somewhere near $\boldsymbol{x}_0$.

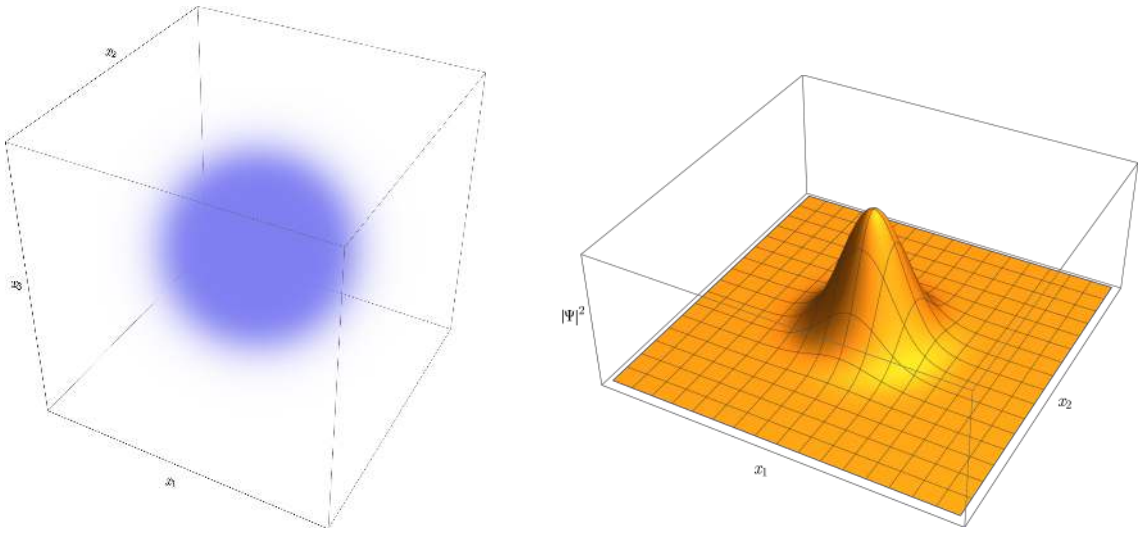


Figure 2.1: Left: a 3D density plot of the probability density (2.9) The diameter of the fuzzy ball is roughly $2\Delta x$. Right: a plot of $|\Psi|^2$ from (2.9) at $x_3 = 0$.

To quantify how large is $\Psi(\boldsymbol{x}, 0)$ let us consider its absolute value

$$|\Psi(\boldsymbol{x}, 0)|^2 = \Psi(\boldsymbol{x}, 0)^* \Psi(\boldsymbol{x}, 0) = A^2 e^{-\frac{(\boldsymbol{x}-\boldsymbol{x}_0)^2}{2\Delta x^2}}. \tag{2.9}$$

The plot of this function is shown in figure 2.1. We see that it is indeed large near $\boldsymbol{x} = \boldsymbol{x}_0$ and decays quickly away from it. What does it tell us about the position of the particle at $t = 0$? Initially physicists tried to interpret the wave-function like this as a "spread-out" particle. That is, the value of $|\Psi(\boldsymbol{x}, 0)|^2$ would determine "how much" of the particle is at a point x. This interpretation turned out to be **incorrect** because, as we will see in later examples, the wave-function can spread out in ways in which the real particles have not been observed to behave.

What is now understood to be the **correct** interpretation of the wave-function is that $|\Psi(\boldsymbol{x}, t)|^2$ gives the probability that the particle's position, if measured, will be $\boldsymbol{x}$. More precisely, $|\Psi(\boldsymbol{x}, t)|^2$ is the probability density, meaning that the probability $P(V)$ of finding the particle in some set $V \in \mathbb{R}^3$ is[4]

[4]Our notation in the integral $\int_V d^3x$ is a shorthand for the volume integral $\int\int\int dx_1 dx_2 dx_3$ over the volume $V \in \mathbb{R}^3$.

$$\boxed{P(V) = \int_V d^3x \, |\Psi(\boldsymbol{x}, t)|^2.} \tag{2.10}$$

How to interpret this fact is a matter of real debate - there are different philosophical takes on what the probability distribution derived from the wave-function "means" - but in this course we will adopt the "standard" or "Copenhagen" interpretation. It is important to note that the different interpretations do not differ on physical predictions, they all agree on what experimenters will see, just on what it all "means".

In the Copenhagen interpretation, the particle is still point-like, but we have to give up the classical idea that a particle always has a definite value of the position $\boldsymbol{x}$. This is different from probability theory or statistical physics, which try to capture our ignorance of the complete description of the system. In quantum theory we really believe that Ψ describes all there is to know about the particle. For this reason we will often refer to Ψ as the "state" or the "quantum state" of the particle.

Notice that the total probability $P(\mathbb{R}^3)$ of finding the particle anywhere in $\mathbb{R}^3$ must be 1, so we should have

$$1 = P(\mathbb{R}^3) = \int_{\mathbb{R}^3} d^3x \, |\Psi(\boldsymbol{x}, t)|^2. \tag{2.11}$$

We will often call this equation the normalisation condition for Ψ and we will refer to the functions that satisfy it as normalised. Sometimes we might have a wave-function Ψ which does not satisfy (2.11), but for which the integral

$$\mathcal{N} = \int_{\mathbb{R}^3} d^3x \, |\Psi(\boldsymbol{x}, t)|^2 \tag{2.12}$$

is finite, $\mathcal{N} < \infty$. We will call such wave-functions normalisable, because they can be normalised by defining a new wave-function

$$\Phi(\boldsymbol{x}, t) = \mathcal{N}^{-\frac{1}{2}} \Psi(\boldsymbol{x}, t). \tag{2.13}$$

Our example wave-function (2.8) is not normalised unless we choose the constant A carefully. The normalisation condition takes the form

$$1 = \int_{\mathbb{R}^3} d^3x \, |\Psi(\boldsymbol{x}, t)|^2 = A^2 \int_{\mathbb{R}^3} d^3x \, e^{-\frac{(\boldsymbol{x}-\boldsymbol{x}_0)^2}{2\Delta x^2}}. \tag{2.14}$$

We can compute the integral using the identity

$$\int_{-\infty}^{+\infty} dx e^{-x^2/2} = \sqrt{2\pi}. \tag{2.15}$$

For this we first change the integration variable to $\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ by substituting

$$\boldsymbol{x} = \Delta x\, \boldsymbol{y} + \boldsymbol{x}_0, \tag{2.16}$$

recalling that Δx and $\boldsymbol{x}_0$ are constants, to get

$$\int_{\mathbb{R}^3} d^3x\, e^{-\frac{(\boldsymbol{x}-\boldsymbol{x}_0)^2}{2\Delta x^2}} = \Delta x^3 \int_{\mathbb{R}^3} d^3y e^{-\frac{\boldsymbol{y}^2}{2}} = \Delta x^3 \left(\int_{-\infty}^{+\infty} dy_1\, e^{-y_1^2/2} \right)^3 = \Delta x^3 \left(2\pi \right)^{3/2}. \tag{2.17}$$

Here we have used (2.15) and the fact that $\boldsymbol{y}^2 = y_1^2 + y_2^2 + y_3^2$,[5] which allowed us to split the 3d integral into a product of 3 identical 1d integrals (over y_1, y_2, and y_3). Overall, we find for (2.14)

$$1 = A^2 \Delta x^3 (2\pi)^{3/2}, \tag{2.18}$$

which gives

$$A = \frac{1}{(\Delta x)^{3/2}(2\pi)^{3/4}}. \tag{2.19}$$

The probability density (2.9) then becomes

$$|\Psi(\boldsymbol{x}, 0)|^2 = \frac{1}{(2\pi)^{3/2}(\Delta x)^3} e^{-\frac{(\boldsymbol{x}-\boldsymbol{x}_0)^2}{2\Delta x^2}}. \tag{2.20}$$

The normalisation condition has been found for a wave-function at time $t = 0$ but we will need it to hold for all t and this will have come interesting consequences.

2.3 The Free Particle

The interpretation of the wave-function (2.8) is that it describes a particle which is likely to be found at time $t = 0$ near $\boldsymbol{x} = \boldsymbol{x}_0$. What happens to this wave-function at later times $t > 0$?

The free Schrödinger equation (2.2) expresses the time derivative of Ψ in terms of its spatial derivatives, which means that it can be used to determine $\Psi(\boldsymbol{x}, t)$ for all t starting with the expression (2.8) at $t = 0$. In the problem set you will verify that the solution takes the following form

$$\Psi(\boldsymbol{x}, t) = \frac{A}{\left(1 + \frac{i\hbar t}{2m\Delta x^2}\right)^{\frac{3}{2}}} \exp\left(-\frac{(\boldsymbol{x} - \boldsymbol{x}_0 - \boldsymbol{p}t/m)^2}{4(\Delta x^2 + i\hbar t/2m)} + i\boldsymbol{p}\cdot\boldsymbol{x}/\hbar - iEt/\hbar \right), \tag{2.21}$$

where $E = \frac{p^2}{2m}$, and A is given as before by (2.19).

[5] We are using the notation v_i with $i = 1, 2, 3$ for the components of a vector $\boldsymbol{v}$.

Let us see what this means for the probability density $|\Psi(x,t)|^2$. First of all, we compute

$$|\Psi(x,t)|^2 = \frac{A^2}{|1+\frac{i\hbar t}{2m\Delta x^2}|^3}\exp\left(-\frac{(\boldsymbol{x}-\boldsymbol{x}_0-\boldsymbol{p}t/m)^2}{4(\Delta x^2+i\hbar t/2m)}-\frac{(\boldsymbol{x}-\boldsymbol{x}_0-\boldsymbol{p}t/m)^2}{4(\Delta x^2-i\hbar t/2m)}\right) \tag{2.22}$$

$$= \frac{1}{(2\pi)^{3/2}\left(\Delta x^2+\frac{\hbar^2t^2}{4m^2\Delta x^2}\right)^{3/2}}\exp\left(-\frac{(\boldsymbol{x}-\boldsymbol{x}_0-\boldsymbol{p}t/m)^2}{4(\Delta x^2+\hbar^2t^2/4m^2\Delta x^2)}\right) \tag{2.23}$$

$$= \frac{1}{(2\pi)^{3/2}(\Delta x(t))^3}\exp\left(-\frac{(\boldsymbol{x}-\boldsymbol{x}_0(t))^2}{4(\Delta x(t))^2}\right), \tag{2.24}$$

where

$$\Delta x(t)^2 = \Delta x^2 + \frac{\hbar^2t^2}{4m^2\Delta x^2}, \tag{2.25}$$

$$\boldsymbol{x}_0(t) = \boldsymbol{x}_0 + \frac{\boldsymbol{p}}{m}t. \tag{2.26}$$

Just like $|\Psi(\boldsymbol{x},0)|^2$ in (2.20) the probability density $|\Psi(\boldsymbol{x},t)|^2$ is a Gaussian, but now centered at $\boldsymbol{x}_0(t)$ and with the growing "uncertainty" $\Delta x(t)$ given above.

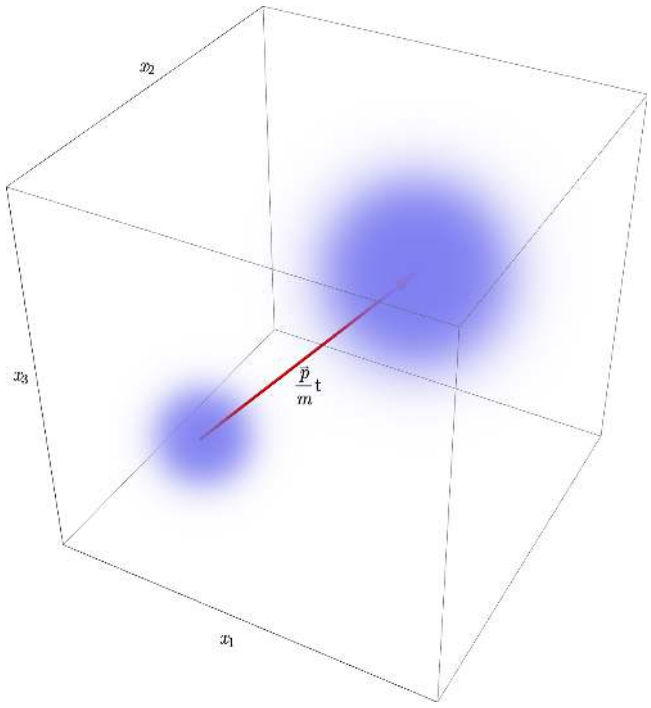


Figure 2.2: A schematic representation of the time evolution of the density (2.24): the center moves by $\frac{\boldsymbol{p}}{m}t$ at the same time as the uncertainty $\Delta x(t)$ grows according to (2.25).

Two things are therefore happening to our wave-function. Firstly, the center $\boldsymbol{x}_0(t)$ is moving with the velocity $\boldsymbol{v} = \boldsymbol{p}/m$, which is just the classical velocity of the particle. Secondly, the

probability distribution is spreading out as dictated by $\Delta x(t)$. Both effects are schematically illustrated in figure 2.2. We will come to this spreading of the wave-function in a second. Ignoring it for now, we see that the Schrödinger equation succeeds in describing the linear motion of a free particle, thus reproducing the simplest bit of classical physics.

2.4 Uncertainty and wave-function spreading

Let us now discuss the spreading of the wave-function. First of all, let's see how badly it spreads.

At time t, the uncertainty is given by (2.25). How much time does it take for $\Delta x(t)$ to double, i.e. what is the t for which $\Delta x(t) = 2\Delta x$?[6] We find from (2.25)

$$t = \sqrt{12}\frac{\Delta x^2 m}{\hbar}. \tag{2.27}$$

Let us estimate this for a person weighting $m = 80\,kg$ and initial uncertainty Δx of one angstrom, i.e. $\Delta x = 10^{-10}m$ – a typical atomic scale. We find, using $\hbar \approx 10^{-34}\,m^2 kg/s$, the order of magnitude

$$t \approx 10^{16}s \approx 10^8\,years. \tag{2.28}$$

This is a hundred million years, so clearly the spread of the wave-function is absolutely negligible for most applications in this setup. If we use instead the mass of an electron, $m \approx 10^{-30}kg$, we get

$$t \approx 10^{-16}s, \tag{2.29}$$

which is quite fast! This clearly shows that the spreading of the wave-function is much more important for electrons in atomic systems than for people.

How do we interpret this spreading of the wave-function? Let's look again at (2.25). For large times t we can ignore the first term (the initial uncertainty Δx) and write

$$\Delta x(t) \approx \frac{\hbar t}{2m\Delta x}. \tag{2.30}$$

Given the probabilistic interpretation of the wave-function, we can say that after a large time t, the particle's coordinate x_i will most likely be in

$$x_i \in \left[(x_0)_i + \frac{p_i}{m}t - \frac{\hbar t}{2m\Delta x},\ (x_0)_i + \frac{p_i}{m}t + \frac{\hbar t}{2m\Delta x}\right]. \tag{2.31}$$

[6] Note though that the growth of $\Delta x(t)$ is not exponential!

This looks very much as if we are uncertain about the particle's velocity $\boldsymbol{v}$ or, equivalently, its momentum $\boldsymbol{p} = m\boldsymbol{v}$, i.e. $p_i \in [p_i - \Delta p, p_i + \Delta p]$. That is, we can describe this growing uncertainty in particle's position by an initial uncertainty in its momentum,

$$\Delta p = \frac{\hbar}{2\Delta x}. \tag{2.32}$$

Equivalently, we find

$$\Delta p \Delta x = \frac{\hbar}{2}. \tag{2.33}$$

This is a special case of Heisenberg's uncertainty principle, which says that in all situations the uncertainty in particle's position and momentum satisfy

$$\Delta p \Delta x \geq \frac{\hbar}{2} \tag{2.34}$$

We will derive this principle in full generality later in this module.

We can now understand why we couldn't tell the location of the particle for the plane wave (2.6): we constructed that wave-function to have a definite momentum $\boldsymbol{p}$, which means that $\Delta p = 0$. But then (2.34) implies that $\Delta x = \infty$ and we shouldn't be able to tell where the particle is, which is indeed what we observed.

2.5 Operators and Expectation Values

We have seen that for a typical wave-function $\Psi(\boldsymbol{x}, t)$ there's some uncertainty not just in the position of the particle, but also in its momentum. The uncertainty in the position is easy to see even at fixed t, since $|\Psi(\boldsymbol{x}, t)|^2$ has the interpretation of the probability density in $\boldsymbol{x}$. To see the uncertainty in $\boldsymbol{p}$ we have to study how the wave-function evolved over time. Can we tell what Δp is by just looking at $\Psi(\boldsymbol{x}, t)$ at one t, e.g. $t = 0$?

The answer to this question is yes. As we will see later in this module, analogously to $\boldsymbol{x}$, one can compute a probability density for $\boldsymbol{p}$ from $\Psi(\boldsymbol{x}, t)$. For now, we will discuss a simpler way of dealing with the "observables" which are not $\boldsymbol{x}$, such as $\boldsymbol{p}$.

First, note that we can compute the expectation value of $\boldsymbol{x}$ in a normalised state Ψ using the formula

$$\langle \boldsymbol{x} \rangle = \int_{\mathbb{R}^3} d^3x \; \boldsymbol{x} |\Psi(\boldsymbol{x}, t)|^2. \tag{2.35}$$

We can rewrite this in the following way,

$$\langle \boldsymbol{x} \rangle = \int_{\mathbb{R}^3} d^3x \; \Psi(\boldsymbol{x}, t)^* \boldsymbol{x} \Psi(\boldsymbol{x}, t). \tag{2.36}$$

The idea is now to make the following liberal interpretation of this equation: we view the expression

$$\boldsymbol{x}\Psi(\boldsymbol{x},t) \tag{2.37}$$

as a *linear operator* $\hat{\boldsymbol{x}}$ acting on the function $\Psi(\boldsymbol{x},t)$.

Definition 2.5.1. A linear operator $\hat{\mathcal{O}}$ is a rule that takes a wave-function Ψ and produces a new wave-function that we denote $\hat{\mathcal{O}}\Psi$. Linearity means that it satisfies the property

$$\hat{\mathcal{O}}(\alpha\Psi_1+\beta\Psi_2)=\alpha\hat{\mathcal{O}}\Psi_1+\beta\hat{\mathcal{O}}\Psi_2. \tag{2.38}$$

The definition of the operator $\hat{\boldsymbol{x}}$ is just this: it multiplies the wave-function by $\boldsymbol{x}$. To be more precise, $\hat{\boldsymbol{x}}$ is a actually a vector of three operators $\hat{x}_1,\hat{x}_2,\hat{x}_3$. Each of these operators act by

$$(\hat{x}_i\Psi)(\boldsymbol{x},t)\equiv x_i\Psi(\boldsymbol{x},t). \tag{2.39}$$

We now postulate that every observable, for example the momentum $\boldsymbol{p}$, is described in the quantum theory by a linear operator. For $\boldsymbol{p}$, the operator is a vector with components

$$\hat{p}_i=-i\hbar\frac{\partial}{\partial x_i}. \tag{2.40}$$

or

$$\hat{\boldsymbol{p}}=-i\hbar\nabla. \tag{2.41}$$

To motivate this definition, consider the plane wave (2.6),[7]

$$\Psi(\boldsymbol{x},t)=e^{i(\boldsymbol{p}\cdot\boldsymbol{x}-Et)/\hbar}. \tag{2.42}$$

By construction it has a definite momentum $\boldsymbol{p}$. If we act on it with $\hat{p}_i$ given in (2.40), we find

$$(\hat{p}_i\Psi)(\boldsymbol{x},t)=-i\hbar\frac{\partial}{\partial x_i}e^{i(\boldsymbol{p}\cdot\boldsymbol{x}-Et)/\hbar}=p_ie^{i(\boldsymbol{p}\cdot\boldsymbol{x}-Et)/\hbar}=p_i\Psi(\boldsymbol{x},t). \tag{2.43}$$

So $\hat{p}_i$ acting on a Ψ with a definite momentum p_i just multiplies it by the value of p_i. We say that Ψ is an *eigenfunction* of $\hat{p}_i$ with eigenvalue p_i.

More generally, we say that a wave-function Ψ is an eigenfunction of $\hat{\mathcal{O}}$ with eigenvalue $\mathcal{O}\in\mathbb{R}$ if

$$\hat{\mathcal{O}}\Psi=\mathcal{O}\Psi. \tag{2.44}$$

[7] A cautious reader will notice that a plane wave is not normalised. In fact, it is not normalisable since for it $\int_{\mathbb{R}^3}d^3x|\Psi(\boldsymbol{x},t)|^2=\int_{\mathbb{R}^3}d^3x\,1=\infty$. We will learn how to deal with this problem later in the module, but for now note that it was from the plane wave that we derived the Schödinger equation, i.e. despite its problems, it is a very useful tool.

We interpret this to mean that the observable described by $\hat{\mathcal{O}}$ has a definite value in the state described by Ψ.

Finally, we can generalise (2.36) to a general observable $\hat{\mathcal{O}}$: the expectation value of $\hat{\mathcal{O}}$ in a normalised state Ψ is given by

$$\langle \mathcal{O} \rangle = \int_{\mathbb{R}^3} d^3x \Psi(x,t)^* (\hat{\mathcal{O}}\Psi)(x,t). \tag{2.45}$$

You will check in the problem set that the expectation value of $\hat{\boldsymbol{p}}$ in the state given by the wave-function Ψ in (2.8) is indeed $\boldsymbol{p}$,

$$\langle \hat{\boldsymbol{p}} \rangle = \boldsymbol{p}, \tag{2.46}$$

but Ψ is not an eigenfunction of $\hat{\boldsymbol{p}}$, so we should expect some uncertainty.

2.6 The Hamiltonian Operator and the Schrödinger equation

Using the language of operators, we can neatly rewrite the Schrödinger equation (2.5). First we notice that, using (2.40),

$$\hat{\boldsymbol{p}}^2 = \hat{p}_1^2 + \hat{p}_2^2 + \hat{p}_3^2 = -\hbar^2 \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} \right) = -\hbar^2 \nabla^2. \tag{2.47}$$

Using this, the Schrödinger equation (2.5) can be written as

$$i\hbar \frac{\partial}{\partial t} \Psi(\boldsymbol{x},t) = \frac{\hat{\boldsymbol{p}}^2}{2m} \Psi(\boldsymbol{x},t) + V(\boldsymbol{x})\Psi(\boldsymbol{x},t) = \left(\frac{\hat{\boldsymbol{p}}^2}{2m} + V(\boldsymbol{x}) \right) \Psi(\boldsymbol{x},t) = \hat{H}\Psi(\boldsymbol{x},t), \tag{2.48}$$

where

$$\hat{H} \equiv \frac{\hat{\boldsymbol{p}}^2}{2m} + V(\boldsymbol{x}), \tag{2.49}$$

is called the Hamiltonian operator. It generalises the classical expression

$$E = \frac{\boldsymbol{p}^2}{2m} + V(\boldsymbol{x}) \tag{2.50}$$

to a linear operator $\hat{H}$ by simply replacing $\boldsymbol{p} \to \hat{\boldsymbol{p}}$. To summarise, with the definition (2.49), the Schrödinger equation takes the form

$$\boxed{i\hbar \frac{\partial}{\partial t} \Psi(\boldsymbol{x},t) = \hat{H}\Psi(\boldsymbol{x},t).} \tag{2.51}$$

2.7 Conservation of Probability

We have seen that in order to interpret $\rho \equiv |\Psi(\boldsymbol{x}, t)|^2$ as the probability density, we need the normalisation condition (2.11),

$$\int_{\mathbb{R}^3} d^3x \rho = \int_{\mathbb{R}^3} d^3x |\Psi(\boldsymbol{x}, t)|^2 = 1, \tag{2.52}$$

to hold. At the same time, the Schrödinger equation (2.5) (or (2.51)) predicts the time derivative of Ψ, i.e. it dictates how Ψ evolves in time. It could happen that this evolution is inconsistent with the normalisation condition $\int_{\mathbb{R}^3} d^3x |\Psi(\boldsymbol{x}, t)|^2 = 1$. In other words, the value of $\Psi(\boldsymbol{x}, 0)$ determines, through the Schrödinger equation, $\Psi(\boldsymbol{x}, t)$ for all times t. Is it guaranteed that if $\Psi(\boldsymbol{x}, 0)$ is normalised, then $\Psi(\boldsymbol{x}, t)$ is normalised for all t?

To check this, let us compute the time derivative of $\rho = \Psi^* \Psi$ where for notational simplicity we have written Ψ in place of $\Psi(\boldsymbol{x}, t)$. Then we have

$$\begin{aligned}
\frac{\partial \rho}{\partial t} &= \frac{\partial \Psi^*}{\partial t} \Psi + \Psi^* \frac{\partial \Psi}{\partial t} \\
&= \frac{i}{\hbar} \left((\hat{H}\Psi)^* \Psi - \Psi^* (\hat{H}\Psi) \right) \\
&= \frac{i}{\hbar} \left(\left(-\frac{\hbar^2}{2m} \nabla^2 \Psi + V(\hat{x}) \Psi \right)^* \Psi - \Psi^* \left(-\frac{\hbar^2}{2m} \nabla^2 \Psi + V(\hat{x}) \Psi \right) \right) \\
&= \frac{i\hbar}{2m} \left(-(\nabla^2 \Psi^*) \Psi + \Psi^* \left(\nabla^2 \Psi \right) \right) \\
&= \boldsymbol{\nabla} \cdot \left(-\frac{i\hbar}{2m} \left((\boldsymbol{\nabla} \Psi^*) \Psi - \Psi^* (\boldsymbol{\nabla} \Psi) \right) \right) \\
&\equiv -\boldsymbol{\nabla} \cdot \boldsymbol{J}.
\end{aligned} \tag{2.53}$$

Altogether we can write the equation above as

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{J} = 0 \tag{2.54}$$

and (2.54) is called the continuity equation for probability, while

$$\boldsymbol{J} = \frac{i\hbar}{2m} \left((\boldsymbol{\nabla} \Psi^*) \Psi - \Psi^* (\boldsymbol{\nabla} \Psi) \right) \tag{2.55}$$

is known as the probability current, as it encodes the flow of probability across space. In particular consider a finite volume $V \subset \mathbb{R}^3$ with boundary S and consider the rate of change

of the probability of finding a particle (recall equation (2.10)) with wave-function $\Psi(\boldsymbol{x}, t)$ in V:

$$\begin{aligned}\frac{dP(V)}{dt} &= \int_V \frac{\partial \rho}{\partial t}\, d^3x \\ &= -\int_V \boldsymbol{\nabla}\cdot\boldsymbol{J} d^3x \\ &= -\oint_S \boldsymbol{J}\cdot d\boldsymbol{a}\end{aligned} \tag{2.56}$$

where we have used the divergence theorem (also known as Gauss's theorem) from vector calculus to convert the volume integral to the integral over the surface in the last line. The equation above implies that the rate of change of $P(V)$ is equal to the amount of $\boldsymbol{J}$ flowing outwards at the the surface of V, this is why $\boldsymbol{J}$ is called the probability current. If instead we allow V to grow to be all of space, i.e. $V \to \mathbb{R}^3$ then if total probability of finding a particle associated to Ψ somewhere in space is to remain equal to one, then it must be that there is no loss of probability at infinity, i.e. if $\frac{dP(\mathbb{R}^3)}{dt} = 0$ then at infinity (the boundary of $\mathbb{R}^3$) we must have that the probability current vanishes: $\boldsymbol{J} = \boldsymbol{0}$.

We can ensure that $\boldsymbol{J}$ vanishes at infinity by imposing the following conditions on any wave-function:

$$\boxed{\begin{aligned}\lim_{x\to\infty}\left[\Psi(\boldsymbol{x}, t)\right] &= 0 \qquad (2.57)\\ \lim_{x\to\infty}\left[\boldsymbol{\nabla}\Psi(\boldsymbol{x}, t)\right] &\in \mathbb{R}^3 \qquad (2.58)\end{aligned}}$$

i.e. the wave-function should vanish at infinity, and its derivative must be finite in any direction. These conditions are consistent with the idea that a wave-function should be normalisable.

The condition that the probability current vanishes at infinity has consequences for the Hamiltonian operator:

$$\frac{dP(\mathbb{R}^3)}{dt} = \frac{i}{\hbar}\int_{\mathbb{R}^3} d^3x \left((\hat{H}\Psi)^*\Psi - \Psi^*(\hat{H}\Psi)\right) = 0 \tag{2.59}$$

or if we rearrange the equation:

$$\int_{\mathbb{R}^3} d^3x \left((\hat{H}\Psi)^*\Psi\right) = \int_{\mathbb{R}^3} d^3x\ \Psi^*(\hat{H}\Psi) \tag{2.60}$$

any such operator (not only $\hat{H}$) which satisfies the above condition is called Hermitian.

Definition 2.7.1. An operator $\hat{\mathcal{O}}$ for which

$$\int_{\mathbb{R}^3} d^3x \Psi_1(\boldsymbol{x}, t)^*\hat{\mathcal{O}}\Psi_2(\boldsymbol{x}, t) = \int_{\mathbb{R}^3} d^3x \left(\hat{\mathcal{O}}\Psi_1(\boldsymbol{x}, t)\right)^* \Psi_2(\boldsymbol{x}, t) \tag{2.61}$$

for any Ψ_1 and Ψ_2 is said to be Hermitian.[8]

For example, let us check that the momentum operator $\hat{p}_i$ is Hermitian,

$$\begin{aligned}
&\int_{\mathbb{R}^3} d^3x \left(\hat{p}_i \Psi_1(\boldsymbol{x},t)\right)^* \Psi_2(\boldsymbol{x},t) = \int_{\mathbb{R}^3} d^3x \left(-i\hbar \frac{\partial}{\partial x_i}\Psi_1(\boldsymbol{x},t)\right)^* \Psi_2(\boldsymbol{x},t) \\
&= i\hbar \int_{\mathbb{R}^3} d^3x \left(\frac{\partial}{\partial x_i}\Psi_1(\boldsymbol{x},t)^*\right) \Psi_2(\boldsymbol{x},t) \\
&= -i\hbar \int_{\mathbb{R}^3} d^3x \Psi_1(\boldsymbol{x},t)^* \frac{\partial}{\partial x_i}\Psi_2(\boldsymbol{x},t) \\
&= \int_{\mathbb{R}^3} d^3x \left(\Psi_1(\boldsymbol{x},t)\right)^* \hat{p}_i \Psi_2(\boldsymbol{x},t). \qquad (2.62)
\end{aligned}$$

Here we used integration by parts to arrive to the second to last line, in which we dropped the boundary terms because we have used that $\Psi(\boldsymbol{x},t)$ tends to 0 as $\boldsymbol{x} \to \infty$. We will learn more about Hermitian operators a bit later when we develop better notation.

In the problem set, you will show in a very similar way that the Hamiltonian $\hat{H}$ given by (2.49) is a Hermitian operator. Using this fact and the definition 2.7.1 (with $\Psi_1 = \Psi_2 = \Psi$), we can directly conclude that (2.59) is zero. Note that all of our arguments are consistent: the vanishing of the probability current $\boldsymbol{J}$ at infinity agrees with the Hamiltonian operator being Hermitian.

We have shown that the norm

$$\int_{\mathbb{R}^3} d^3x |\Psi(\boldsymbol{x},t)|^2 \qquad (2.63)$$

is constant if the Hamiltonian is Hermitian, the wave-function Ψ vanishes at infinity and has only finite derivatives at infinity. If it is equal to 1 at one moment in time, it is 1 for all other t. We say that the total probability is conserved under time evolution, or that the time evolution is unitary. The origin of the word "unitary" will become clearer later in the module.

2.8 Evolution of expectation values and the Ehrenfest theorem

A similar exercise that we can do is to consider how expectation values of various quantities evolve with time. Let's focus on the expectation value of some operator $\hat{\mathcal{O}}$ (which we assume does not explicitly depend on t),

$$\langle \hat{\mathcal{O}} \rangle = \int_{\mathbb{R}^3} d^3x \Psi(\boldsymbol{x},t)^* \hat{\mathcal{O}} \Psi(\boldsymbol{x},t). \qquad (2.64)$$

[8]This at least is how physicists use the term, and we will adopt this definition in this module. In functional analysis, the definitions are more refined and more technical. We will briefly mention them later in this module.

Taking the time derivative, we find

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{\mathcal{O}}\rangle &= \int_{\mathbb{R}^3} d^3x \left(\frac{\partial}{\partial t}\Psi(\boldsymbol{x},t)\right)^* \hat{\mathcal{O}}\Psi(\boldsymbol{x},t) + \int_{\mathbb{R}^3} d^3x \Psi(\boldsymbol{x},t)^* \hat{\mathcal{O}}\frac{\partial}{\partial t}\Psi(\boldsymbol{x},t) \\ &= \frac{i}{\hbar}\left(\int_{\mathbb{R}^3} d^3x \left(\hat{H}\Psi(\boldsymbol{x},t)\right)^* \hat{\mathcal{O}}\Psi(\boldsymbol{x},t) - \int_{\mathbb{R}^3} d^3x \Psi(\boldsymbol{x},t)^* \hat{\mathcal{O}}\hat{H}\Psi(\boldsymbol{x},t)\right). \end{aligned} \tag{2.65}$$

The Hermiticity of $\hat{H}$ implies (using the definition 2.7.1 with $\Psi_1 = \Psi$ and $\Psi_2 = \hat{\mathcal{O}}\Psi$)

$$\int_{\mathbb{R}^3} d^3x \left(\hat{H}\Psi(\boldsymbol{x},t)\right)^* \hat{\mathcal{O}}\Psi(\boldsymbol{x},t) = \int_{\mathbb{R}^3} d^3x \Psi(\boldsymbol{x},t)^* \hat{H}\hat{\mathcal{O}}\Psi(\boldsymbol{x},t). \tag{2.66}$$

Plugging this into the above, we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{\mathcal{O}}\rangle &= \frac{i}{\hbar}\left(\int_{\mathbb{R}^3} d^3x \left(\Psi(\boldsymbol{x},t)\right)^* \hat{H}\hat{\mathcal{O}}\Psi(\boldsymbol{x},t) - \int_{\mathbb{R}^3} d^3x \Psi(\boldsymbol{x},t)^* \hat{\mathcal{O}}\hat{H}\Psi(\boldsymbol{x},t)\right) \\ &= \frac{i}{\hbar}\left(\int_{\mathbb{R}^3} d^3x \left(\Psi(\boldsymbol{x},t)\right)^* (\hat{H}\hat{\mathcal{O}} - \hat{\mathcal{O}}\hat{H})\Psi(\boldsymbol{x},t)\right) \\ &= \frac{i}{\hbar}\langle\hat{H}\hat{\mathcal{O}} - \hat{\mathcal{O}}\hat{H}\rangle. \end{aligned} \tag{2.67}$$

Expressions such as $\hat{H}\hat{\mathcal{O}} - \hat{\mathcal{O}}\hat{H}$ are extremely common in quantum theory and have their special notation,

$$[\hat{H}, \hat{\mathcal{O}}] \equiv \hat{H}\hat{\mathcal{O}} - \hat{\mathcal{O}}\hat{H}, \tag{2.68}$$

called the commutator. We have thus proved the following

Theorem 2.1. *For any time-independent observable $\hat{\mathcal{O}}$, the time derivative of its expectation value is given by*

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{\mathcal{O}}\rangle = \frac{i}{\hbar}\langle[\hat{H}, \hat{\mathcal{O}}]\rangle, \tag{2.69}$$

where $\hat{H}$ is the Hamiltonian operator.

Comment(s). *1. Note that we only used the Schrödinger equation in the form* (2.51)*, and did not have to specify $\hat{H}$ (besides assuming its Hermiticity).*

2. Compare this expression with the similar statement in the Hamiltonian formulation of classical mechanics for an explicitly time-independent function $f = f(q,p)$ on phase space (i.e. $\frac{\partial f}{\partial t} = 0$):

$$\begin{aligned}\frac{d}{dt}(f(q,p)) &= \sum_i \left(\frac{\partial f}{\partial q_i}\frac{\partial q_i}{\partial t} + \frac{\partial f}{\partial p_i}\frac{\partial p_i}{\partial t}\right) \\ &= \sum_i \left(\frac{\partial f}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q_i}\right) \\ &= \{f, H\}\end{aligned}$$

where H is the Hamiltonian of the system and $\{f, g\}$ denotes the Poisson bracket of the functions f and g. In quantum mechanics, the function is replaced by an expectation value of an operator, and up to coefficients ($\frac{1}{i\hbar}$) the Possion bracket has been replaced with the expectation of a a commutator.

Let us apply theorem 2.1 to the operators $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{p}}$. This will also give us the first example of non-commutativity of quantum-mechanical observables: in the classical theory where all observables are numbers, expressions like (2.68) vanish. In quantum theory observables are linear operators (and so expressible as matrices), and these don't have to commute.

We first start with $\hat{\mathcal{O}} = \hat{p}_i$. Recall the expression for the Hamiltonian (2.49), which implies

$$[\hat{H}, \hat{p}_i] = \left[\frac{\hat{\boldsymbol{p}}^2}{2m} + V(\boldsymbol{x}), \hat{p}_i\right]. \tag{2.70}$$

At this point we can use the fact that $[\alpha\hat{A}+\beta\hat{B}, \hat{C}] = \alpha[\hat{A}, \hat{C}]+\beta[\hat{B}, \hat{C}]$ for any linear operators $\hat{A}, \hat{B}, \hat{C}$ and numbers $\alpha, \beta \in \mathbb{C}$ (check it using the definition 2.5.1 and (2.68)!) to get

$$[\hat{H}, \hat{p}_i] = \frac{1}{2m}\left[\hat{\boldsymbol{p}}^2, \hat{p}_i\right] + \left[V(\boldsymbol{x}), \hat{p}_i\right], \tag{2.71}$$

so we can compute each of the terms separately.

We first compute $[V(\boldsymbol{x}), \hat{p}_i]$,

$$[V(\boldsymbol{x}), \hat{p}_i] = -i\hbar\left[V(\boldsymbol{x}), \frac{\partial}{\partial x_i}\right], \tag{2.72}$$

where we used (2.40). To understand the commutator, consider how it acts on a function $\Psi(\boldsymbol{x}, t)$,

$$\begin{aligned}\left[V(\boldsymbol{x}), \frac{\partial}{\partial x_i}\right]\Psi(\boldsymbol{x}, t) &= V(\boldsymbol{x})\frac{\partial}{\partial x_i}\Psi(\boldsymbol{x}, t) - \frac{\partial}{\partial x_i}\left(V(\boldsymbol{x})\Psi(\boldsymbol{x}, t)\right) \\ &= V(\boldsymbol{x})\frac{\partial}{\partial x_i}\Psi(\boldsymbol{x}, t) - \frac{\partial V(\boldsymbol{x})}{\partial x_i}\Psi(\boldsymbol{x}, t) - V(\boldsymbol{x})\frac{\partial}{\partial x_i}\Psi(\boldsymbol{x}, t) \\ &= -\frac{\partial V(\boldsymbol{x})}{\partial x_i}\Psi(\boldsymbol{x}, t). \end{aligned} \tag{2.73}$$

So we conclude that

$$\left[V(\boldsymbol{x}), \frac{\partial}{\partial x_i}\right] = -\frac{\partial V(\boldsymbol{x})}{\partial x_i}, \tag{2.74}$$

$$[V(\boldsymbol{x}), \hat{p}_i] = i\hbar\frac{\partial V(\boldsymbol{x})}{\partial x_i}. \tag{2.75}$$

Here we understand functions $f(\boldsymbol{x})$ as operators which simply multiply by $f(\boldsymbol{x})$.

We can similarly compute $\left[\hat{\boldsymbol{p}}^2, \hat{p}_i\right]$ by writing $\hat{\boldsymbol{p}}^2 = \hat{p}_1^2 + \hat{p}_2^2 + \hat{p}_3^2$, and noting that for any k, from (2.40),

$$[\hat{p}_k^2, \hat{p}_i] = i\hbar^3 \left[\frac{\partial^2}{\partial x_k^2}, \frac{\partial}{\partial x_i}\right]. \tag{2.76}$$

Acting on a function $\Psi(\boldsymbol{x}, t)$ with the above commutator we find that it is zero,

$$\left[\frac{\partial^2}{\partial x_k^2}, \frac{\partial}{\partial x_i}\right] = 0 \tag{2.77}$$

because the order in which we take the derivatives does not matter (check these steps yourself!).

We find then that $\left[\hat{\boldsymbol{p}}^2, \hat{p}_i\right] = 0$, and combining this with (2.71) and (2.75) we find

$$[\hat{H}, \hat{p}_i] = i\hbar \frac{\partial V(\boldsymbol{x})}{\partial x_i}. \tag{2.78}$$

We can use this in theorem (2.1) to conclude that in any state we have

$$\boxed{\frac{\partial}{\partial t}\langle \hat{p}_i \rangle = -\left\langle \frac{\partial V(\boldsymbol{x})}{\partial x_i} \right\rangle.} \tag{2.79}$$

In the problem set you will show that

$$[\hat{H}, \hat{x}_i] = -\frac{i\hbar \hat{p}_i}{m}, \tag{2.80}$$

which then implies, using theorem (2.1),

$$\boxed{\frac{\partial}{\partial t}\langle \hat{x}_i \rangle = \frac{\langle \hat{p}_i \rangle}{m}.} \tag{2.81}$$

The two equations (2.79) and (2.81) comprise **the Ehrenfest theorem**. The importance of this theorem is that it can be contrasted with the classical equations of motion

$$\frac{\partial}{\partial t} x_i = \frac{p_i}{m}, \tag{2.82}$$

$$\frac{\partial}{\partial t} p_i = -\frac{\partial V(\boldsymbol{x})}{\partial x_i}. \tag{2.83}$$

These would be the same as (2.79) and (2.81) under the identification $x_i = \langle \hat{x}_i \rangle$ and $p_i = \langle \hat{p}_i \rangle$ if we had

$$\left\langle \frac{\partial V(\boldsymbol{x})}{\partial x_i} \right\rangle \stackrel{?}{=} \frac{\partial V(\langle \boldsymbol{x} \rangle)}{\partial x_i}, \tag{2.84}$$

which is not true for general V. However, if the wave-function Ψ in which we compute the expectation values is tightly localised near just one point, similar the case considered in section 2.3, i.e. the uncertainly Δx of x is sufficiently small so that $V(\boldsymbol{x})$ varies only a little over this scale,[9] then we do have

$$\left\langle \frac{\partial V(\boldsymbol{x})}{\partial x_i} \right\rangle \approx \frac{\partial V(\langle \boldsymbol{x} \rangle)}{\partial x_i} \tag{2.85}$$

and we recover the classical equations of motion for the expectation values $\langle \hat{x}_i \rangle$ and $\langle \hat{p}_i \rangle$.

[9] And if we can neglect the wave-function spreading over the range of times in which we are studying the system.

3. Simple Quantum-Mechanical Systems

In this section we will solve the the Schrödinger equation for some simple one-dimensional systems and analyze the meaning of the solutions.

3.1 Separation of Variables and Energy Levels

The wave-functions satisfy the Schrödinger equation (2.5),

$$i\hbar\frac{\partial\Psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,t)}{\partial x^2} + V(x)\Psi(x,t). \tag{3.1}$$

We will use the following trick, known as separation of variables, to solve this equation. First, we assume that $\Psi(x,t)$ factorises as

$$\Psi(x,t) = \phi(t)\varphi(x), \tag{3.2}$$

where ϕ and φ are some new unknown functions. Plugging this into the Schrödinger equation above we find

$$i\hbar\frac{\partial\phi(t)}{\partial t}\varphi(x) = -\frac{\hbar^2}{2m}\frac{\partial^2\varphi(x)}{\partial x^2}\phi(t) + V(x)\phi(t)\varphi(x). \tag{3.3}$$

Dividing by $\phi(t)\varphi(x)$, we find

$$i\hbar\frac{\partial\phi(t)}{\partial t}\frac{1}{\phi(t)} = \frac{1}{\varphi(x)}\left(-\frac{\hbar^2}{2m}\frac{\partial^2\varphi(x)}{\partial x^2} + V(x)\varphi(x)\right). \tag{3.4}$$

Now the left-hand side only depends on t, while the right-hand side only depends on x. This can happen only if both sides are equal to a constant which we can call E,

$$i\hbar\frac{\partial\phi(t)}{\partial t}\frac{1}{\phi(t)} = E, \tag{3.5}$$

$$\frac{1}{\varphi(x)}\left(-\frac{\hbar^2}{2m}\frac{\partial^2\varphi(x)}{\partial x^2} + V(x)\varphi(x)\right) = E. \tag{3.6}$$

It is easy to solve the first equation, the solution is just[1]

$$\phi(t) = e^{-iEt/\hbar}. \tag{3.7}$$

So we find that there are solutions of the equation (3.1) of the form

$$\Psi(x,t) = e^{-iEt/\hbar}\varphi(x), \tag{3.8}$$

where $\varphi(x)$ solves (3.6) that can also be written as

$$-\frac{\hbar^2}{2m}\frac{\partial^2\varphi(x)}{\partial x^2} + V(x)\varphi(x) = E\varphi(x). \tag{3.9}$$

We can interpret the above equation (3.9) as follows. Recall the Hamiltonian operator given (2.49), which allows us to write (3.9) as

$$\boxed{\hat{H}\varphi(x) = E\varphi(x).} \tag{3.10}$$

Equation (3.10) (or its version with $\hat{H}$ plugged in) is often called **the time-independent Schrödinger equation**. This equation just says that $\varphi(x)$ is an eigenfunction of $\hat{H}$, i.e. a particle in the state $\varphi(x)$ has a definite value of energy and this value is equal to E. The functions Note that the full time-dependent wave-function, given by (3.8), at any given time t differs from $\varphi(x)$ only by an x-independent pure phase $e^{-iEt/\hbar}$. Consequently the probability density described by $|\Psi(x,t)|^2 = |\varphi(x)|^2$ is independent of time, such a state is called a stationary state. In particular, a particle described by $\Psi(x,t) = \phi(t)\varphi(x)$ is an eigenfunction of $\bar{H}$ and has a definite value of energy, equal to E, at any time t.

When we study example of solving (3.10) for some simple potentials we will find that we only have normalisable solutions $\varphi(x)$ for a discrete set of values of E, which we call the energy levels of our system. This is the analogue of the set of energy levels that we discussed in the context of Bohr's atom, but now described in a systematic way using wave-functions and operators. While the time-dependent Schrödinger equation tells us how states evolve in time, the time-independent version tells us which values of the energy E are allowed for the system.

3.2 The Infinite Potential Well

Consider a particle that can move in only one dimension, parametrised by a coordinate x. Suppose that we confine our particle to a length L, given by $0 < x < L$. The particle is free to

[1]We could multiply this $\phi(t)$ by any constant to get another solution, but such constants can always be absorbed into $\varphi(x)$ since the only important combination is $\Psi(x,t) = \phi(t)\varphi(x)$

move inside the (one-dimensional) box, but cannot exit it (i.e. $V(x) = 0$ for $0 < x < L$, and $V(x) = \infty$ for $x \leq 0$ and $x \geq L$ - i.e. there is an infinite potential well used to confine the particle). We will model this situation by requiring the wave-function $\Psi(x,t)$ to vanish outside of the box,

$$\Psi(x,t) = 0 \quad \text{if } x \leq 0 \text{ or } x \geq L. \tag{3.11}$$

The general solution of (3.9) is given by

$$\varphi(x) = A\cos(kx) + B\sin(kx), \tag{3.12}$$

where

$$k = \sqrt{2mE}/\hbar, \tag{3.13}$$

and A, B are arbitrary constants. This might make it seem like any value of $E = \frac{\hbar^2 k^2}{2m}$ is allowed. However, we have insisted that our particle is confined to a box, and so the wave-function should satisfy (3.11), this will restrict which values of k are possible. The solution (3.12) only describes $\varphi(x)$ for $0 \leq x \leq L$, so it is allowed to be non-vanishing, but it still has vanish at the end points $x = 0$ and $x = L$. This means

$$\varphi(0) = A = 0, \tag{3.14}$$

$$\varphi(L) = B\sin kL = 0. \tag{3.15}$$

The first equation implies that $A = 0$. The second implies that either $B = 0$, or $\sin(kL) = 0$. We cannot have $B = 0$ because then the whole wave-function would vanish, and so we conclude

$$\sin(kL) = 0. \tag{3.16}$$

This implies

$$kL = n\pi, \tag{3.17}$$

where $n \in \mathbb{Z}_{>0}$ is a positive integer.[2] Expressing this in terms of E, we find

$$E = E_n \equiv \frac{\pi^2\hbar^2 n^2}{2mL^2}. \tag{3.18}$$

These are the allowed energy levels for the particle in a box of size L in one dimension.

[2]It is non-negative since by construction k and L are non-negative. It can't be 0 because that would imply $k = 0$, and so $\sin kx = 0$, and then the whole wave-function would vanish.

It remains to normalise the wave-function. We have[3]

$$1 = \int_0^L |\varphi(x)|^2 dx \tag{3.19}$$

$$= B^2 \int_0^L (\sin kx)^2 dx \tag{3.20}$$

$$= B^2 \int_0^L \frac{1-\cos 2kx}{2} dx$$

$$= B^2 \left[\frac{x}{2} - \frac{\sin 2kx}{4k}\right]_0^L \tag{3.21}$$

$$= \frac{B^2 L}{2}, \tag{3.22}$$

where we used that $\sin 2kL = 2\sin kL \cos kL = 0$ due to (3.16). We conclude $B = \sqrt{2/L}$.

Combining the above results with (3.8), we find the time-dependent solutions of Schrödinger equation

$$\Psi_n(x,t) = e^{-iE_n t/\hbar}\varphi_n(x), \tag{3.23}$$

$$\varphi_n(x) = \sqrt{\frac{2}{L}} \sin(\tfrac{\pi n}{L} x), \tag{3.24}$$

where E_n is defined in (3.18). The first few $\varphi_n(x)$ (equivalently, $\Psi_n(x, t = 0)$) are plotted in figure 3.1. The corresponding probability densities $|\Psi(x,t)|^2$ are time-independent and are

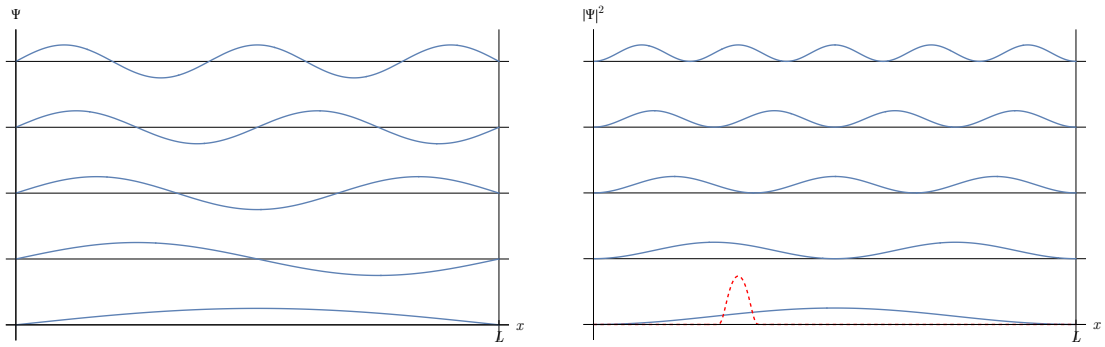


Figure 3.1: Left: the wave-functions φ_n for $n = 1, \cdots, 5$ (from bottom up). Right: the probability densities $|\varphi_n|^2$ for $n = 1, \cdots, 5$ (from bottom up). Red, dashed: the localised classical-like probability density $|\varphi_{\text{loc}}|^2$.

[3]An alternative way to compute $I = \int_0^L \sin^2(kx)dx$ is to observe that $2I = \int_0^L (\sin^2(kx) + \cos^2(kx))dx = L$ and so $I = L/2$.

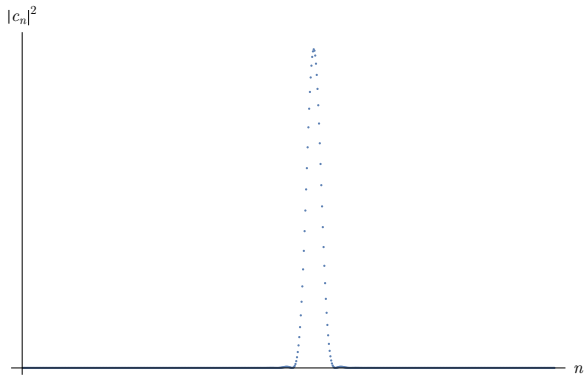


Figure 3.2: Typical behavior of $|c_n|^2$ for a wave-function well-localised in energy.

given by

$$|\Psi_n(x,t)|^2 = |\varphi_n(x)|^2 = \frac{2}{L}\sin^2(\tfrac{\pi n}{L}x). \tag{3.25}$$

The first few are plotted in figure 3.1. These look nothing like the classical expectation of a particle bouncing back and forth. So what's going on? What about the correspondence principle[4]?

The wave-functions in figure 3.1 (which are each energy eigenfunctions with precise energy eigenvalues for $\hat{H}$) have time-independent probability densities do not look like that of a classical particle. As we saw when we introduced plane waves as a candidate for the matter wave for a free particle, their probability density is not localised about a point. A wave-function for a particle (localised about some point) does not have a definite energy, i.e. it is not equal to a single energy eigenfunction, but a linear sum of them.

Let $\varphi_{\rm loc}(x)$ be a wave-function with a localised probability density $|\varphi_{\rm loc}(x)|^2$ as shown by the red, dashed line in the bottom graph on the right of figure 3.1, similar to the states that we studied in section 2.3. We can express $\varphi_{\rm loc}(x)$ as a sum of eigenfunctions of $\hat{H}$ that we found above[5],

$$\varphi_{\rm loc}(x) = \sum_{n=1}^{\infty} c_n \varphi_n(x) \tag{3.26}$$

for some coefficients c_n. Typical values of the coefficients c_n are shown in figure 3.2: they may be peaked sharply around some n, but they are non-vanishing for many n. We interpret these coefficients as follows: $|c_n|^2$ gives the probability that when the energy of the particle is measured, the outcome will be E_n as given in (3.18). We say that the state $\varphi_{\rm loc}$ is a *superposition* of many energy eigenstates φ_n.

[4]See problem set 3 for a problem which addresses this question.

[5]The energy eigenfunctions here form a complete basis for the wave-function, we have not proved this yet, but we make use of this fact here.

Let us note how one can use the decomposition (3.26) to understand how $\varphi_{\text{loc}}(x)$ evolves in time. Say we have $\Psi(x, t = 0) = \varphi_{\text{loc}}(x)$, what is $\Psi(x, t)$ for general t? Since we know how each φ_n evolves in time, see (3.23), we can use the linearity of the Schrödinger equation (a sum of solutions is a solution) to write

$$\Psi(x,t) = \sum_{n=1}^{\infty} c_n e^{-iE_n t/\hbar} \varphi_n(x). \tag{3.27}$$

This time $|\Psi(x,t)|^2$ will be time-dependent (unlike $|\Psi_n(x,t)|^2$) because we cannot simply factor out the phases $e^{-iE_n t/\hbar}$ from all terms simultaneously. In fact, $\Psi(x,t)$ will continue to be localised as $\Psi(x, t = 0) = \varphi_{\text{loc}}(x)$, just as we found for the free-particle wave-function and its maximum will bounce around in the box just like a classical particle with speed $\frac{p}{m}$ as indicated in (2.26), but the width of the Gaussian wave-packet will spread out until eventually the wave-function spreads out to fill the whole box (which may happen slowly or quickly depending on situation, see section 2.4).

3.3 The Free Particle Again

It may seem surprising that we discuss the free particle now, but it is worth revisiting in the context of the time-independent Schrödinger equation. In the free case, in one dimension, we have $V(x) = 0$ and we aim to solve

$$-\frac{\hbar^2}{2m}\frac{\partial^2 \varphi}{\partial x^2} = E\varphi \tag{3.28}$$

for all $x \in \mathbb{R}$. The solution has the form

$$\varphi(x) = Ae^{ikx} + Be^{-ikx} \tag{3.29}$$

where $k = \frac{\sqrt{2mE}}{\hbar}$,but now we run into some problems: φ is not normalisable, and there are no boundary conditions which we can use to eliminate the constants A and B in favour of E. If we form the time-dependent solution we have

$$\Psi(x,t) = Ae^{i(kx - Et/\hbar)} + Be^{-i(kx + Et/\hbar)} \tag{3.30}$$

$$= Ae^{ik(x - (\hbar k/2m)t)} + Be^{-ik(x + (\hbar k/2m)t)}. \tag{3.31}$$

Note that we might also write this in terms of momentum using $p = \hbar k$.

These are travelling plane wave solutions, the first term travels in the positive x direction with speed $\hbar k/2m = p/2m$ while the second term denotes a wave travelling in the negative x

direction with the same speed. By allowing k to be positive or negative we can combine the two terms, so that

$$\Psi(x,t) = Ae^{ik(x-(\hbar k/2m)t)} \tag{3.32}$$

and for $k > 0$ the solution travels to the right (towards $x = +\infty$), while for $k < 0$ then the solution travels to the left (towards $x = -\infty$). The wavespeed for the quantum wave $p/2m$ is half the speed of a classical particle with momentum $p = mv$. The various problems with these free solutions were addressed when we introduced the free particle, the resolution is to take a continuous sum (an integral) of such plane wave solutions, each term weighted such that the total integral gives a Gaussian wave function, which is normalisable and whose peak travels at the classical particle speed.

At $t = 0$ we might write this sum as[6]

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \int_{k=-\infty}^{\infty} dk\, \widetilde{\varphi}(k) e^{ikx} \tag{3.33}$$

where $\widetilde{\varphi}(k)$ are the amplitudes of each plane wave solution (at $t = 0$), and the coefficient $\frac{1}{\sqrt{2\pi}}$ is a convention[7] chosen so the we might see the sum above as the inverse Fourier transform of another function $\widetilde{\Psi}(k)$. With this convention the Fourier transform is written

$$\widetilde{\varphi}(k) = \frac{1}{\sqrt{2\pi}} \int_{x=-\infty}^{\infty} dx\, \varphi(x) e^{-ikx}. \tag{3.34}$$

One can substitute the expression for $\widetilde{\varphi}(k)$ in (3.33) into (3.34) to recover $\varphi(x)$ using

$$\delta(x - x') = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{ik(x-x')} \tag{3.35}$$

where $\delta(x - x')$ is called the Dirac delta function and is an example of a distribution, or generalised function, which is infinite when its argument is zero but zero otherwise, and its integral over all x is one i.e.

$$\int_{-\infty}^{\infty} dx\, \delta(x) = 1 \tag{3.36}$$

so that

$$\int_{-\infty}^{\infty} dx'\, \delta(x - x') f(x') = f(x) \int_{-\infty}^{\infty} dx'\, \delta(x - x') \tag{3.37}$$

$$= f(x) \tag{3.38}$$

[6]The interested reader is encouraged to read through Appendix A for a review of complex waves and the Fourier transform which will be relevant in the discussion that follows.

[7]Conventions differ for the choice of coefficient, the only real constraint being that the product of the coefficients on the Fourier transform and the inverse Fourier transform is $\frac{1}{2\pi}$.

where we have used that $\delta(x-x')$ vanishes everywhere except $x=x'$ in the first equality to allow us to replace the variable $f(x')$ with $f(x)$, and in the second equality we used that the delta function is normalised to one. The Dirac delta function is the continuous analogue for integrals of the Kronecker delta used to pick out terms in discrete sums, it allows one to pick out particular parts of an integral. However the key observation here is that there is no information lost going from $\varphi(x)$ to $\widetilde{\varphi}(k)$ or vice versa using the Fourier transform.

The wave-function $\varphi(x)$ is made up of waves of different wavelengths whose amplitudes are $\widetilde{\varphi}(k)$ and the dispersion relation for the Schrödinger equation, $\omega(k)$ tells us the speed of each constituent wave (as wavespeed $c=k\omega$). In the notation above $\widetilde{\varphi}(k)$ enter the description as amplitudes of plane waves used to make up $\varphi(x)$, but when we realise that we can go back and forth between $\varphi(x)$ in physical space and $\widetilde{\varphi}(k)$ in momentum space[8] and see the symmetry between the two descriptions we may wonder whether $\widetilde{\varphi}(k)$ is another wave-function on a par with $\varphi(x)$. First we think about the normalisation of $\widetilde{\varphi}(k)$

$$\int_{-\infty}^{\infty} dk\,|\widetilde{\varphi}(k)|^2 = \int_{-\infty}^{\infty} dk\,\left(\frac{1}{\sqrt{2}\pi}\int_{x=-\infty}^{\infty} dx\,\varphi(x)e^{-ikx}\right)^*\left(\frac{1}{\sqrt{2}\pi}\int_{x=-\infty}^{\infty} dx'\,\varphi(x')e^{-ikx'}\right) \tag{3.39}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\int_{x=-\infty}^{\infty} dx\,\varphi^*(x)e^{ikx}\int_{x=-\infty}^{\infty} dx'\,\varphi(x')e^{-ikx'} \tag{3.40}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\int_{x=-\infty}^{\infty} dx\,\varphi^*(x)\int_{x=-\infty}^{\infty} dx'\,\varphi(x')e^{ik(x-x')} \tag{3.41}$$

$$= \int_{x=-\infty}^{\infty} dx\,\varphi^*(x)\int_{x=-\infty}^{\infty} dx'\,\varphi(x')\delta(x-x') \tag{3.42}$$

$$= \int_{x=-\infty}^{\infty} dx\,|\varphi(x)|^2. \tag{3.43}$$

This is known as Plancherel's theorem[9], and if $\varphi(x)$ is normalised (or normalisable) then $\widetilde{\varphi}(k)$ is also normalised (or normalisable). This result means that $|\widetilde{\varphi}(k)|^2$ has the key property of being normalisable that we associate with a probability density, and leads us to wonder if $\widetilde{\varphi}(k)$ might be a wave-function. Rather than work with wavenumber we will now make the switch to momentum.

Rewriting the Fourier transform explicitly in terms of $p=k\hbar$ leads to some asymmetry in

[8] We might have said wavenumber space, but as $p=k\hbar$ we choose to refer to momentum space, as in quantum mechanics we will work with the physical quantities energy and momentum and so this is the conventional, preferred name.

[9] It was named for the 1910 work of Michel Plancherel but, as it is a generalisation of Parseval's theorem from 1799, it is often called the Parseval-Plancherel theorem.

the equations, and rather than live with this, we rescale the $\widetilde{\varphi}(k) \to \sqrt{\hbar}\widetilde{\varphi}(k)$ to give

$$\varphi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{p=-\infty}^{\infty} dp\, \widetilde{\varphi}(p) e^{ipx/\hbar} \tag{3.44}$$

$$\widetilde{\varphi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{x=-\infty}^{\infty} dx\, \varphi(x) e^{-ipx/\hbar} \tag{3.45}$$

for the Fourier transforms with

$$\int_{-\infty}^{\infty} dp\, |\widetilde{\varphi}(p)|^2 = \int_{x=-\infty}^{\infty} dx\, |\varphi(x)|^2. \tag{3.46}$$

In (3.44) we can understand the wavefunction $\varphi(x)$ as a sum of momentum eigenstates $e^{ipx/\hbar}$ each of amplitude $\widetilde{\varphi}(p)$ and so we can interpret $|\widetilde{\varphi}(p)|^2$ as a probability density in momentum space (i.e. $\int_{p=p_0}^{p_1} dp |\widetilde{\varphi}(p)|^2$ is the probability of finding the particle described by $\varphi(x)$ having momentum $p \in [p_0, p_1]$).

The Fourier transform has given us a dual picture of our state $\varphi(x)$, now in momentum space as $\widetilde{\varphi}(p)$. In momentum space we understand $e^{-ipx/\hbar}$ as the constituent plane waves at time $t = 0$, but now it is x which for each plane wave is fixed and p which is the variable. By considering the action of simple operators on $e^{-ipx/\hbar}$ we can find the position operator, in momentum space, is $\hat{x} = i\hbar\frac{\partial}{\partial p}$, and we define a momentum operator, in momentum space, as $\hat{p}$ satisfying $\hat{p}\widetilde{\varphi}(p) = p\widetilde{\varphi}(p)$. In momentum space the operators look different to those we have developed in position space, and consequently the Schrödinger equation will also look a little different in momentum space, it becomes:

$$i\hbar\frac{\partial\widetilde{\Psi}(p,t)}{\partial t} = \left[\frac{\hat{p}^2}{2m} + V(i\hbar\frac{\partial}{\partial p})\right]\widetilde{\Psi}(p,t) \tag{3.47}$$

where we have put back in the time dependence i.e. from $\Psi(x,t) = \varphi(x)e^{-iEt/\hbar}$ in (3.44) we have $\widetilde{\Psi}(p,t) = \widetilde{\varphi}(p)e^{-iEt/\hbar}$, i.e. $\Psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} dp \widetilde{\Psi}(p,t)e^{ipx/\hbar}$.

In three dimensions the Fourier transforms that move between position and momentum space are[10]

$$\boxed{\Psi(\boldsymbol{x}) = \frac{1}{(2\pi\hbar)^{\frac{3}{2}}} \int_{\mathbb{R}^3} d^3p\, \widetilde{\Psi}(\boldsymbol{p}) e^{i\boldsymbol{p}\cdot\boldsymbol{x}/\hbar}} \tag{3.48}$$

$$\boxed{\widetilde{\Psi}(\boldsymbol{p}) = \frac{1}{(2\pi\hbar)^{\frac{3}{2}}} \int_{\mathbb{R}^3} d^3x\, \Psi(\boldsymbol{x}) e^{-i\boldsymbol{p}\cdot\boldsymbol{x}/\hbar}} \tag{3.49}$$

[10]Here we use $\Psi(\boldsymbol{x})$ rather than $\varphi(\boldsymbol{x})$ and we understand it is distinguished from $\Psi(\boldsymbol{x},t)$ by the number of arguments, and explicitly $\Psi(\boldsymbol{x}) = \Psi(\boldsymbol{x},0)$ and this motivates the notation.

and the one-dimensional Dirac delta function $\delta(x)$ becomes the Dirac delta function in three-dimensions $\delta^3(\boldsymbol{x})$ with the identity

$$\delta^3(\boldsymbol{x}-\boldsymbol{x'}) = \frac{1}{(2\pi\hbar)^3}\int_{\mathbb{R}^3} d^3p\, e^{i(\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{x'})/\hbar)} \tag{3.50}$$

one can confirm the relations (3.48) and (3.49).

Returning to the free particle, in our earlier discussion of the free particle we have developed a Gaussian function peaked around $x_0(t)$ and with a width $\Delta x(t)$ indicating an uncertainty in the position of the particle about $x_0(t)$. We commented that by careful choice of coefficients $\widetilde{\Psi}(k)$ we can form such a localised wave-packet for $\varphi(x)$ and now we see that the Fourier transform tells us how to extract those coefficients $\widetilde{\varphi}(k)$. Let us develop some instincts for such a wavepacket and the plane waves it is made out of.

Consider the choice

$$\widetilde{\varphi}(k) = Ae^{-(\Delta x)^2(k-k_0)^2} \tag{3.51}$$

where $A = \left(\frac{4(\Delta x)^2}{2\pi}\right)^{\frac{1}{4}}$ so that $\widetilde{\Psi}$ is a normalised, Gaussian function that is symmetric about $k = k_0$ where it peaks. Now from (3.33) we find

$$\varphi(x) = \frac{1}{\sqrt{2\pi}}\int_{k=-\infty}^{\infty} dk\, \left(Ae^{-(\Delta x)^2(k-k_0)^2}\right) e^{ikx}. \tag{3.52}$$

The argument of the exponential above can be written as

$$-(\Delta x)^2(k-k_0)^2 + ikx = -\left(\Delta x(k-k_0) - i\frac{x}{2\Delta x}\right)^2 - \frac{x^2}{4(\Delta x)^2} + ik_0x \tag{3.53}$$

motivating the substitution $l = \Delta x(k-k_0) - i\frac{x}{2\Delta x}$ giving

$$\varphi(x) = \frac{A}{\Delta x\sqrt{2\pi}}\int_{l=-\infty}^{\infty} dl\, e^{-l^2} e^{-\frac{x^2}{4(\Delta x)^2}} e^{ik_0x} \tag{3.54}$$

$$= \frac{A}{\Delta x\sqrt{2}} e^{-\frac{x^2}{4(\Delta x)^2}} e^{ik_0x} \tag{3.55}$$

and we see that a Gaussian $\widetilde{\Psi}(k)$ peaked at k_0 gives rise to a wavefunction $\Psi(x)$ which has a definite initial wavenumber k_0. Putting back in the time-dependence gives

$$\Psi(x,t) = \frac{1}{\sqrt{2\pi}}\int_{k=-\infty}^{\infty} dk\, \widetilde{\varphi}(k) e^{i(kx-\omega t)} \tag{3.56}$$

the contributions where $\widetilde{\Psi}(k)$ is stationary and nearby (in our example this is for $k \approx k_0$) give the dominant contributions to the integral above. However since waves which are out of phase will interfere in order for there to be an overall large contribution to the integral near $k = k_0$

we must have that the phase $kx - \omega t$ must also be stationary at $k = k_0$. This idea is often called the stationary phase principle. When this occurs we find that

$$\frac{d}{dk}(kx - wt)\big|_{k=k_0} = x - \frac{d\omega}{dk}\big|_{k=k_0} t = 0 \tag{3.57}$$

which tells us that the peak of such a wavepacket moves with group velocity $\frac{d\omega}{dk}|_{k=k_0}$. Using $\omega = E/\hbar = \hbar k^2/2m$ we see that $\frac{d\omega}{dk}|_{k=k_0} = \hbar k_0/m = p_0/m$ i.e. the wave-packet has the classical velocity expected for the initial momentum p_0. Notice that in order to solve the problem of normalisability and the particle velocity we necessarily had to adopt a wave-function which leads to uncertainty in position, and via the inverse Fourier transform we also acquire uncertainty in the momentum.

The comments in this section are essentially identical to those made in the section on the free particle, but the observation that $\Psi(k)$ is also Gaussian is new. It is generally true that Gaussian functions are Fourier transformed into other Gaussian functions. We have also seen a general process of obtaining $\Psi(x,t)$ from $\widetilde{\varphi}(k)$: use the inverse Fourier transform on $\widetilde{\varphi}(k)$ to find $\varphi(x) = \Psi(x,0)$ and then $\Psi(x,t) = \Psi(x,0)e^{-i\omega t}$.

Following our digression to discuss momentum space and Fourier transforms, let us return to solving the time-independent Schrödinger equation for simple potentials.

3.4 Bound States and Scattering States

Classically energy $E = T + V$ where T is the kinetic energy and V is the potential energy. As $T \geq 0$ then classical dynamics occurs for $E \geq V$. Classically a bound states describes a particle whose kinetic energy is not sufficient to allow it to escape potential barriers, and is trapped in some finite region of space, and unable to escape to infinity. A scattering state is one which can reach infinity.

In quantum mechanics the analysis is more subtle. The time-dependent Schödinger equation (in one dimension) rearranges to read

$$\frac{d^2\varphi}{dx^2} = \frac{2m}{\hbar^2}(V - E)\varphi \equiv \alpha\varphi. \tag{3.58}$$

Classically a scattering state is one for which $E > V$ in a region extending to infinity. In this region the stationary state $\exp(\sqrt{\alpha}x)$ has $\alpha < 0$ and so is oscillating as x varies and is not normalisable. Scattering states will be free states at $\pm\infty$ and so not normalisable. On the other hand bound states will be normalisable states. The picture in quantum mechanics is complicated because quantum states may tunnel through potential barriers, but we will take a bound state to be one for which $E < V$ at $\pm\infty$. In cases where $V = 0$ at $\pm\infty$, a bound state is

one with $E < 0$ while a scattering states is one with $E > 0$. The sign of the eigenvalue E in the time-independent Schrödinger equation will frequently be sufficient to classify a state as either bounded or a scattering state, and we will find the bound states have a finite number of discrete energy eigenvalues, while scattering states have a continuum of energy eigenvalues. At this stage we have seen that the infinite potential well contains an infinite number of bound states labelled by their discrete energy eigenvalues $E_n = \frac{n^2\pi^2\hbar^2}{2mL^2}$ (equation (3.18)) and no scattering states, while the free particle has a wave-function which is not normalisable for a continuum of energy eigenvalues $E = \frac{\hbar^2 k}{2m}$. Now we will look at potentials which admit both bound states and scattering states.

3.5 The Delta-Function Potential Well

Let

$$V(x) = -\alpha\delta(x) \tag{3.59}$$

where $\alpha \in \mathbb{R}^+$ and

$$\delta(x) = \begin{cases} \infty & x = 0 \\ 0 & x \neq 0 \end{cases} \tag{3.60}$$

is the Dirac delta function satisfying $\int_{-\infty}^{\infty} \delta(x) = 1$. The potential well may be visualised as an infinite (negative) spike at $x = 0$ weight by α as in figure (3.3). It seems odd to include α at this stage, but we will see it plays an important role.

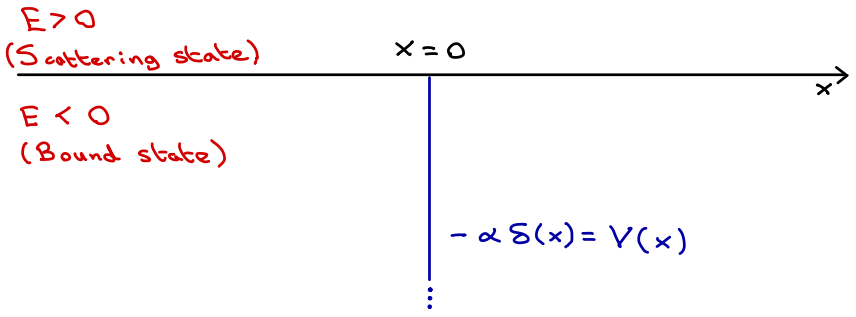


Figure 3.3: The shape of $V(x)$ for a delta function potential well given by (3.59) (blue line), it is an infinite negative spike at $x = 0$ weighted by the real, positive parameter α. States with $E > 0$ are scattering states, while states with $E < 0$ are bound states.

The time-independent Schrödinger equation that we must solve is

$$-\frac{\hbar^2}{2m}\frac{d^2\varphi}{dx^2} - \alpha\delta(x)\varphi = E\varphi. \tag{3.61}$$

We will split our search for solutions into two classes: bound states (for which $E < 0$) and scattering states (for which $E > 0$).

3.5.1 Bound states in the delta-function well

For $E < 0$, in the region $x < 0$ where $V(x) = 0$, (3.61) becomes

$$\frac{d^2\varphi}{dx^2} = -\frac{2mE}{\hbar^2}\varphi \equiv \kappa^2\varphi \tag{3.62}$$

where

$$\kappa = \frac{\sqrt{-2mE}}{\hbar} \tag{3.63}$$

and since $E < 0$ then $-2mE > 0$, and so $\kappa \in \mathbb{R}^+$. The stationary states in this region are

$$\varphi(x) = Ae^{\kappa x} + Be^{-\kappa x} \quad \text{for } x < 0 \tag{3.64}$$

where A and B are constants to be determined. We note that for $\varphi \to 0$ as $x \to -\infty$ we must have that $B = 0$ (as the term $e^{-\kappa x}$ diverges as $x \to -\infty$). Consequently we have

$$\varphi(x) = Ae^{\kappa x} \quad \text{for } x < 0. \tag{3.65}$$

For the region $x > 0$ we find (3.61) is solved by

$$\varphi(x) = Ce^{\kappa x} + De^{-\kappa x} \quad \text{for } x > 0. \tag{3.66}$$

and arguing that $\varphi \to 0$ for $x \to \infty$ leads us to deduce that $C = 0$. Hence, we have,

$$\varphi(x) = De^{-\kappa x} \quad \text{for } x > 0. \tag{3.67}$$

and $\kappa = \frac{\sqrt{-2mE}}{\hbar}$ is unaltered.

In general the properties of the potential $V(x)$ dictate the properties of φ through the time-independent Schrödinger equation. In general φ and $\frac{d\varphi}{dx}$ will be everywhere continuous except where the potential is infinite. Let us look at the implications of the potential becoming infinite at $x = 0$ for the δ-function potential well.

Integrating (3.61) in a region $[-\epsilon, \epsilon]$ about $x = 0$ we have

$$-\frac{\hbar^2}{2m}\int_{-\epsilon}^{\epsilon} dx\, \frac{d^2\varphi}{dx^2} - \alpha\int_{-\epsilon}^{\epsilon} \delta(x)\varphi(x) = E\int_{-\epsilon}^{\epsilon} \varphi. \tag{3.68}$$

then taking the limit $\epsilon \to 0$ we have

$$\lim_{\epsilon\to 0}\left[-\frac{\hbar^2}{2m}\left(\left.\frac{d\varphi}{dx}\right|_{x=\epsilon} - \left.\frac{d\varphi}{dx}\right|_{x=-\epsilon}\right)\right] - \alpha\varphi(0) = 0 \tag{3.69}$$

We see that $\frac{d\varphi}{dx}$ is discontinuous at $x = 0$, it jumps by an amount

$$\Delta(\frac{d\varphi}{dx}) = -\frac{2m\alpha}{\hbar^2}\varphi(0) \tag{3.70}$$

at $x = 0$. If V did not have a delta-function (or worse) we would find that $\frac{d\varphi}{dx}$ is continuous everywhere.

We may wonder if we can now argue that φ is continuous. Suppose that φ was not continuous and suppose it had a finite discontinuity at $x = x_0$, i.e. $\varphi(x)$ makes a finite jump (like a step function) at $x = x_0$. In such a case we would find that $\frac{d\varphi}{dx} \propto \delta(x_0)$, which would in turn imply, $\frac{d^2\varphi}{dx^2} \propto \frac{d}{dx}(\delta(x)) = \delta'(x)$, and so, by the Schrödinger equation, we would have $V(x) \propto \delta'(x)$. So long as we do not consider any potentials containing the derivative of a δ-function then φ will be continuous. We will not consider any such potentials here and so we have the general rules that:

$$\boxed{\begin{array}{l}\varphi(\boldsymbol{x}) \text{ is continuous everywhere, and,} \\ \dfrac{d\varphi}{dx} \text{ is continuous apart from at points where } V \text{ is infinite.}\end{array}} \tag{3.71}$$

Returning to our δ-function potential well, we have that φ must be continuous at $x = 0$, i.e. from (3.65) and (3.67), we have $A = C$ and

$$\varphi(x) = \begin{cases} Ae^{\kappa x} & x \leq 0 \\ Ae^{-\kappa x} & x > 0. \end{cases} \tag{3.72}$$

Then

$$\frac{d\varphi(x)}{dx} = \begin{cases} \kappa Ae^{\kappa x} & x \leq 0 \\ -\kappa Ae^{-\kappa x} & x > 0. \end{cases} \tag{3.73}$$

Hence $\lim_{x\to 0^-}\left[\frac{d\varphi(x)}{dx}\right] = \kappa A$ and $\lim_{x\to 0^+}\left[\frac{d\varphi(x)}{dx}\right] = -\kappa A$ and so we have

$$\begin{aligned}\Delta(\frac{d\varphi}{dx})\Big|_{x=0} &= -2\kappa A \\ &= -\frac{2m\alpha}{\hbar^2}\varphi(0) \\ &= -\frac{2m\alpha}{\hbar^2}A.\end{aligned} \tag{3.74}$$

Hence using (3.63) we find

$$E = -\frac{m\alpha^2}{2\hbar^2}. \tag{3.75}$$

We have found the energy of the single bound state in the δ-function potential well, and we see it depends on α, the coefficient weighting the δ-function potential.

Finally we can normalise the bound state wave-function to determine A:

$$\begin{aligned}\int_{-\infty}^{\infty} dx\,|\varphi(x)|^2 &= \int_{-\infty}^{0} dx\,|A|^2 e^{2\kappa x} + \int_{0}^{\infty} dx\,|A|^2 e^{-2\kappa x} \\ &= 2|A|^2 \int_{-\infty}^{0} dx\, e^{2\kappa x} \\ &= |A|^2/\kappa \\ &= 1 \end{aligned} \tag{3.76}$$

i.e. $A = \sqrt{\kappa} = \frac{\sqrt{m\alpha}}{\hbar}$ and the single bound state, sketched in figure (3.4) has wave-function and energy:

$$\varphi(x) = \frac{\sqrt{m\alpha}}{\hbar} e^{-m\alpha|x|/\hbar^2} \tag{3.77}$$

$$E = -\frac{m\alpha^2}{2\hbar^2} \tag{3.78}$$

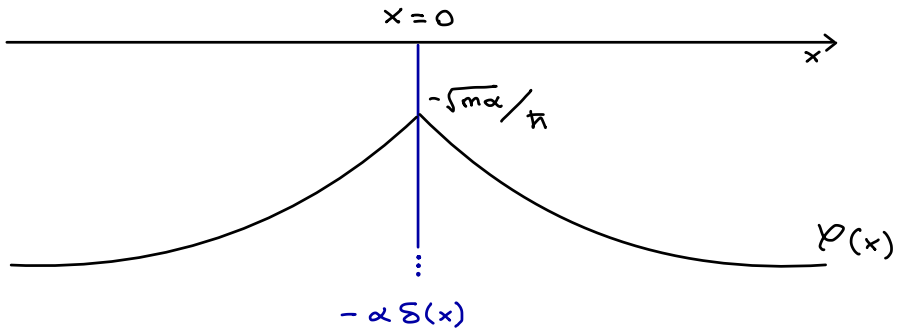


Figure 3.4: The single bound state wave-function for the δ-function potential well.

3.5.2 Scattering states of the delta-function well

To uncover the scattering states we take $E > 0$ in the time-independent Schrödinger equation. In the region $x < 0$ where $V(x) = 0$, (3.61) becomes equation:

$$\frac{d^2\varphi}{dx^2} = -\frac{2mE}{\hbar^2}\varphi \equiv -k^2\varphi \tag{3.79}$$

where

$$k = \frac{\sqrt{2mE}}{\hbar} \tag{3.80}$$

and we emphasise that $E > 0$. We again split the equation into two regions about $x = 0$, and for $x < 0$ we find

$$\varphi(x) = Ae^{ikx} + Be^{-ikx} \quad \text{for } x < 0 \tag{3.81}$$

where A and B are two (new) arbitrary, real constants. Notice that the impact of considering $E > 0$ is to map $\kappa \to ik$, and the wave-function for this energy eigenstate (defined for a single value of $E > 0$) is now oscillating, and not normalisable. There is no condition that we can impose on A or B so that $\phi \to 0$ as $x \to -\infty$. For the region $x > 0$ we have

$$\varphi(x) = Ce^{ikx} + De^{-ikx} \quad \text{for } x > 0 \tag{3.82}$$

where C and D are real constants. Continuity of φ at $x = 0$ implies that

$$A + B = C + D. \tag{3.83}$$

Due to the infinite potential well at $x = 0$ we find the same discontinuity in $\frac{d\varphi}{dx}$ at $x = 0$ as when we solved for the bound state in equation (3.70) which now reads (using φ we have determined for the scattering states so far)

$$\begin{aligned} \Delta(\frac{d\varphi}{dx})\Big|_{x=0} &= -\frac{2m\alpha}{\hbar^2}\varphi(0) \\ &= -\frac{2m\alpha}{\hbar^2}(A + B). \end{aligned} \tag{3.84}$$

While we directly compute that

$$\frac{d\varphi(x)}{dx} = \begin{cases} ikAe^{ikx} - ikBe^{-ikx} & x \leq 0 \\ ikCe^{ikx} - ikDe^{-ikx} & x > 0. \end{cases} \tag{3.85}$$

from where we compute another expression for

$$\begin{aligned} \Delta(\frac{d\varphi}{dx})\Big|_{x=0} &= \frac{d\varphi(x)}{dx}\Big|_{x\to 0^+} - \frac{d\varphi(x)}{dx}\Big|_{x\to 0^-} \\ &= ik(C - D - A + B). \end{aligned} \tag{3.86}$$

Equating (3.84) and (3.86) gives us

$$\begin{aligned} C - D &= A - B + \frac{2im\alpha}{k\hbar^2}(A + B) \\ &\equiv A(1 + 2i\beta) - B(1 - 2i\beta) \end{aligned} \tag{3.87}$$

where we have introduced $\beta \equiv \frac{m\alpha}{k\hbar^2}$. We also know that $A + B = C + D$ from (3.83) . Together with the energy $E > 0$ we have five unknowns and only two equations relating them. We also

know that the states are not normalisable and so we will not get any further way to eliminate the coefficients A, B, C, D from a normalisation condition. What can we do?

Let us pause to think about the scattering states. Each term in the wave functions we have developed is associated with a plane wave at time $t = 0$. If we add in the time dependence to reconstruct $\Phi(x,t) = \phi(t)\varphi(x)$ (recalling that $\phi(t) = \exp(-iEt/\hbar) = \exp(-i\omega t)$ we have:

$$\Psi(x,t) = Ae^{i(kx-\omega t)} + Be^{-i(kx+\omega t)} \quad \text{for } x < 0 \tag{3.88}$$

$$\Psi(x,t) = Ce^{i(kx-\omega t)} + De^{-i(kx+\omega t)} \quad \text{for } x > 0. \tag{3.89}$$

The part of the wave function with coefficient A is a plane wave travelling to the right (from $x = -\infty$ towards $x = 0$), while the part with coefficient B is travelling to the left (from $x = 0$ towards $x = -\infty$). Similarly for $x > 0$ the term with coefficient C is a right-travelling wave (from $x = 0$ towards $x = \infty$), while the part with coefficient D is a left-travelling wave (from $x = \infty$ towards $x = 0$). To simplify our problem we may choose to study a situation where

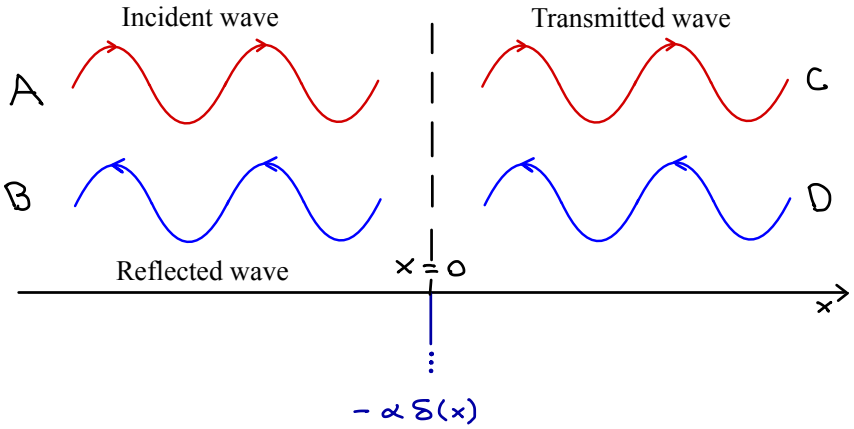


Figure 3.5: Potential scattering solutions for the δ-function potential well. In the sketch above we imagine a scenario in which incident waves are incoming from the left (A), some waves are reflected at $x = 0$ (B), and some are transmitted (C). As there are no waves incoming from the right, we set $D = 0$. If we had considered instead incident from the right but not the left, we would have set $A = 0$ and the roles of reflected and transmitted waves as depicted in the image above would have been swapped.

there are waves incoming only from $x = -\infty$ which then scatter off of the δ-function potential, some waves are reflected (these correspond to the part of the wave-function for $x < 0$ with coefficient B) and some are transmitted (the part of wave-function for $x > 0$ with coefficient C) as sketched in figure (3.5). There are no waves incoming from the right, and we impose

this by setting the appropriate amplitude to zero (i.e. we set $D = 0$)[11]. With $D = 0$ we have $A + B = C$ and so

$$\begin{aligned} C &= A(1 + 2i\beta) - (C - A)(1 - 2i\beta) \\ &= 2A - C(1 - 2i\beta) \end{aligned} \tag{3.90}$$

hence

$$\begin{aligned} C &= \frac{A}{1 - i\beta} \\ B &= \frac{i\beta}{1 - i\beta} A \end{aligned} \tag{3.91}$$

where we have used $B = C - A$ to find the second expression above. Notice that the expressions above, which we refer to collectively as the scattering data, would each have a pole if we allow β to become imaginary. The pole would occur for $\beta = -i = \frac{m\alpha}{k\hbar^2}$ i.e. when $k = \frac{im\alpha}{\hbar^2}$. For this imaginary k we see that $E = -\frac{m\alpha^2}{2\hbar^2}$ which is the energy of the bound state we found in equation (3.78). This is a general feature: poles in the scattering data indicate the existence of bound states.

The scattering states are not normalisable and so we cannot deduce absolute probabilities from the amplitudes A, B and C. Instead we can talk about the relative probabilities[12]:

$$\text{(Reflection probability)} \qquad R \equiv \frac{|B|^2}{|A|^2} = \frac{\beta^2}{1 + \beta^2} \tag{3.92}$$

$$\text{(Transmission probability)} \qquad T \equiv \frac{|C|^2}{|A|^2} = \frac{1}{1 + \beta^2}. \tag{3.93}$$

We note that $R + T = 1$ as we would expect for the probabilities where reflection or transmission are the only two possible outcomes. At this point we give a health warning: these are not yet the full definitions of reflection and transmission probabilities (or coefficients as they are often called), we will soon meet their generalisation - but in cases where there is only a single wavenumber these expressions above are correct. It is interesting to write the reflection and transmission probabilities in terms of the energy E for the scattering state (we recall that $\beta = \frac{m\alpha}{k\hbar^2}$ and $k = \sqrt{2mE}\hbar$):

$$R = \frac{1}{1 + \frac{2E\hbar^2}{m\alpha^2}} \tag{3.94}$$

$$T = \frac{1}{1 + \frac{m\alpha^2}{2E\hbar^2}}. \tag{3.95}$$

[11] This is a choice for the situation we are studying, and allows us to remove one of the four coefficients, but we might equally well have considered waves incident from the right and set the amplitude $A = 0$.

[12] By considering relative probabilities we can effectively eliminate another parameter, the incoming wave amplitude, A.

We see that as $E \to \infty$ then $T \to 1$ and $R \to 0$ as we would expect. More interesting, arguably, is the observation that there is a non-zero probability that a free particle might be reflected off of a potential well. It is worth emphasising that this is rather unusual classically, where we might expect a particle to be slowed down by a potential well but not for it to be reflected (we might have expected a reflection off of a potential barrier, but not off a well). One might wonder if this is a consequence of using an unphysical δ-function potential well, but we will see that this feature persists in more reasonable potentials.

3.6 The Finite Square Potential Well

Next we consider the particle in a finite potential well, this is similar to the particle in the infinite potential well but differs in that the walls of the potential well (or box) are no longer hard walls of infinite potential, but are defined by a finite potential:

$$V(x) = \begin{cases} -V_0, & |x| \leq L/2, \\ 0, & |x| > L/2 \end{cases}, \tag{3.96}$$

where $V_0 > 0$ is a real constant. The potential is shown in figure 3.6, it has an infinite gradient at $|x| = L/2$ which may not be the case for physical potentials, but it serves a an idealised potential well. A more realistic potential would avoid having an infinite gradient such as one shown in figure 3.6 by the (orange) dashed line. This potential will admit both bound states

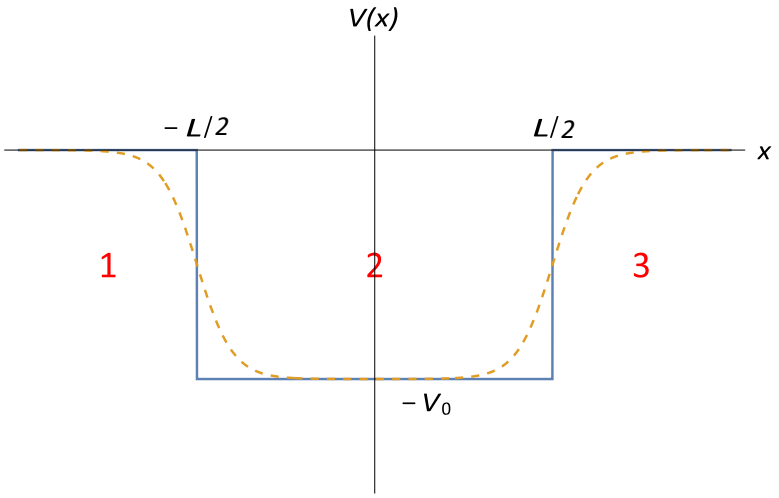


Figure 3.6: The shape of $V(x)$ for a simple potential well given by (3.96) (blue, solid), and a generic $V(x)$ that can be qualitatively modelled by (3.96) (orange, dashed). In the main text we split the range of x into 3 regions, labeled 1 through 3 in this figure.

and scattering states. The potential is zero at infinity and so bound states are those with $E < 0$ while scattering states are those with $E > 0$.

3.6.1 Bound states in a finite square potential well

We aim to solve

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\varphi(x) + V(x)\varphi(x) = E\varphi(x). \tag{3.97}$$

for $E < 0$, but we will also impose that $-V_0 < E$ - we well think about this lower bound on E later, but for now we recognise that it is a sensible condition for dynamics in classical physics.

There are now three regions in which we must solve (3.97), which we label 1 for $x < -L/2$, 2 for $-L/2 < x < L/2$ and 3 for $x > L/2$. In region 1, where $V(x) = 0$. We have

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\varphi(x) = E\varphi(x). \tag{3.98}$$

As $E < 0$ we find

$$\varphi(x) = Ae^{\kappa x} + Be^{-\kappa x}, \qquad \text{for } x < -L/2, \tag{3.99}$$

where

$$\kappa = \frac{\sqrt{-2mE}}{\hbar}. \tag{3.100}$$

Now if $B \neq 0$, then $\varphi(x)$ will grow exponentially as $x \to -\infty$, but we require that $\varphi \to 0$ at $\pm\infty$, and so we must take $B = 0$, then.

$$\varphi(x) = Ae^{\kappa x}. \qquad \text{for } x < -L/2, \tag{3.101}$$

We can make a similar argument in region 3 (where $V(x)$ is also 0) and determine

$$\varphi(x) = De^{-\kappa x}, \qquad \text{for } x > L/2, \tag{3.102}$$

where D is a real constant.

In region 2, $-L/2 < x < L/2$, the equation becomes

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\varphi(x) - V_0\varphi(x) = E\varphi(x), \tag{3.103}$$

or equivalently

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\varphi(x) = (E + V_0)\varphi(x). \tag{3.104}$$

As we have taken $-V_0 < E < 0$ then $E + V_0 > 0$, and so we find

$$\varphi(x) = B\sin(kx) + C\cos(kx), \qquad \text{for } -L/2 \leq x \leq L/2, \tag{3.105}$$

where B and C are arbitrary, real constants and $k = \sqrt{2m(E+V_0)}/\hbar$. Since region 2 is a finite interval, there is no normalisability issue arising from the solution.

The potential is nowhere infinite, and so both $\varphi(x)$ and $\frac{d\varphi}{dx}$ are both continuous. For φ to be continuous at $x = L/2$ and $x = -L/2$ we find

$$Ae^{-\frac{\kappa L}{2}} = -B\sin(\frac{kL}{2}) + C\cos(\frac{kL}{2}) \qquad \text{at } x = -L/2 \tag{3.106}$$

$$B\sin(\frac{kL}{2}) + C\cos(\frac{kL}{2}) = De^{-\frac{\kappa L}{2}} \qquad \text{at } x = L/2. \tag{3.107}$$

The derivative of the wavefunction is

$$\frac{d\varphi}{dx} = \begin{cases} \kappa Ae^{\kappa x} & x < -L/2 \\ kB\cos(kx) - kC\sin(kx) & -L/2 \leq x \leq L/2 \\ -\kappa De^{-\kappa x} & x > L/2. \end{cases} \tag{3.108}$$

Imposing continuity on $\frac{d\varphi}{dx}$ at $x = -L/2$ and $x = L/2$ gives

$$\kappa Ae^{-\frac{\kappa L}{2}} = kB\cos(\frac{kL}{2}) + kC\sin(\frac{kL}{2}) \qquad \text{at } x = -L/2 \tag{3.109}$$

$$kB\cos(\frac{kL}{2}) - kC\sin(\frac{kL}{2}) = -\kappa De^{-\frac{\kappa L}{2}} \qquad \text{at } x = L/2. \tag{3.110}$$

Summarising, we now have

$$\varphi(x) = \begin{cases} Ae^{\kappa x} & x < -L/2 \\ B\sin(kx) + C\cos(kx) & -L/2 \leq x \leq L/2 \\ De^{-\kappa x} & x > L/2. \end{cases} \tag{3.111}$$

with four unknowns, and we also have four equations in (3.106), (3.107), (3.109), and (3.110), so we are in a good position to now eliminate three of the constants, and we have the expectation that normalisation will allow us to fix the final constant. However in order for the work not to become overwhelming we will organise our solutions into even ($\varphi(x) = \varphi(-x)$) and odd ($\varphi(x) = -\varphi(-x)$) solutions. We can see that the solutions will split into these two classes due to the potential being an even function i.e. $V(x) = V(-x)$. This has the consequence that if $\varphi(x)$ solves the time-independent Schrödinger equation, stated in (3.97), then $\varphi(-x)$ does too. Since (3.97) is linear and homogeneous in φ, we can add and subtract solutions, and form new solutions,

$$\varphi_+(x) \equiv \varphi(x) + \varphi(-x) \tag{3.112}$$

and

$$\varphi_-(x) \equiv \varphi(x) - \varphi(x) \tag{3.113}$$

(one of them might be zero, but not both). The first of these is an even solution as $\varphi_+(-x) = \varphi(-x)+\varphi(x) = \varphi_+(x)$, and the second is an odd solution as $\varphi_-(-x) = \varphi(-x)-\varphi(x) = -\varphi_-(x)$. These new solutions have the property that the original solution can be recovered from them,

$$\varphi(x) = \frac{1}{2}\left(\varphi_+(x) + \varphi_-(x)\right), \tag{3.114}$$

This means that, without any loss of generality, we can organise the solutions to the time-independent Schrödinger equation with an even potential $V(x) = V(-x)$ into even and odd wave-functions.

Even solutions

Suppose that the solution that we are after, $\varphi(x)$, is even then we must have $A = D$ and $B = 0$. We need only consider one set of the continuity conditions, as the behaviour of the wave-function at $x = -L/2$ determines its behaviour at $x = L/2$, i.e. it is even and so $\varphi(-L/2) = \varphi(L/2)$ while $\frac{d\varphi}{dx}(x) = -\frac{d\varphi}{dx}(-x)$. The equations (3.106) and (3.109) then simplify to

$$Ae^{\frac{-\kappa L}{2}} = C\cos(\frac{kL}{2}), \tag{3.115}$$

$$\kappa Ae^{-\frac{\kappa L}{2}} = kC\sin(\frac{kL}{2}). \tag{3.116}$$

Taking the ratio of these two equations, we find

$$\kappa = k\tan\tfrac{kL}{2}. \tag{3.117}$$

Recall that $\kappa = \sqrt{-2mE}/\hbar$ and $k = \sqrt{2m(E+V_0)}/\hbar$, which means

$$\kappa^2 = \frac{2mV_0}{\hbar^2} - k^2. \tag{3.118}$$

Using (3.118) in (3.117), we find an equation for k only,

$$\sqrt{\frac{2mV_0}{\hbar^2} - k^2} = k\tan(\tfrac{kL}{2}). \tag{3.119}$$

Solving this equation will allow us to find k (and thus E and κ too). We will discuss how we solve this equation momentarily, after determining the odd solutions. Let us write out the wave-function in terms of a single constant, we will use (3.115) to express A in terms of C and we find

$$\varphi(x) = C\begin{cases}\cos(\frac{kL}{2})e^{\kappa(x+L/2)}, & x < -L/2,\\ \cos(kx), & |x| \le L/2,\\ \cos(\frac{kL}{2})e^{-\kappa(x-L/2)}, & x < -L/2,\end{cases} \tag{3.120}$$

where C will be determined from the normalisation condition for $\varphi(x)$.

Odd solutions

Suppose now that we seek the odd solutions for $\varphi(x)$. In this case we have $A = -D$ and $C = 0$. Now equations (3.106) and (3.109) reduce to

$$Ae^{-\kappa L/2} = -B\sin(\tfrac{kL}{2}), \tag{3.121}$$

$$\kappa Ae^{-\kappa L/2} = kB\cos(\tfrac{kL}{2}). \tag{3.122}$$

Taking the ratio of these two equations, we find

$$\kappa = -k\frac{\cos(\frac{kL}{2})}{\sin(\frac{kL}{2})} = k\tan(\tfrac{kL+\pi}{2}). \tag{3.123}$$

Using (3.118) in (3.123), we find an equation for k only,

$$\sqrt{\frac{2mV_0}{\hbar^2} - k^2} = k\tan(\tfrac{kL+\pi}{2}). \tag{3.124}$$

Using (3.121) to express A in terms of B we find

$$\varphi(x) = B\begin{cases} -\sin(\frac{kL}{2})e^{\kappa(x+L/2)}, & x < -L/2, \\ \sin(kx), & |x| \le L/2, \\ \sin(\frac{kL}{2})e^{-\kappa(x-L/2)}, & x > -L/2, \end{cases} \tag{3.125}$$

where B will be determined using the normalisation condition for $\varphi(x)$. For both cases, even and odd solutions, it remains to understand the equations (3.119) and (3.124) in order to determine k. We can do so using a graphical method. Note that we assumed $-V_0 < E < 0$, which means $0 < \kappa$ and $0 < k < \sqrt{\frac{2mV_0}{\hbar^2}}$. In this interval, the plot of the left-hand side of (3.119) and (3.124) is simply a quarter-circle of radius $\sqrt{\frac{2mV_0}{\hbar^2}}$, see figure 3.7. The plots of the right-hand sides of (3.119) and (3.124) are also shown in figure 3.7 and are V_0-independent.

Each intersection gives a value of k for respectively the even-case or odd-case equations (3.119) or (3.124). As V_0 is increased, the radius of the quarter-circle grows, and more solutions become available, but for any V_0 there are only finitely many solutions.

Since each value of k determines the respective value of E through

$$E = \frac{\hbar^2 k^2}{2m} - V_0, \tag{3.126}$$

we see that there are only finitely many energy levels with $-V_0 < E < 0$. They are schematically shown in figure 3.8. The wave-functions $\varphi(x)$ corresponding to these energy levels are computed by (3.120) and (3.125). Their plots are shown in figure 3.9. As is clear from these plots and

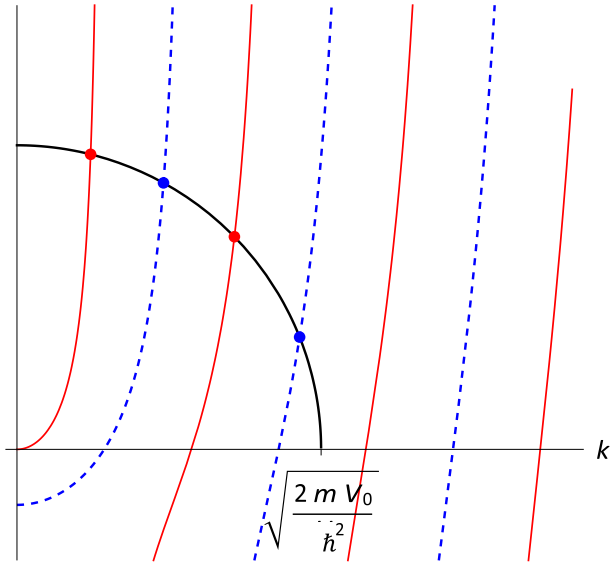


Figure 3.7: A graphical solution of (3.119) and (3.124). The left-hand side of these equations is plotted in solid black (the quarter-circle). The right-hand side of (3.119) is plotted in solid red, and the right-hand side of (3.124) in dashed blue. The red and blue circles mark the even and odd solutions, respectively. The number of solutions depends on the parameters of the problem, and in this case is four.

the explicit formulae (3.120) and (3.125), the wave-functions decay exponentially outside of the well. These states are bound states and although they mostly correspond to the particle being trapped in the potential well there is still a non-zero probability to find it outside the well.

Let us now revisit the assumption that $-V_0 < E < 0$. First, what happens if $E \leq -V_0$? This case corresponds to pure imaginary $k \in i\mathbb{R}$ (including $k = 0$). One can check that we did not have to assume that k is real to arrive at the equations (3.119) or (3.124). Plugging $k = is$ with $s \in \mathbb{R}$ into, for example, (3.119), we find

$$\sqrt{\frac{2mV_0}{\hbar^2} + s^2} = -s \tanh \tfrac{sL}{2}. \tag{3.127}$$

This has no solutions for $s \in \mathbb{R}$ because the left-hand side is strictly positive and the right-hand-side is negative (or zero). Equation (3.124) can be verified in a similar way.

3.6.2 Scattering States in a finite square potential well

We now seek solutions for $E > 0$, these solutions are free at infinity and are called the scattering states. We must solve

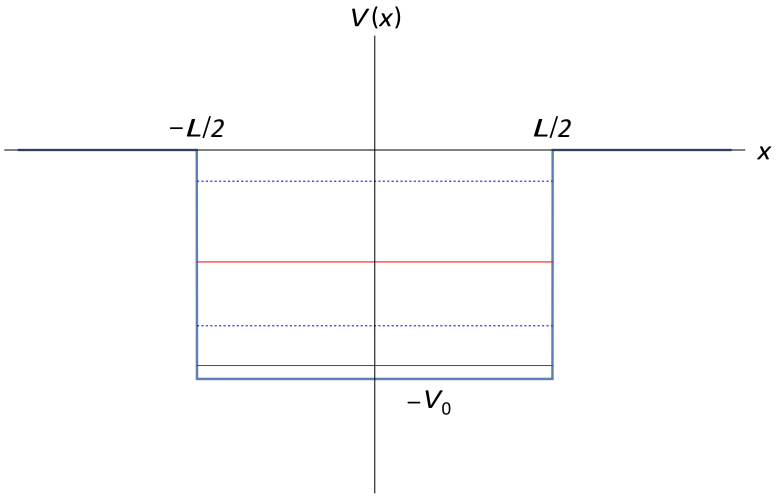


Figure 3.8: Values of the bound state energy levels E shown by horizontal lines in the potential well. Solid red represents even wave-functions, dashed blue – odd wave-functions.

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\varphi(x)+V(x)\varphi(x)=E\varphi(x). \tag{3.128}$$

for $E>0$. We will assume that there is a single incident wave in region 1 ($x<L/2$) and no wave incoming from $x=\infty$ in region 3, (in regions 1 and 3 $V(x)=0$) and so we have

$$\varphi(x)=Ae^{ikx}+Be^{-ikx} \qquad x<-L/2 \tag{3.129}$$

$$\varphi(x)=Fe^{ikx} \qquad x>\ L/2 \tag{3.130}$$

where A, B and F are real constants and

$$k=\frac{\sqrt{2mE}}{\hbar}. \tag{3.131}$$

Inside the potential well (i.e. for $|x|\le L/2$) we must solve

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\varphi(x)=(E+V_0)\varphi(x). \tag{3.132}$$

and we find

$$\varphi(x)=C\sin(k'x)+D\cos(k'x) \qquad |x|\le L/2 \tag{3.133}$$

where C and D are real constants and

$$k'=\frac{\sqrt{2m(E+V_0)}}{\hbar} \tag{3.134}$$

and we emphasise that $E+V_0>0$. Continuity of φ at $x=\pm L/2$ implies that

$$Ae^{-ikL/2}+Be^{ikL/2}=-C\sin(k'L/2)+D\cos(k'L/2) \tag{3.135}$$

$$C\sin(k'L/2)+D\cos(k'L/2)=Fe^{ikL/2}. \tag{3.136}$$

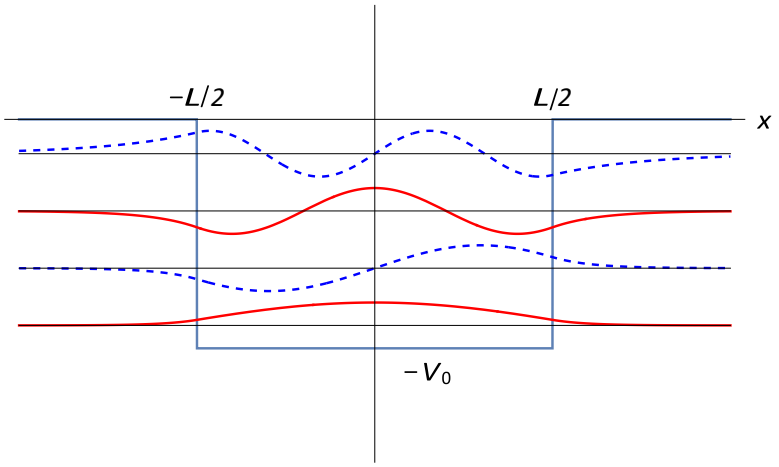


Figure 3.9: Wave-functions of the bound states in a finite potential well. Solid red represents even wave-functions, dashed blue represents odd wave-functions. The energy level is increasing from bottom up.

While continuity of $\frac{d\varphi}{dx}$ at $x = \pm L/2$ gives us

$$ikAe^{-ikL/2} - ikBe^{ikL/2} = k'C\cos(k'L/2) + k'D\sin(k'L/2) \tag{3.137}$$

$$k'C\cos(k'L/2) - k'D\sin(k'L/2) = ikFe^{ikL/2}. \tag{3.138}$$

We can use these four equations to rewrite the scattering data F (the amplitude of the transmitted wave) and B (the amplitude of the reflected wave) in terms of A (the amplitude of the incident wave). We do this so that we can compare the amplitudes of the set of waves at $\pm\infty$, to find the probabilities of reflection by the well potential and transmission past the entire well. Afterwards we will look at the reflection and transmission probabilities at one edge of the well, and refine our thinking about calculating these probabilities.

The computation is a little laborious but we give an outline here. First combine (3.136) and (3.138) to find

$$C = Fe^{ikL/2}\left(\sin(k'L/2) + \frac{ik}{k'}\cos(k'L/2)\right) \tag{3.139}$$

$$D = Fe^{ikL/2}\left(\cos(k'L/2) - \frac{ik}{k'}\sin(k'L/2)\right). \tag{3.140}$$

Combining the above with (3.135) and (3.137) and using the two resulting equations we find

$$F = \frac{Ae^{-ikL}}{\cos(k'L) - \frac{i(k^2+k'^2)}{2kk'}\sin(k'L)} \tag{3.141}$$

$$B = \frac{\frac{i}{2kk'}(k'^2 - k^2)\sin(k'L)Ae^{-ikL}}{\cos(k'L) - \frac{i(k^2+k'^2)}{2kk'}\sin(k'L)}. \tag{3.142}$$

Hence the reflection and transmission probabilities are

$$R = |B|^2/|A|^2 = \frac{\left(\frac{k'^2-k^2}{2kk'}\right)^2 \sin^2(k'L)}{\cos^2(k'L) + \left(\frac{k^2+k'^2}{2kk'}\right)^2 \sin^2(k'L)} \tag{3.143}$$

$$T = |F|^2/|A|^2 = \frac{1}{\cos^2(k'L) + \left(\frac{k^2+k'^2}{2kk'}\right)^2 \sin^2(k'L)} \tag{3.144}$$

and we can check that $R+T=1$. We can use (3.131) and (3.134) to rewrite the reflection and transmission probabilities in terms of E and V_0:

$$R = \frac{V_0^2 \sin^2(\frac{\sqrt{2m(E+V_0)}}{\hbar}L)}{4E(E+V_0) + V_0^2 \sin^2(\frac{\sqrt{2m(E+V_0)}}{\hbar}L)} \tag{3.145}$$

$$T = \frac{4E(E+V_0)}{4E(E+V_0) + V_0^2 \sin^2(\frac{\sqrt{2m(E+V_0)}}{\hbar}L)}. \tag{3.146}$$

Note that as $E \to \infty$, then $T \to 1$ and $R \to 0$, but also notice that when $\frac{\sqrt{2m(E+V_0)}}{\hbar}L = n\pi$ for $n \in \mathbb{Z}$, or equivalently when

$$E + V_0 = \frac{n^2\pi^2\hbar^2}{2L^2m} \tag{3.147}$$

then $T=1$ and $R=0$. The finite square well is invisible to scattering states with this energy E, and, curiously, these are also the same as the energy levels of the infinite potential well found in equation (3.18).

Refining our notion of reflection and transmission probabilities

Our computation of reflection and transmission probabilities have worked well so far, but in the example of the scattering states of the finite potential well we can find an indication that we must refine our definitions.

Consider the far edge at $x = L/2$ of the potential well above. At this boundary there are only three waves which we rewrite as:

$$\varphi(x) = \begin{cases} C'e^{ik'x} + D'e^{-ik'x} & -L/2 < x < L/2 \\ Fe^{ikx} & x > L/2 \end{cases} \tag{3.148}$$

hence compared to our previous coefficients C and D these new amplitudes are

$$C' = \frac{1}{2}(D - iC) \tag{3.149}$$

$$D' = \frac{1}{2}(D + iC). \tag{3.150}$$

We re-write the wave-function inside the potential well in this way to make sense of the notions of incident, reflected and transmitted waves. We will need

$$D \pm iC = Fe^{ikL/2}\left(1 \mp \frac{k}{k'}\right)(\cos(k'L/2) \pm i\sin(k'L/2)) \tag{3.151}$$

which we found using (3.139) and (3.140). The incident wave at $x = L/2$ is $C'e^{ik'x}$ and we may compute the probability ratios to find:

$$\frac{|F|^2}{|C'|^2} = \left(\frac{2k'}{k+k'}\right)^2 \tag{3.152}$$

$$\frac{|D'|^2}{|C'|^2} = \left(\frac{k'-k}{k+k'}\right)^2. \tag{3.153}$$

Something has clearly gone wrong: the probabilities no longer add to one! The problem is in our definition of the reflection and transmission probabilities which have worked well until now. The difference in this case is that the wavenumbers k and k' (and hence momentum) are different on either side of $x = L/2$. Let us look a little closer at what is happening using the probability current in one dimension

$$J = \frac{i\hbar}{2m}\left(\frac{\partial\Psi^*}{\partial x}\Psi - \Psi^*\frac{\partial\Psi}{\partial x}\right) \tag{3.154}$$

which we derived in (2.55) and which we have rewritten here for a wave-function in one dimension with coordinate x. We can find the probability current for each wave either side of $x = L/2$. We will give the various waves generic names, $\varphi_I(x) = C'e^{ik'x}$, $\varphi_R(x) = D'e^{-ik'x}$ and $\varphi_T(x) = Fe^{ikx}$ where the labels are for incident, reflected and transmitted respectively. Then

$$J_I = \frac{\hbar k'}{m}|C'|^2 \tag{3.155}$$

$$J_R = -\frac{\hbar k'}{m}|D'|^2 \tag{3.156}$$

$$J_T = \frac{\hbar k}{m}|F|^2 \tag{3.157}$$

note that because of the reality of the probability current, the time-dependence has dropped out.

The continuity equation for probability (derived in (2.54)) in one dimension is

$$\frac{\partial\rho}{\partial t} + \frac{\partial J}{\partial x} = 0. \tag{3.158}$$

The solutions we have found for the wave-function are stationary states: the probability density $\rho = |\Psi|^2$ is constant in time for these states, and so we must have $\frac{\partial J}{\partial x} = 0$, i.e. J is constant

as a function of x. Further if the probability current took different constant values on either side of any point then there would be a build-up or drop of probability density at that point, as this can't happen (ρ is constant) then we must have that the probability current is balanced on either side of $x = L/2$ and we can check this is the case:

$$J_I + J_R = \frac{\hbar k'}{m}(|C'|^2 - |D'|^2) \tag{3.159}$$

$$= \frac{\hbar k'}{m}\left(\frac{4kk'}{(k+k')^2}\right)|C'|^2 \tag{3.160}$$

$$= \frac{\hbar k'}{m}\left(\frac{k}{k'}\frac{|F|^2}{|C'|^2}\right)|C'|^2 \tag{3.161}$$

$$= J_T. \tag{3.162}$$

It is this computation that gives us the insight to redefine the reflection and transmission probabilities in terms of the probability current. We can see that it contains the relation

$$|C'|^2 - |D'|^2 = \frac{k}{k'}|F|^2 \tag{3.163}$$

or

$$1 = \frac{|D'|^2}{|C'|^2} + \frac{k}{k'}\frac{|F|^2}{|C'|^2} \tag{3.164}$$

and we see it is the missing factor k/k' that we need to make our previous reflection and transmission probabilities sum to one. Consequently, we redefine our definitions of reflection and transmission probabilities[13]

$$\text{(Reflection probability)} \qquad R \equiv \frac{|J_R|}{|J_I|} \tag{3.165}$$

$$\text{(Transmission probability)} \qquad T \equiv \frac{|J_T|}{|J_I|}. \tag{3.166}$$

The expression above apply for any set of incoming, reflected and transmitted wave-functions and their probability currents.

With these definitions we find that at $x = L/2$ we have

$$R = \frac{|D'|^2}{|C'|^2} = \left(\frac{k'-k}{k+k'}\right)^2 \tag{3.167}$$

$$T = \frac{k}{k'}\frac{|F|^2}{|C'|^2} = \left(\frac{4kk'}{(k+k')^2}\right). \tag{3.168}$$

and $R + T = 1$.

[13]It is common to call these the reflection and transmission coefficients.

3.7 The Potential Barrier

Having studied the potential well, we will now look at the potential barrier for a finite potential. We approach this by seemingly repeating the computation for the finite well but now with $V_0 \to -V_0$ so that the potential well becomes a potential barrier. We study the potential

$$V(x) = \begin{cases} V_0, & -L/2 \leq x \leq L/2, \\ 0, & x > L/2 \text{ or } x < -L/2 \end{cases}, \tag{3.169}$$

for $V_0 > 0$¿ This potential is plotted in figure 3.10.

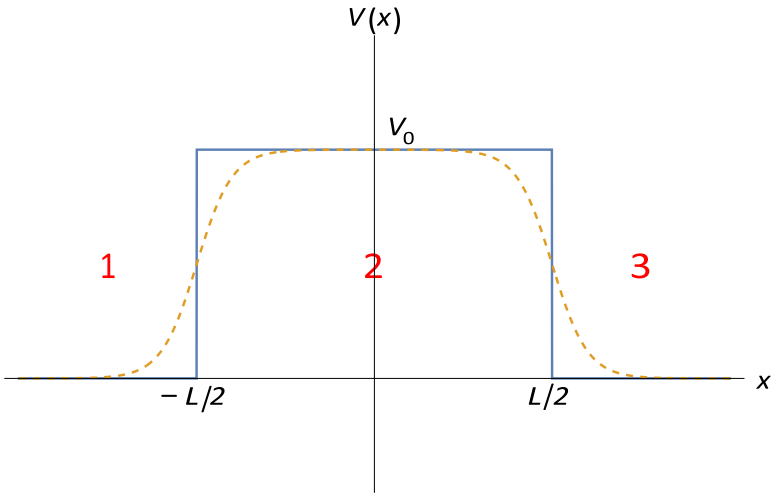


Figure 3.10: The shape of $V(x)$ for a simple potential barrier given by (3.169) (blue, solid). In orange (dashed) we sketch a generic smooth potential barrier, which is a more realistic potential. Our idealised finite potential well is a useful toy model for investigating the realistic potentials. We split the range of x into three regions, labelled 1, 2 and 3 in the figure.

A classical particle incoming from the left would need to have $E > V_0$ in order to pass through it. If the energy is $0 < E < V_0$, then the classical particle would get reflected.

As in the case of the potential well, we solve the Schrödinger equation in the three regions separately (as labelled in figure 3.10). In regions 1 and 3, where $V = 0$, we have

$$-\frac{\hbar^2}{2m}\frac{\partial \varphi(x)}{\partial x^2} = E\varphi(x), \qquad |x| > L/2. \tag{3.170}$$

As before, we will consider an incident wave only from the left, so that, the equation is solved by

$$\varphi(x) = \begin{cases} Ae^{ikx} + Be^{-ikx} & x < -L/2 \\ Fe^{ikx} & x > L/2 \end{cases} \tag{3.171}$$

where

$$k = \frac{\sqrt{2mE}}{\hbar}. \tag{3.172}$$

These scattering states are not normalisable, so we cannot further determine the coefficients.

In region 2, we aim to solve

$$-\frac{\hbar^2}{2m}\frac{\partial\varphi(x)}{\partial x^2} + V_0\varphi(x) = E\varphi(x), \qquad |x| \leq L/2, \tag{3.173}$$

which is equivalent to

$$\frac{\partial\varphi(x)}{\partial x^2} = -\frac{2m(E-V_0)}{\hbar^2}\varphi(x), \qquad |x| \leq L/2. \tag{3.174}$$

For a full solution we ought now to consider two cases: (1) $0 \leq E < V_0$, and (2) $V_0 \leq E$, but we will restrict our enquiry to the first case. The second case can be solved in the usual way[14] to find scattering states in all three regions, but will not reveal any new phenomena to us.

Taking $0 \leq E < V_0$ we solve the equation above to find

$$\varphi(x) = Ce^{\kappa x} + De^{-\kappa x}, \qquad |x| \leq L/2, \tag{3.175}$$

where

$$\kappa = \frac{\sqrt{2m(V_0-E)}}{\hbar}. \tag{3.176}$$

Imposing that $\varphi(x)$ is continuous at $x = \pm L/2$ gives us

$$Ae^{-ikL/2} + Be^{ikL/2} = Ce^{-\kappa L/2} + De^{\kappa L/2} \tag{3.177}$$

$$Ce^{\kappa L/2} + De^{-\kappa L/2} = Fe^{ikL/2}. \tag{3.178}$$

Next we compute

$$\frac{d\varphi}{dx} = \begin{cases} ikAe^{ikx} - ikBe^{-ikx} & x < -L/2 \\ \kappa Ce^{\kappa x} - \kappa De^{-\kappa x} & |x| \leq L/2 \\ ikFe^{ikx} & x > L/2 \end{cases} \tag{3.179}$$

and impose the condition that $\frac{d\varphi}{dx}$ is continuous (this is valid because the potential is nowhere infinite in this example) to obtain the relations

$$ikAe^{-ikL/2} - ikBe^{ikL/2} = \kappa Ce^{-\kappa L/2} - \kappa De^{\kappa L/2} \tag{3.180}$$

$$\kappa Ce^{\kappa L/2} - \kappa De^{-\kappa L/2} = ikFe^{ikL/2}. \tag{3.181}$$

[14] I.e. Solve the time-independent Schrödinger equation in each of the three regions separately, impose continuity of φ and $\frac{d\varphi}{dx}$, and then restrict our interest to the case of a single incident wave.

In order to think about scattering and reflection probabilities we should write the amplitudes B, C, D and F in terms of A, the amplitude of our incident wave. The problem is identical to that of the finite potential well[15], but now with $\kappa = ik'$. As before, one uses (3.178) and (3.181) to eliminate D giving C in terms of F, and then using (3.178) again to find D in terms of F:

$$C = F\left(\frac{\kappa + ik}{2\kappa}\right)e^{(ik-\kappa)L/2} \tag{3.182}$$

$$D = F\left(\frac{\kappa - ik}{2\kappa}\right)e^{(ik+\kappa)L/2}. \tag{3.183}$$

Using (3.177) and (3.180), together with (3.182) and (3.183), we find,

$$Ae^{-ikL/2} + Be^{ikL/2} = F\left(\frac{\kappa + ik}{2\kappa}\right)e^{(ik-2\kappa)L/2} + F\left(\frac{\kappa - ik}{2\kappa}\right)e^{(ik+2\kappa)L/2} \tag{3.184}$$

$$ikAe^{-ikL/2} - ikBe^{ikL/2} = F\left(\frac{\kappa + ik}{2}\right)e^{(ik-2\kappa)L/2} - F\left(\frac{\kappa - ik}{2}\right)e^{(ik+2\kappa)L/2}. \tag{3.185}$$

Combining (3.184) and (3.185) gives

$$A = F\frac{1}{4ik\kappa}\left((\kappa + ik)^2e^{-\kappa L} + (k + i\kappa)^2e^{\kappa L}\right)e^{ikL}. \tag{3.186}$$

Observing that

$$(\kappa + ik)^2e^{-\kappa L} + (k + i\kappa)^2e^{\kappa L} = (k^2 - \kappa^2)2\sinh(\kappa L) + 4ik\kappa\cosh(\kappa L) \tag{3.187}$$

we find

$$F = A\left(\frac{2ik\kappa}{(k^2 - \kappa^2)\sinh(\kappa L) + 2ik\kappa\cosh(\kappa L)}\right)e^{-ikL} \tag{3.188}$$

$$B = A\left(\frac{(k^2 + \kappa^2)\sinh(\kappa L)}{(k^2 - \kappa^2)\sinh(\kappa L) + 2ik\kappa\cosh(\kappa L)}\right)e^{-ikL}.. \tag{3.189}$$

Noting that in regions 1 and 3 the waves have the same wavenumber, then we have, for the reflection and transmission probabilities,

$$R = \frac{|B|^2}{|A|^2} = \frac{(k^2 + \kappa^2)^2\sinh^2(\kappa L)}{(k^2 - \kappa^2)\sinh^2(\kappa L) + 4k^2\kappa^2\cosh^2(\kappa L)} \tag{3.190}$$

$$T = \frac{|F|^2}{|A|^2} = \frac{4k^2\kappa^2}{(k^2 - \kappa^2)\sinh^2(\kappa L) + 4k^2\kappa^2\cosh^2(\kappa L)}. \tag{3.191}$$

[15] And equally laborious. The computations along the way are included here in case you wish to check your working, but to follow the ideas one only needs the final expression stating F and B in terms of A.

Then observing that[16]

$$(k^2+\kappa^2)^2\sinh^2(\kappa L)+4k^2\kappa^2=(k^2-\kappa^2)^2\sinh^2(\kappa L)+4k^2\kappa^2\cosh^2(\kappa L) \tag{3.192}$$

we find that $R+T=1$ as expected.

By consdering

$$T^{-1}=\left(\frac{k^2+\kappa^2}{2k\kappa}\right)^2\sinh^2(\kappa L)+1 \tag{3.193}$$

and recalling that k,κ,L are all positive, we see that as $L\to\infty$ then $T^{-1}\to\infty$ and so $T\to 0$ as expected. But we also realise that for finite L the transmission probability is greater than zero. There is a non-zero probability of the wave-function being transmitted through the potential barrier, even though $E<V_0$. This is a purely quantum phenomenon called quantum tunnelling: the wave-function is able to reach the classically forbidden region (and there is a non-zero probability for this to occur). Quantum tunnelling has a very large impact on the world around us, it allows atomic nuclei to overcome the repulsion due to electrostatic forces and tunnel through the Coulomb barrier which enables nuclear fusion reactions such as those which power the Sun. It also plays a crucial role in alpha decay in which a nucleus emits an alpha particle (or helium nucleus), the forces which bind the nucleus together present a potential barrier and it is quantum tunnelling through this barrier which allows alpha particle to escape. Quantum tunnelling plays a crucial role in modern electronics, from semiconductors and transistors, to flash memory storage and even quantum computing, the modern world we know simply wouldn't have existed without quantum tunnelling.

For completeness we now write the full time-dependent solution for a given E

$$\Psi(x,t)=e^{-iEt/\hbar}\begin{cases}Ae^{ix\sqrt{2mE}/\hbar}+Be^{-ix\sqrt{2mE}/\hbar}, & x<-L/2\\ Ce^{x\sqrt{2m(V_0-E)}/\hbar}+De^{-x\sqrt{2m(V_0-E)}/\hbar}, & -L/2<x<L/2\\ Fe^{ix\sqrt{2mE}/\hbar}, & x>L/2\end{cases} \tag{3.194}$$

these expressions are cumbersome, so it is slightly simpler (and commonplace) to rewrite them in terms of the momentum of the incoming free particle using $E=p^2/2m$:

$$\Psi(x,t)=e^{-i(p^2/2m\hbar)t}\begin{cases}Ae^{ixp/\hbar}+Be^{-ixp/\hbar}, & x<-L/2\\ Ce^{x(\sqrt{2mV_0}-p)/\hbar}+De^{-x(\sqrt{2mV_0}-p)/\hbar}, & -L/2<x<L/2\\ Fe^{ixp/\hbar}, & x>L/2.\end{cases} \tag{3.195}$$

[16] Use $\cosh^2\alpha-\sinh^2\alpha=1$.

We should emphasise that each of the coefficients is related to A by functions of k (via our equations (3.182), (3.183), (3.189), (3.188)), or equivalently the momentum p or the energy E, and it is normal to see the coefficients with explicit k or p dependence indicated.

We have plotted the wave-functions for a couple of representative energies in figure 3.11. Since the wave-functions are intrinsically complex, this plot shows the real and imaginary parts of the wave-functions (3.194) (at $t = 0$) for energies $E = \frac{1}{2}V_0$ and $E = 2V_0$. We did not go through the case where $E > V_0$ but the sketch of the wave-functions in figure 3.11 should not surprise you.

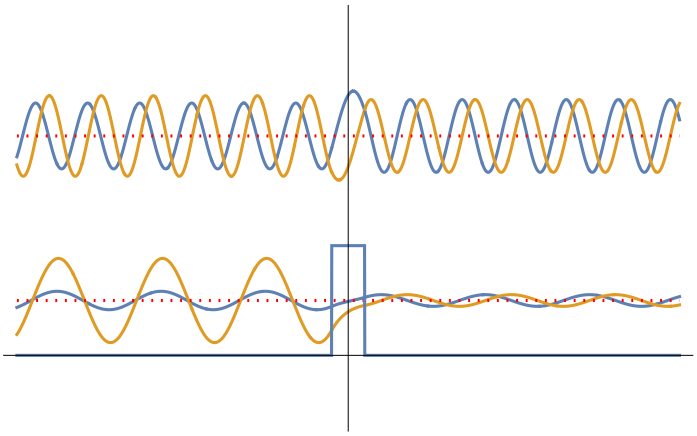


Figure 3.11: Wave-functions in the presence of a barrier of height V_0. The dotted red lines are the energies and examples are shown for $E = \frac{1}{2}V_0$ (which exhibit quantum tunnelling through the potential barrier) and $E = 2V_0$, the blue and yellow lines are the real and imaginary parts of the wave-functions (centred on the energy of the state).

The scattering wave-functions that we have worked with are stationary states of the time-independent Schrödinger equation, or in other words they are energy eigenfunctions of the Hamiltonian operator for a particular energy E. We have met such states several times so far, in particular when we considered the free particle. We were at pains to emphasise that the travelling wave does is not a wave-function of a free particle: it is neither localised nor is it normalisable. The remedy was to form a Gaussian wavepacket from an integral of the stationary states, which was both localised and normalisable. Yet here and earlier when we have discussed transmission and reflection probabilities we have restricted our discussion to the stationary states. What we ought to do is instead of taking a plane wave as our incident wave-function, we should instead form a wave-packet corresponding to a free particle from the stationary states. Once we have done that we should return to thinking about quantum tunnelling of particles. It is not an obvious statement, but we can complete this computation, i.e. we can use the stationary states as a basis to form a Gaussian wave-packet. We will study this idea once we

meet Hilbert spaces, it is the idea of completeness. But for now, recall the wave-function of the free particle which we argued was expressed as a Gaussian wavepacket. It is formed from an infinite and continuous sum of stationary states, and so contains a range of energies. These are sharply peaked about the classical energy of the particle, but they are spread about this energy. Our transmission and reflection probabilities are functions of the energy, and if we were to repeat our analysis, with all the complication it entails, of the scattering about the finite potential barrier for example, we would find that our transmission and reflection probabilities are affected by the range of energies in the incoming Gaussian wavepacket. However while our results we have found won't be the precise results for a Gaussian wavepacket centred about the energy E to be reflected or transmitted, they will be roughly correct as the Gaussian wavepacket is dominated by the stationary states of energy E and the energies nearby. Consequently we should think of the various reflection and transmission probabilities we have worked out for stationary states, as being only approximately correct for a free particle with a Gaussian wavepacket as its wavefunction.

4. A Mathematical Framework for Quantum Mechanics

In the work so far we have looked at the Schrödinger equation, and we have developed a set of ideas that mirror concepts in linear algebra: We have emphasised that we can take linear

Linear algebra	**Quantum mechanics**
Vector	Wave-function
Matrix	Linear Operator
Inner product (to compute $\|\boldsymbol{v}\|^2$)	Integral (to compute $\|\Psi\|^2$)
Eigenvector	Eigenfunction
Vector Space	Hilbert Space

Table 4.1: Analogous concepts in linear algebra and quantum mechanics.

combinations of solutions to the Schrödinger equation to find new solutions. We have also seen in several constructions how we can use certain stationary wave-functions as basis states to construct other solutions, in particular the Gaussian wavepacket of the free particle. We have been concerned with the normalisation of wave-functions, and we have referred to stationary states as eigenfunctions of the Hamiltonian operator. Finally from the first discussions of quantum mechanics we have been making use of linear operators (as differential operators) which act on wave-functions. There is so much similarity with vector spaces and linear algebra that we are impelled to develop the analogy into a sturdy formalism, roughly outlined in table (4.1), and the setting for the objects in quantum mechanics is called a Hilbert space.

In the case of quantum mechanics, the wave-functions will be the elements of the vector space, and we will be interested in finding a complete set of functions which form a basis for the space. The space should be equipped with an inner product, so that we can normalise our wave-functions, and linear operators should act on states in that space to extract physical

information.

In this chapter[1] we will define such a space, the Hilbert space, and transfer much of our thinking about quantum mechanics to this setting.

4.1 Hilbert spaces

Definition 4.1.1. A Hilbert space, $\mathcal{H}$, is a vector space equipped with an inner product that defines a distance function for which the space is a complete metric space.

Comment(s). *(On Hilbert Spaces.)*

- *Hilbert spaces are named[2] after David Hilbert who along with Erhard Schmidt and Friyes Riesz, introduced the spaces in functional analysis at the start of the twentieth century, as an extension of the ideas of Euclidean spaces to functional spaces.*

- *An inner product must satisfy the properties[3]:*

  - $\langle \boldsymbol{u}, \boldsymbol{v} \rangle^* = \langle \boldsymbol{v}, \boldsymbol{u} \rangle$ *(conjugate symmetry),*
  - $\langle \boldsymbol{u}, \alpha \boldsymbol{v} \rangle = \alpha \langle \boldsymbol{u}, \boldsymbol{v} \rangle$ *(linearity in the second entry), and*
  - $\langle \boldsymbol{v}, \boldsymbol{v} \rangle \geq 0$ *and* $\langle \boldsymbol{v}, \boldsymbol{v} \rangle = 0$ *iff* $\boldsymbol{v} = \boldsymbol{0}$.

- *A norm must satisfy the following properties:*

  - $||\boldsymbol{v}|| \geq 0$ *and* $||\boldsymbol{v}|| = 0$ *iff* $\boldsymbol{v} = \boldsymbol{0}$,
  - $||\alpha \boldsymbol{v}|| = |\alpha| ||\boldsymbol{v}||$, *and*
  - $||\boldsymbol{u} + \boldsymbol{v}|| \leq ||\boldsymbol{u}|| + ||\boldsymbol{v}||$.

[1] This chapter will attempt to an introduction to several topics which are properly met as part of a subject called functional analysis. Our aim here being to introduce quantum mechanics we will not derive all the analytical results, nor expect the reader to follow all details and comments. At first pass, the reader may wish to skip over some of the technical comments here. However, one should aim to understand what a Hilbert space is, why it is relevant for quantum mechanics, and be familiar with the two main examples used in quantum mechanics $\mathbb{C}^N$ and $L^2(\mathbb{R})$.

[2] By John von Neumann, who following the rapid development of quantum mechanics in the 1920s, wrote several papers formalising the key mathematical ideas. The results in his papers were published in the book Mathematical Foundations of Quantum Mechanics in 1932.

[3] We have adopted the conventions whereby the inner product is linear in the second entry, and conjugate linear in the first, one could have adopted the alternative convention i.e. linearity in the first entry and conjugate linear in the second. The inner product we have defined is often called a Hermitian inner product.

- *An inner product may be used to define a norm* $||\boldsymbol{v}|| = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v} \rangle}$ *and a norm may be used to define a distance function on the space as* $D(\boldsymbol{u} - \boldsymbol{v}) = ||\boldsymbol{u} - \boldsymbol{v}||$*. Consequently Hilbert space is always a normed vector space.*

- *The probabilistic interpretation of quantum mechanics means that a wave-function should be normalisable. This evidently requires a norm, what we have not seen so far is why we will need an inner product as well as a norm*[4]*. But we can make a couple of motivating comments here: if we cast our thoughts back to the discussion of expectation values* $\langle O \rangle = \int dx\, \Psi^*(x)(\hat{\mathcal{O}}\Psi)(x)$ *we can see the need to have a structure that might compute products between different states, e.g.* $\hat{\mathcal{O}}\Psi$ *and* Ψ*, rather than just a single state as is used in the norm. We have also had a brief look at the expansion of a state in terms of a (infinite, continuous) sum of stationary states (e.g. the Gaussian wavepacket of stationary states used to describe the free particle) - we know that the wave-function encodes all possible outcomes of a measurement of a system, and if we wished to extract the amplitude squared of a single stationary state we would use an inner product to project out that part of the wave-function and its amplitude. So, for now, let us say an inner product space is on our mathematical wish-list for making physical predictions based on the wave-function of a system.*

- *A Hilbert space may be finite or infinite-dimensional. Completeness means that every Cauchy sequence*[5] *in a space converges, with respect to the norm, to an element in the space. The requirement we aim for in quantum mechanics is that any wave-function can be expressed*[6] *as*

$$\Psi = \lim_{N \to \infty} \left(\sum_{i=1}^{N} \alpha_i \psi_i \right) \tag{4.1}$$

for an independent set of vectors ψ_i*. That is, we can find* α_i *such that* $\sum_{i=1}^{\infty} \alpha_i \psi_i$ *is convergent*[7]*. The completeness of the space means that there is no state of the system*

[4] A space which is complete, and has a norm which is not derived from an inner product, is called a Banach space. So a Hilbert space is a special example of a Banach space (Hilbert spaces $\subset$ Banach spaces). How could you have a normed space, without an inner product? Take the space of continuous functions on an interval $[a, b]$, and use the supremum norm $||f|| = \max_{x \in [a,b]}(|f(x)|)$. A norm induced from an inner product satisfies the parallelogram law: $||\boldsymbol{u} + \boldsymbol{v}||^2 + ||\boldsymbol{u} - \boldsymbol{v}||^2 = 2||\boldsymbol{u}||^2 + 2||\boldsymbol{v}||^2$, and the supremum norm does not satisfy the parallelogram law.

[5] A Cauchy sequence in $\mathbb{R}$ is a sequence of numbers x_1, x_2, x_3, ... such that for any positive real number ϵ there exists $n, m > N$ such that $|x_m - x_n| < \epsilon$. Hence a convergent series in $\mathbb{R}$ converges to a number in $\mathbb{R}$.

[6] To be able to label the basis of an infinite-dimensional space by a countable label, as we have done here, we further require that the space is separable. The only infinite-dimensional Hilbert space we will look at is the L^2 space of square-integrable functions, and this is a separable space.

[7] This implies that the sequence of partial sums $\Omega_1, \Omega_2, \Omega_3, \ldots$ where $\Omega_N = \sum_{i=1}^{N} \alpha_i \psi_i$ forms a Cauchy

that cannot be expressed in terms of a sum of independent wave-functions ψ_i. For a finite-dimensional space, any basis of the space will give the expansion we seek, and consequently every finite dimensional vector space equipped with an inner product is a Hilbert space. For infinite-dimensional spaces we will restrict our interest to the space of square-integrable functions over $\mathbb{R}$.

Example 4.1. *The complex vector space $\mathbb{C}^N$ equipped with the inner product*

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \sum_{i=1}^{N} u_i^* v_i = \boldsymbol{u}^\dagger \boldsymbol{v} \tag{4.2}$$

is a Hilbert space.

Comment(s). *(On $\mathbb{C}^N$ as a Hilbert space.)*

- *The elements of $\mathbb{C}^N$ are vectors $\boldsymbol{v} \in \mathbb{C}^N$ which consist of N complex entries $v_i \in \mathbb{C}$ respectively which satisfy the properties of addition: if $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^N$ then $\boldsymbol{u} + \boldsymbol{v} \in C^N$ and the vector $\boldsymbol{u} + \boldsymbol{v}$ has components $u_i + v_i \in \mathbb{C}$ and vectors in $\mathbb{C}^N$ can be multiplied by a complex scalar $\alpha \in \mathbb{C}$ to give another vector in $\mathbb{C}^N$, $\alpha \boldsymbol{v} \in \mathbb{C}^N$, where the components of $\alpha \boldsymbol{v}$ are $\alpha v_i \in \mathbb{C}$.*

- *The norm derived from the inner product is $|\boldsymbol{v}| = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v} \rangle} \in \mathbb{R}$.*

- *As $\dim(\mathbb{C}^N) = N$ (complex-dimensional, or $2N$ real-dimensional) is finite, and because $\mathbb{C}$ is complete, then $\mathbb{C}^N$ is always complete.*

Definition 4.1.2. $L^2(\mathbb{R})$ is the space of square integrable functions $f : \mathbb{R} \to \mathbb{C}$ i.e. those satisfying

$$||f(x)||^2 = \int_{x=-\infty}^{\infty} dx \, |f(x)|^2 < \infty, \tag{4.3}$$

and where addition and scalar multiplication are defined by

$$(f + g)(x) = f(x) + g(x) \tag{4.4}$$

$$(\alpha f)(x) = \alpha f(x) \tag{4.5}$$

for $f, g \in L^2(\mathbb{R})$ and $\alpha \in \mathbb{C}$.

sequence: $||\Omega_n - \Omega_m|| < \epsilon$ for any pair $n, m > M$, for any $\epsilon \in \mathbb{R}^+$ and some $M \in \mathbb{Z}^+$. In this way, we can see how, up to a certain granularity defined by ϵ, we can treat infinite-dimensional spaces in the same way as finite-dimensional spaces.

Comment(s). *It is not immediately evident that if $f, g \in L^2(\mathbb{R})$ then $f + g \in L^2(\mathbb{R})$. What guarantees that $f + g$ is square-integrable? Consider*

$$\begin{aligned}
\int_{-\infty}^{\infty} dx\, |(f+g)(x)|^2 &= \int_{-\infty}^{\infty} dx\, (f(x)+g(x))^*(f(x)+g(x)) \\
&= \int_{x=-\infty}^{\infty} dx\, \left(|f(x)|^2 + |g(x)|^2 + f(x)^* g(x) + g(x)^* f(x)\right) \\
&\leq \int_{x=-\infty}^{\infty} dx\, \left(|f(x)|^2 + |g(x)|^2 + |f(x)^*||g(x)| + |g(x)^*||f(x)|\right) \\
&= \int_{x=-\infty}^{\infty} dx\, \left(|f(x)|^2 + |g(x)|^2 + 2|f(x)||g(x)|\right) \\
&\leq 2\int_{x=-\infty}^{\infty} dx\, \left(|f(x)|^2 + |g(x)|^2\right) \\
&\leq \infty
\end{aligned} \tag{4.6}$$

where we took the absolute value of the integrand in going from the second to third line, and introduced the inequality as e.g. $f(x) \leq |f(x)|$, and we used $|f(x)^*| = |f(x)|$ in moving from the third to the fourth line. In the final line we made use of the arithmetic-geometric mean inequality[8] *to deduce that $2|f(x)||g(x)| \leq |f(x)|^2 + |g(x)|^2$.*

Example 4.2. *The space $L^2(\mathbb{R})$ equipped with the inner product*

$$\langle f, g\rangle = \int_{x=-\infty}^{\infty} dx\, f(x)^* g(x) \tag{4.7}$$

is a Hilbert space.

Comment(s). *(On $L^2(\mathbb{R})$ as a Hilbert space.)*

- *One property of the inner product is not self-evident: $\langle f, f\rangle \geq 0$ and $\langle f, f\rangle = 0$ iff $f(x) = 0$. However a function that is zero almost everywhere apart from at isolated points would also give $\langle f, f\rangle = 0$. This highlights a technical issue that we have avoided until now, the consequence is that when defining the L^2 space, those functions which are equal almost everywhere but differ only at a countable number of isolated points (more strictly: over a set of measure zero) should be considered identical, i.e. instead of a single function $f(x)$ we are really dealing with an equivalence class of $f(x)$ when we speak of $f \in L^2(\mathbb{R})$. This technicality won't affect us in this module, but if you find it interesting and wish to think more about the idea, you should consider taking a course in measure theory, where many of the ideas we have introduced here will be refined.*

[8] For two real, non-negative numbers a and b, consider $(a-b)^2 \geq 0$, then $a^2+b^2-2ab \geq 0$ and so $a^2+b^2 \geq 2ab$.

- *We can confirm it is finite: using the arithmetic-geometric mean inequality, exactly as we did before, we have* $2|f(x)||g(x)| \leq |f(x)|^2 + |g(x)|^2$*, and so*

$$\begin{aligned}\int_{x=-\infty}^{\infty} dx\, |f(x)^*||g(x)| &\leq \frac{1}{2}\int_{x=-\infty}^{\infty} dx\, (|f(x)|^2 + |g(x)|^2) \\ &\leq \infty \end{aligned} \tag{4.8}$$

and so we can see that $\langle f, g\rangle$ *converges as the integral defining it converges absolutely.*

- *The norm is defined as*

$$||f|| = \sqrt{\langle f, f\rangle} = \sqrt{\int_{-\infty}^{\infty} |f(x)|^2} \tag{4.9}$$

and the triangle inequality follows from the Cauchy-Schwarz inequality. This latter inequality comes from considering

$$\begin{aligned} ||f + \alpha g||^2 &= \langle f + \alpha g, f + \alpha g\rangle \\ &= \langle f, f\rangle + \langle f, \alpha g\rangle + \langle \alpha g, f\rangle + \langle \alpha g, \alpha g\rangle \\ &= \langle f, f\rangle + \alpha\langle f, g\rangle + \alpha^*\langle g, f\rangle + |\alpha|^2\langle g, g\rangle \\ &= \langle f, f\rangle + \langle g, g\rangle|\alpha + \frac{\langle g, f\rangle}{\langle g, g\rangle}|^2 - \frac{|\langle g, f\rangle|^2}{\langle g, g\rangle} \\ &\geq 0 \end{aligned} \tag{4.10}$$

as $||f + \alpha g||^2 \geq 0$*. The expression is minimised when* $\alpha = -\frac{\langle g,f\rangle}{\langle g,g\rangle}$*, hence we have the inequality:*

$$\langle f, f\rangle \geq \frac{|\langle g, f\rangle|^2}{\langle g, g\rangle} \tag{4.11}$$

or

$$\sqrt{\langle f, f\rangle\langle g, g\rangle} \geq |\langle f, g\rangle| \tag{4.12}$$

which is the Cauchy-Schwarz inequality. To show the triangle inequality is satisfied by the norm on $L^2(\mathbb{R})$ *consider*

$$||f + g||^2 = ||f||^2 + ||g||^2 + \langle f, g\rangle + \langle g, f\rangle. \tag{4.13}$$

Now taking the absolute value gives

$$\begin{aligned} ||f + g||^2 &= ||f||^2 + ||g||^2 + 2|\langle f, g\rangle| \\ &\leq ||f||^2 + ||g||^2 + 2\sqrt{\langle f, f\rangle\langle g, g\rangle} \\ &= ||f||^2 + ||g||^2 + 2||f||||g|| \\ &= (||f|| + ||g||)^2 \end{aligned} \tag{4.14}$$

i.e. $||f + g|| \leq ||f|| + ||g||$ *as required.*

- *$L^2(\mathbb{R})$ is a complete, separable space[9]. This has the consequence that it has a countable basis.*

- *The $L^2(\mathbb{R})$ space is evidently a sensible setting for wave-functions which are normalisable, but we have already met the stationary states of the free particle, the eigenfunction of momentum, which are not normalisable. We will return to these states shortly and find a way to make sense of them.*

- *We can find a set of orthonormal functions ψ_i (i.e. satisfying $\langle\psi_i, \psi_i\rangle = 1$ and $\langle\psi_i, \psi_j\rangle = \delta_{ij}$) such that any normalisable wave-function can be expressed as*

$$\Psi = \sum_{i=1}^{\infty} \alpha_i \psi_i \tag{4.15}$$

in exact analogue of our experience in a Euclidean vector space. Moreover $\langle\Psi, \psi_j\rangle = \alpha_i$, the amplitude of the wave-function ψ_i in Ψ.

4.2 Observables

The expectation value of an operator is very naturally written in terms of the inner product of a Hilbert space[10]. If the system is described by the normalised wave-function Ψ then

$$\langle\hat{\mathcal{O}}\rangle = \int dx\, \Psi^*(x)(\hat{\mathcal{O}}\Psi)(x) = \langle\Psi, \hat{\mathcal{O}}\Psi\rangle. \tag{4.16}$$

If the outcome of a measurement is to be a real number then we must have that $\langle\hat{\mathcal{O}}\rangle^* = \langle\hat{\mathcal{O}}\rangle$ implying that

$$\langle\hat{\mathcal{O}}\Psi, \Psi\rangle = \langle\Psi, \hat{\mathcal{O}}\Psi\rangle \tag{4.17}$$

we have defined this property of an operator previously: such an operator is called Hermitian[11]. Real quantities that can be measured are called observables and are represented by Hermitian operators. We saw previously that the Hamiltonian $\hat{H}$ and the momentum operator $\hat{p}$ are both Hermitian operators.

[9] For a proof of this statement, see, for example, Real and Complex Analysis by Walter Rudin.

[10] We will restrict our expressions to one-dimensional quantum systems for notational simplicity, but of course in $\mathbb{R}^3$ the integral should be over all of the three-dimensional space, as we wrote it previously.

[11] It is typical to define a Hermitian operator with respect to two different wave-functions, rather than just one, i.e. $\hat{\mathcal{O}}$ is a Hermitian operator if $\langle\phi, \hat{\mathcal{O}}\varphi\rangle = \langle\hat{\mathcal{O}}\phi, \varphi\rangle$ for all $\phi, \varphi \in \mathcal{H}$. But it can be shown that $\langle\Psi, \hat{\mathcal{O}}\Psi\rangle = \langle\hat{\mathcal{O}}\Psi, \Psi\rangle$ implies $\langle\phi, \hat{\mathcal{O}}\varphi\rangle = \langle\hat{\mathcal{O}}\phi, \varphi\rangle$ by considering for example $\Psi = \phi + \varphi$ and then $\Psi = \phi + i\varphi$.

4.2.1 Hermitian operators and the adjoint operator

In linear algebra we already have a notion of a Hermitian operator in the form of a self-adjoint matrix action on a vector space. Consider the example of $\mathcal{H} = \mathbb{C}^N$, and a matrix operator $\boldsymbol{O}$ on the space, whose action is $\boldsymbol{v} \mapsto \boldsymbol{O}\,\boldsymbol{v}$ by matrix multiplication. The inner product in $\mathbb{C}^N$ can also be written in a way that emphasises the matrix multiplication of a row vector and a column vector as

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \sum_{i=1}^{N} u_i^* v_i = \boldsymbol{u}^\dagger \boldsymbol{v} \tag{4.18}$$

where $\boldsymbol{u}^\dagger = (\boldsymbol{u}^*)^T = (\boldsymbol{u}^T)^*$ is used to denote the complex-conjugate transpose. Now consider the inner product of $\boldsymbol{u}$ and $\boldsymbol{O}\,\boldsymbol{v}$:

$$\begin{aligned} \langle \boldsymbol{u}, \boldsymbol{O}\,\boldsymbol{v} \rangle &= \boldsymbol{u}^\dagger \boldsymbol{O}\,\boldsymbol{v} \\ &= (\boldsymbol{O}^\dagger \boldsymbol{u})^\dagger\, \boldsymbol{v} \\ &= \langle \boldsymbol{O}^\dagger\, \boldsymbol{u}, \boldsymbol{v} \rangle. \end{aligned} \tag{4.19}$$

The operation of taking the complex-conjugate transpose is called the adjoint operator in $\mathbb{C}^N$. A self-adjoint operator in $\mathbb{C}^N$ is a Hermitian operator, in this case $\boldsymbol{O}^\dagger = \boldsymbol{O}$. In the finite-dimensional we also have $\langle \boldsymbol{u}, \boldsymbol{v} \rangle^\dagger = \langle \boldsymbol{u}, \boldsymbol{v} \rangle^* = \langle \boldsymbol{v}, \boldsymbol{u} \rangle$, where we observe that $\langle \boldsymbol{u}, \boldsymbol{v} \rangle^T = \langle \boldsymbol{u}, \boldsymbol{v} \rangle$ as it is a complex number (and not a matrix).

In quantum mechanics the operators we have seen so far have been differential operators. In a finite dimensional Hilbert space, the differential operators can be written as a matrix action with respect to a basis of $\mathcal{H}$, i.e. as each linear differential operator acts to map wave-functions to other wave-functions (in a linear way) we can adopt a basis for the Hilbert space to write the differential operator as a matrix action on the basis states. In this way we can think of our differential operators as matrices in a finite-dimensional Hilbert space. For an infinite-dimensional Hilbert space we cannot expect to write down the matrix representation of an operator, as it would be a square matrix with an infinite number of rows and columns. Nevertheless we define the adjoint operation with the same notation (the dagger) that we use to mean complex-conjugate transpose in finite-dimensional spaces, i.e. we define $\hat{\mathcal{P}}$ to be the adjoint of an operator $\hat{\mathcal{O}}$ if it satisfies

$$\langle \Phi, (\hat{\mathcal{O}}\Psi) \rangle = \langle (\hat{\mathcal{P}}\Phi), \Psi \rangle \tag{4.20}$$

with respect to the inner product of the Hilbert space, for $\Phi, \Psi \in \mathcal{H}$, or in the notation we have adopted: $\hat{\mathcal{P}} = \hat{\mathcal{O}}^\dagger$. We will revisit the adjoint operator when we introduce Dirac's bra-ket notation.

4.2.2 Eigenfunction of Hermitian Operators

The outcome of measuring an observable will have a standard deviation, or uncertainty $\Delta\mathcal{O}$, which we define by

$$\Delta\mathcal{O}^2 = \langle(\hat{\mathcal{O}} - \langle\hat{\mathcal{O}}\rangle)^2\rangle. \tag{4.21}$$

Consider the Hamiltonian operator when the wave-function of the system is a stationary state, that is $\hat{\mathcal{O}} = \hat{H}$ and $\Psi(x,t) = \phi(t)\varphi(x)$ satisfying the time-independent Schrödinger equation $\hat{H}\varphi = E\varphi$. Then the uncertainty of measuring the energy vanishes as

$$\begin{aligned}
\Delta H^2 &= \langle\Psi, (\hat{H} - \langle\hat{H}\rangle)^2\Psi\rangle && (4.22)\\
&= \langle\Psi, (\hat{H} - E)^2\Psi\rangle && (4.23)\\
&= \langle(\hat{H} - E)\Psi, (\hat{H} - E)\Psi\rangle && (4.24)\\
&= 0 && (4.25)
\end{aligned}$$

where we used that $\langle\hat{H}\rangle = \int dx\, \Psi^*\hat{H}\Psi = E$ and that the Hamiltonian operator is a Hermitian operator. Consequently for a stationary state, an eigenfunction of a Hermitian operator, there is zero uncertainty in the measurement of the energy for the state.

This observation generalises to any eigenfunction of an Hermitian operator, e.g. let $\hat{Q}$ be an Hermitian operator and let the wave-function be an eigenfunction of $\hat{\mathcal{O}}$, i.e. $\hat{\mathcal{O}}\Psi = \lambda\Psi$ and repeat the steps above to find that there is no uncertainty in the outcome of measuring the observable associated to the operator $\hat{\mathcal{O}}$. Eigenfunctions of Hermitian operators have no uncertainty in their measurement.

There are two classes of eigenfunction of Hermitian operators: those with discrete eigenvalues, and those with continuous eigenvalues. The set of eignvalues of an operator is called the spectrum of the operator. From our investigation of bound states and scattering states, we are aware that bound states have discrete energy eigenvalues and have normalisable wave-functions while scattering states have continuous eigenvalues and non-normalisable wave-functions. We will split our discussion into these two classes:

Discrete Eigenvalues

Theorem 4.1. *Normalisable eigenfunctions of Hermitian operators have real eigenvalues.*

Proof. Let $\hat{\mathcal{O}}$ be an Hermitian operator with $\mathcal{O}\Psi = \lambda\Psi$. Then

$$\langle\Psi, \hat{\mathcal{O}}\Psi\rangle = \lambda\langle\Psi, \Psi\rangle \tag{4.26}$$

while we also have, by using that $\hat{\mathcal{O}}$ is Hermitian,

$$\begin{aligned}\langle \Psi, \hat{\mathcal{O}}\Psi\rangle &= \langle \hat{\mathcal{O}}\Psi, \Psi\rangle \\ &= \lambda^* \langle \Psi, \Psi\rangle. \end{aligned} \tag{4.27}$$

Equating the two expressions we have $\lambda = \lambda^*$ and so λ is a real eigenvalue. □

Theorem 4.2. *Eigenfunctions of Hermitian operators with distinct eigenvalues are orthogonal.*

Proof. Suppose $\hat{\mathcal{O}}\Psi_1 = \lambda_1\Psi_1$ and $\hat{\mathcal{O}}\Psi_2 = \lambda_2\Psi_2$ with $\lambda_1 \neq \lambda_2$ (and $\hat{\mathcal{O}}$ is Hermitian), then

$$\langle \Psi_1, \hat{\mathcal{O}}\Psi_2\rangle = \lambda_2 \langle \Psi_1, \Psi_2\rangle \tag{4.28}$$

while we also have, by using that $\hat{\mathcal{O}}$ is Hermitian,

$$\begin{aligned}\langle \Psi_1, \hat{\mathcal{O}}\Psi_2\rangle &= \langle \hat{\mathcal{O}}\Psi_1, \Psi_2\rangle \\ &= \lambda_1^* \langle \Psi_1, \Psi_2\rangle \\ &= \lambda_1 \langle \Psi_1, \Psi_2\rangle \end{aligned} \tag{4.29}$$

as eigenvalues of Hermitian operators are real (and so $\lambda_1^* = \lambda_1$). Taking the difference between (4.28) and (4.29) we find

$$(\lambda_2 - \lambda_1)\langle \Psi_1, \Psi_2\rangle = 0 \tag{4.30}$$

and since $\lambda_1 \neq \lambda_2$ then we must have $\langle \Psi_1, \Psi_2\rangle = 0$ and the two eigenfunctions with different eigenvalues are orthogonal. □

The eigenfunctions span the Hilbert space, although it is not an obvious statement in general (although in the finite-dimensional setting of $\mathbb{C}^N$, we can confirm this claim, as a Hermitian matrix has N distinct eigenvectors). We will express this idea formally when we state the spectral theorem.

Continuous Eigenvalues

As we have seen for scattering states, the stationary states are not normalisable and consequently the inner product used to prove the properties of eigenfunctions of Hermitian operators with discrete eigenvalues cannot work in the same way: the inner product ceases to be defined when the state is not normalisable. However we have already developed the tools needed to derive ideas similar to the results for discrete eigenvalues.

Recall that the stationary states for the free particle[12]are

$$\varphi_p(x) = Ae^{ipx/\hbar} \tag{4.31}$$

where here we allow p to be positive or negative. It is, in fact an eigenfunction of the momentum operator[13]:

$$\hat{p}\,\varphi_p(x) = -i\hbar\frac{d}{dx}\varphi_p(x) = p\varphi_p(x). \tag{4.32}$$

The spectrum of $\hat{p}$ is continuous, the eigenvalues p may be any real number, and conversely for each real number p there is an eigenstate $\varphi_p(x)$. We saw in (3.35) that there is a sense in which the Dirac delta function plays the role of the Kronecker delta function, and it is this notion that we can use to give meaning to orthogonality here.

Now we have, using the $L^2(\mathbb{R})$, inner product:

$$\langle\varphi_{p'}(x), \varphi_p(x)\rangle = \int_{-\infty}^{\infty} dx\,(Ae^{ip'x/\hbar})^* Ae^{ipx/\hbar} \tag{4.33}$$

$$= |A|^2 \int_{-\infty}^{\infty} dx\, e^{ix(p-p')} \tag{4.34}$$

$$= |A|^2 2\pi\hbar\,\delta(p-p') \tag{4.35}$$

where in the last line we have used (3.35) with a change of variables (and also made use of $p = \hbar k$). If we take $A = \frac{1}{\sqrt{2\pi\hbar}}$ then we may say that the continuum of eigenstates of $\hat{p}$

$$\varphi_p(x) = \frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar} \tag{4.36}$$

satisfy a generalised orthonormality condition:

$$\langle\varphi_{p'}(x), \varphi_p(x)\rangle = \delta(p-p'). \tag{4.37}$$

The eigenfunctions are again complete, since we know through the Fourier transform that we can write any square integrable function as an integral of the form

$$\varphi(x) = \int_{-\infty}^{\infty} dp\, C(p)\varphi_p(x) = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} dp\, C(p)e^{ipx/\hbar} \tag{4.38}$$

where we have taken $C(p)$ as the continuous coefficient in the expansion, and we can find the values of $C(p)$ for a given $\varphi(x)$ by taking the inner product and making use of our generalised

[12]See section (3.3) ('The Free Particle Again') for the full details.

[13]We showed previously that $\hat{p}$ is a Hermitian operator, i.e. $\hat{p}^\dagger = \hat{p}$, it is worth reminding ourselves here: $\langle\phi, \hat{p}\psi\rangle = \int_{\mathbb{R}} dx\,\phi^*(-i\hbar)\frac{d\psi}{dx} = \int_{\mathbb{R}} dx\,(i\hbar)\frac{d\phi^*}{dx}\psi = \int_{\mathbb{R}} dx\,(\hat{p}\phi)^*\psi = \langle\hat{p}\phi, \psi\rangle$. We recall also that $\hat{x}$ and so also $\hat{H}$ are all Hermitian, or self-adjoint, operators.

orthogonality:

$$\langle \varphi_{p'}(x), \varphi(x) \rangle = \int_{-\infty}^{\infty} dp\, C(p) \langle \varphi_{p'}(x), \varphi_p(x) \rangle \tag{4.39}$$

$$= \int_{-\infty}^{\infty} dp\, C(p) \delta(p - p') \tag{4.40}$$

$$= C(p'). \tag{4.41}$$

Consequently we can argue that the eigenfunctions with continuous eigenvalues can also be understood to be orthonormal. We can also now use the inner product defined in terms of the delta function to argue that the eigenvalues for Hermitian operators are again real even when there are a continuous set of eigenvalues.

However there is little we can do about the fact the eigenstates are not normalisable, and so are not in the Hilbert space and they individually do not represent possible physical states. Of course, we can use the eigenstates to construct normalisable wave-functions (e.g. the Gaussian wave-packet for the free particle) which are in the Hilbert space.

We may now state the spectral theorem:

Theorem 4.3 (The spectral theorem)**.** *Let $\mathcal{H}$ be a Hilbert space and let $\hat{\mathcal{O}}$ be a Hermitian (self-adjoint) linear operator on $\mathcal{H}$.*

- *The spectrum of $\hat{\mathcal{O}}$ is a subset of $\mathbb{R}$, i.e. all eigenvalues of $\hat{\mathcal{O}}$ are real.*
- *Any two eigenstates of $\hat{\mathcal{O}}$ with distinct eigenvalues are orthogonal to each other.*
- *The eigenstates of $\hat{\mathcal{O}}$ may be used to form a complete basis of the Hilbert space $\mathcal{H}$.*

We emphasise that in the above theorem we no longer distinguish between continuous or discrete eigenvalues, but instead we speak of the full spectrum of an operator. An operator will generally have both a continuous spectrum (for $\lambda \in [\lambda_{\min}, \lambda_{\max}] \subset \mathbb{R}$) and a discrete spectrum (for $\lambda_i \in \mathbb{R}$ and $i \in \mathbb{N}$). For the Hamiltonian, the continuous spectrum corresponded to scattering states and the discrete spectrum to bound states. The statement that all the eigenfunctions form a basis means that for any state $\Psi \in \mathcal{H}$ we have

$$\Psi = \int_{\lambda_{\min}}^{\lambda_{\max}} d\lambda\, C(\lambda) \psi_\lambda + \sum_{i \in \mathbb{N}} C(\lambda_i) \psi_{\lambda_i}. \tag{4.42}$$

It could happen that for each eigenvalue λ there exists more than one eigenstate, i.e. $\hat{\mathcal{O}}\psi_{i,\lambda} = \lambda \psi_{i,\lambda}$ for $i \in \{1, \cdots, n_\lambda\}$. In this case we need n_λ eigenfunctions $\varphi_{i,\lambda}$ labelled by i which satisfy

$\langle \varphi_{i,\lambda}, \varphi_{j,\lambda} \rangle = \delta_{ij}$. In practise a second compatible observable[14] may be used to distinguish between multiple eigenvectors of an observable having the same eigenvalue.

Furthermore if the continuous spectrum is not formed of a simple interval in $\mathbb{R}$ but from a union of intervals then we would need to include one integral, in the expansion of Ψ above, for each interval in the union for the continuous spectrum.

The notation for Ψ and indeed the inner product has become somewhat complicated, being dependent on $\mathcal{H}$, and the properties of the basis of $\mathcal{H}$, and yet the underlying ideas are all very similar. We are now in a position to introduce a universal notation for quantum mechanics.

4.3 Dirac's Bra-Ket Notation

We have introduced two examples of Hilbert spaces ($\mathbb{C}^N$ and $L^2(\mathbb{R})$), in which the wave-function of our quantum mechanical system resides. Furthermore we have seen that we can find a basis for $\mathcal{H}$ formed of eigenstates of a Hermitian (or self-adjoint) operator, however this then introduced different ideas depending upon whether the spectrum was discrete or continuous. While the details of the Hilbert spaces and their different types of bases are quite different, the linear algebra and inner product are almost identical and the idea that eigenstates of self-adjoint operators form a basis for the space is universal. Dirac introduced a notation for quantum mechanics which extracts the central idea and is simple to work with which is called the bra-ket notation[15] .

A wave-function Ψ, or any vector in the Hilbert space, is denoted by a 'ket': $|\Psi\rangle$. Indeed if we have a countable basis for a Hilbert space[16] denoted by ψ_i for $i \in \mathbb{N}$ then a wave-function

$$|\Psi\rangle = \sum_i \alpha_i |\psi_i\rangle. \tag{4.43}$$

We exclude the limits on the summation as the Hilbert space may be finite or infinite-dimensional. For $\mathbb{C}^N$ the statement above is self-evident, but we can make a simple (and partial) argument for the existence of a countable orthonormal basis exists in $L^2(\mathbb{R})$. Consider the Hamiltonian operator for a quadratic potential in x, it is bounded below and at spatial infinity it is infinite. It is a Hermitian operator and the eigenstates of the Hamiltonian will correspond to bound

[14] A pair of compatible observables $\hat{A}$ and $\hat{B}$ commute: $[\hat{A}, \hat{B}] = 0$. This implies that they share eigenvectors but not necessarily eigenvalues. We will meet compatible operators in chapter five.

[15] We will understand the naming as coming from the brackets of the inner product, where a bra and a ket will come together to form a bracket. Don't ask about the 'c'.

[16] If a Hilbert space is separable, this is the case. In this course we are focussed on two classes of Hilbert space relevant to quantum mechanics: $\mathbb{C}^N$ or L^2 spaces. Evidently the basis of $\mathbb{C}^N$ is countable, but it is also the case that L^2 is a separable space and so has a countable basis.

states, and so will have a discrete spectrum. There are no scattering states and so no continuous spectrum. Consequently we can lean on the spectral theorem for this Hermitian operator and say that its eigenstates form an orthonormal basis of $L^2(\mathbb{R})$. The quadratic potential corresponds to the harmonic oscillator, and its eigenfunctions are the Hermite functions, which we will study later. For now we leave it as a comment: the Hermite functions form a countable basis of $L^2(\mathbb{R})$.

Given a Hermitian operator on $\mathcal{H}$ we can find a set of orthonormal (in the generalised sense) eigenstates[17] which we can use to expand Ψ as in equation (4.42). For a state ψ_λ which is an eigenstate of a Hermitian operator $\hat{\mathcal{O}}$ is customary to write the eigenstate just using the eigenvalue as the label on the ket:

$$\psi_\lambda \mapsto |\lambda\rangle \tag{4.44}$$

so that

$$\hat{\mathcal{O}}\,\psi_\lambda = \lambda\,\psi_\lambda \mapsto \hat{\mathcal{O}}|\lambda\rangle = \lambda|\lambda\rangle. \tag{4.45}$$

A 'bra' will denote a continuous linear functional on a Hilbert space $\phi : \mathcal{H} \to \mathbb{C}$. A continuous linear functional ϕ on $\mathcal{H}$ is a (continuous) function that satisfies

$$\phi(\psi_1 + \psi_2) = \phi(\psi_1) + \phi(\psi_2) \tag{4.46}$$

$$\phi(\alpha\psi) = \alpha\phi(\psi) \tag{4.47}$$

where $\alpha \in \mathbb{C}$ and $\psi_1, \psi_2, \psi \in \mathcal{H}$. Any continuous linear functional ϕ on $\mathcal{H}$ is associated with a unique vector in $\mathcal{H}$ given by Φ such that $\phi(\Psi) = \langle\Phi, \Psi\rangle$. This can be seen by writing $\Psi = \sum_i \alpha_i\psi_i$ for an orthonormal basis ψ_i then $\langle\psi_j, \Psi\rangle = \alpha_j$. Now we may write

$$\begin{aligned}
\phi(\Psi) &= \phi(\sum_i \alpha_i\psi_i) \\
&= \phi(\sum_i \langle\psi_i, \Psi\rangle\psi_i) \\
&= \sum_i \langle\psi_i, \Psi\rangle\phi(\psi_i) \\
&= \sum_i \langle\phi(\psi_i)^*\psi_i, \Psi\rangle \\
&= \langle\sum_i \phi(\psi_i)^*\psi_i, \Psi\rangle \\
&\equiv \langle\Phi, \Psi\rangle
\end{aligned} \tag{4.48}$$

[17] We bear in mind the eigenstates need not be in the Hilbert space, we only need that we can find an expansion as in (4.42).

where we used the conjugate linearity of the inner product on $\mathcal{H}$ to bring $\phi(\psi_i)$ inside the left-hand entry of the inner product. The result is that the vector $\Phi \equiv \sum_i \phi(\psi_i)^* \psi_i$ is a vector in $\mathcal{H}$ encoding all the information in the linear functional ϕ. Furthermore the vector Φ associated to the functional ϕ is unique as suppose there was another such vector Φ' then we would have

$$\phi(\Psi) = \langle \Phi, \Psi \rangle = \langle \Phi', \Psi \rangle \tag{4.49}$$

and so $\langle \Phi - \Phi', \Psi \rangle = 0$, but since this implies that $\Phi - \Phi' = \mathbf{0} \in \mathcal{H}$ then the vector is unique. The set of all linear functionals on $\mathcal{H}$ is a vector space, which we call the dual space and is denoted $\mathcal{H}^*$. Indeed using the properties of the inner product on $\mathcal{H}$, we observe that

$$(\phi_1 + \phi_2)(\Psi) = \langle \Phi_1 + \Phi_2, \Psi \rangle = \langle \Phi_1, \Psi \rangle + \langle \Phi_2, \Psi \rangle = \phi_1(\Psi) + \phi_2(\Psi) \tag{4.50}$$

$$(\alpha\phi)(\Psi) = \langle \alpha\Phi, \Psi \rangle = \alpha^* \phi(\Psi). \tag{4.51}$$

The second property above means that we say that the functionals are conjugate linear[18] on $\mathcal{H}$.

Returning to Dirac's bra, we define $\langle \phi |$ to denote the functional ϕ on $\mathcal{H}$, and instead of writing the inner product, we bring the bra and the ket together to form a single bra(c)ket notation $\langle \phi | \psi \rangle$. All in all, the notational changes are:

$$\text{A state in the Hilbert space} \quad \psi \in \mathcal{H} \longrightarrow |\psi\rangle \tag{4.52}$$

$$\text{A continuous linear functional on the Hilbert space} \quad \phi \in \mathcal{H}^* \longrightarrow \langle \phi | \tag{4.53}$$

$$\text{The inner product in } \mathcal{H} \quad \langle \phi, \psi \rangle \longrightarrow \langle \phi | \psi \rangle \tag{4.54}$$

If $|\psi_i\rangle$ are an orthonormal basis for $\mathcal{H}$ then $\langle \psi_i |$ are an orthonormal basis for $\mathcal{H}^*$ and by analogy with $\mathbb{C}^N$ we can see that the adjoint operator takes us from $|\phi_i\rangle$ to $\langle \phi_i | = (|\phi_i\rangle)^\dagger$.

Consider

$$\langle \phi | \equiv \sum_i \langle \phi | \psi_i \rangle \langle \psi_i | \tag{4.55}$$

then

$$\langle \phi | \psi_j \rangle = \sum_i \langle \phi | \psi_i \rangle \langle \psi_i | \psi_j \rangle = \sum_i \langle \phi | \psi_i \rangle \delta_{ij} \tag{4.56}$$

and the action of ϕ specified over the basis fully defines ϕ on $\mathcal{H}$. If we look in the middle of the expression we can read that $\sum_i |\psi_i\rangle\langle\psi_i|$ behaves like the identity operator. Indeed it has an action on any vector in the Hilbert space as

[18] Or sometimes anti-linear.

$$\sum_i |\psi_i\rangle\langle\psi_i|\Psi\rangle = \sum_i |\psi_i\rangle\langle\psi_i \sum_j \alpha_j|\psi_j\rangle \tag{4.57}$$

$$= \sum_i \sum_j \alpha_j|\psi_i\rangle\langle\psi_i|\psi_j\rangle \tag{4.58}$$

$$= \sum_i \alpha_i|\psi_i\rangle \tag{4.59}$$

$$= |\Psi\rangle \tag{4.60}$$

and indeed we see it acts on the Hilbert space as an identity operator. Equally on the dual space, for $\langle\Phi| = \sum_i \beta_i\langle\psi_i|$, we have

$$\langle\Phi| \sum_i |\psi_i\rangle\langle\psi_i| = \sum_j \beta_j\langle\psi_j| \sum_i |\psi_i\rangle\langle\psi_i| \tag{4.61}$$

$$= \sum_j \sum_i \beta_j\langle\psi_j|\psi_i\rangle\langle\psi_i| \tag{4.62}$$

$$= \sum_i \beta_i|\psi_i\rangle \tag{4.63}$$

$$= \langle\Phi|. \tag{4.64}$$

We may deduce that $\sum_i |\psi_i\rangle\langle\psi_i|$ is the identity operator, and in $\mathbb{C}^N$ this expression would represent $\sum_i^N \boldsymbol{e}_i\boldsymbol{e}_i^T$, i.e. the N by N identity matrix. This insight leads us to claim that a matrix operator $\mathcal{O}$ may be written as $\mathcal{O} = \sum_{i,j} |\phi_i\rangle\mathcal{O}_{ij}\langle\psi_j|$. Or, by taking inner products from the left and right, that the matrix elements are $\langle\psi_i|\mathcal{O}|\psi_j\rangle = \mathcal{O}_{ij}$.

4.3.1 Matrix elements in bra-ket notation

Let $\hat{\mathcal{O}} : \mathcal{H} \to \mathcal{H}$ be a linear operator on a Hilbert space $\mathcal{H}$. Then the expectation value of $\hat{\mathcal{O}}$ for a system in a normalisable state $|\Psi\rangle \in L^2(\mathbb{R})$ is defined as

$$\langle\mathcal{O}\rangle \equiv \frac{\langle\Psi|\hat{\mathcal{O}}|\Psi\rangle}{\langle\Psi|\Psi\rangle}. \tag{4.65}$$

If the state $|\Psi\rangle$ is normalised, meaning $\langle\Psi|\Psi\rangle = 1$, then this turns into

$$\langle\mathcal{O}\rangle = \langle\Psi|\hat{\mathcal{O}}|\Psi\rangle. \tag{4.66}$$

If we apply this to $\mathcal{H} = L^2(\mathbb{R})$, then by the definition of the inner product this becomes

$$\langle\mathcal{O}\rangle = \langle\Psi|\hat{\mathcal{O}}|\Psi\rangle = \langle\Psi|(\hat{\mathcal{O}}|\Psi\rangle) = \int_{-\infty}^{+\infty} dx\Psi^*(x)(\hat{\mathcal{O}}\Psi)(x). \tag{4.67}$$

Suppose now that we have an orthonormal basis of $\mathcal{H}$ which we will denote as vectors[19] given by $\{|e_1\rangle, |e_2\rangle, \cdots\} \subset \mathcal{H}$. Every vector $|\Psi\rangle \in \mathcal{H}$ can then be written as

$$|\Psi\rangle = \sum_{i=1}^{\infty} \langle e_i|\Psi\rangle |e_i\rangle \tag{4.68}$$

which we may confirm by taking the inner product $\langle e_j|\Psi\rangle$ and using that $\langle e_i|e_j\rangle = \delta_{ij}$.

We think of $\langle e_i|\Psi\rangle$ as the components of $|\Psi\rangle$ in the basis $\{|e_1\rangle, |e_2\rangle, \cdots\}$. Now $\hat{\mathcal{O}}|\Psi\rangle$ is another vector in $\mathcal{H}$. If we expand $\hat{\mathcal{O}}|\Psi\rangle$ in the $|e_i\rangle$ basis of $\mathcal{H}$ we have

$$\hat{\mathcal{O}}|\Psi\rangle = \sum_{i=1}^{\infty} \langle e_i|\hat{\mathcal{O}}|\Psi\rangle |e_i\rangle. \tag{4.69}$$

The components of $\hat{\mathcal{O}}|\Psi\rangle$ in the same basis are $\langle e_i|\hat{\mathcal{O}}|\Psi\rangle$. Plugging in the above expression for Ψ we get

$$\begin{aligned}
\langle e_k|\hat{\mathcal{O}}|\Psi\rangle &= \langle e_k| \sum_{i=1}^{\infty} \langle e_i|\hat{\mathcal{O}}|\Psi\rangle |e_i\rangle \\
&= \langle e_k| \sum_{i=1}^{\infty} \langle e_i|\hat{\mathcal{O}} \sum_{j=1}^{\infty} \langle e_j|\Psi\rangle |e_j\rangle |e_i\rangle \\
&= \langle e_k| \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \langle e_j|\Psi\rangle \langle e_i|\hat{\mathcal{O}}|e_j\rangle |e_i\rangle \\
&= \sum_{i=1}^{\infty} \langle e_k|\hat{\mathcal{O}}|e_i\rangle \langle e_i|\Psi\rangle.
\end{aligned} \tag{4.70}$$

We then see that if we arrange the components of the vectors $|\Psi\rangle$ and $\hat{\mathcal{O}}|\Psi\rangle$ into column vectors (albeit typically with an infinite number of entries) as

$$|\Psi\rangle = \begin{pmatrix} \langle e_1|\Psi\rangle \\ \langle e_2|\Psi\rangle \\ \vdots \end{pmatrix}, \qquad \hat{\mathcal{O}}|\Psi\rangle = \begin{pmatrix} \langle e_1|\hat{\mathcal{O}}|\Psi\rangle \\ \langle e_2|\hat{\mathcal{O}}|\Psi\rangle \\ \vdots \end{pmatrix}, \tag{4.71}$$

then we can write the identity (4.70) as the following matrix product

$$\begin{pmatrix} \langle e_1|\hat{\mathcal{O}}|\Psi\rangle \\ \langle e_2|\hat{\mathcal{O}}|\Psi\rangle \\ \vdots \end{pmatrix} = \begin{pmatrix} \langle e_1|\hat{\mathcal{O}}|e_1\rangle & \langle e_1|\hat{\mathcal{O}}|e_2\rangle & \cdots \\ \langle e_2|\hat{\mathcal{O}}|e_1\rangle & \langle e_2|\hat{\mathcal{O}}|e_2\rangle & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \langle e_1|\Psi\rangle \\ \langle e_2|\Psi\rangle \\ \vdots \end{pmatrix}. \tag{4.72}$$

[19] In the previous section where we were looking at $L^2(\mathbb{R})$ we adopted the notation φ_i for an eigenfunction of a Hermitian operator, if we were to write this as a vector we would denote it by $|\varphi_i\rangle$. In this section we adopt the more generic basis vector notation $|e_i\rangle$.

Of course, if $\mathcal{H}$ is finite-dimensional, these become finite matrices. The numbers $\langle e_i|\hat{\mathcal{O}}|e_k\rangle$ are called "matrix elements of $\hat{\mathcal{O}}$" because they arise in this matrix product and they may be generalised as described above to infinite-dimensional Hilbert spaces. We can think of the components of the matrix associated to $\hat{\mathcal{O}}$ in some basis as $\boldsymbol{O}_{ij} = \langle e_i|\hat{\mathcal{O}}|e_j\rangle$. If we know the matrix elements of an operator in some basis, we can use them to compute how it acts on the components of the vectors in the same basis.

For any operator we may unpack the bra-ket notation and understand its action directly on the ket $|e_j\rangle$ or via its adjoint action on the bra $\langle e_i|$:

$$\boldsymbol{O}_{ij} = \langle e_i|\hat{\mathcal{O}}|e_j\rangle = \langle e_i, \hat{\mathcal{O}}e_j\rangle = \langle (\hat{\mathcal{O}})^\dagger e_i, e_j\rangle \tag{4.73}$$

and its transpose may be defined as

$$\begin{aligned}(\boldsymbol{O}_{ij})^T = \boldsymbol{O}_{ji} &= \langle e_j|\hat{\mathcal{O}}|e_i\rangle \\ &= \langle e_j, \hat{\mathcal{O}}e_i\rangle \\ &= (\langle \hat{\mathcal{O}}e_i, e_j\rangle)^* \\ &= (\langle e_i, \hat{\mathcal{O}}^\dagger e_j\rangle)^* \\ &= (\langle e_i|\hat{\mathcal{O}}^\dagger|e_j\rangle)^*\end{aligned} \tag{4.74}$$

and so,

$$(\boldsymbol{O}_{ij})^\dagger = \langle e_i|\hat{\mathcal{O}}^\dagger|e_j\rangle \tag{4.75}$$

as we would expect, and this expression which makes direct the link between the adjoint action on a matrix and the adjoint action on an operator.

In a real finite dimensional space we understand $\boldsymbol{u}\boldsymbol{v}^T$ to be a matrix. Indeed if $\boldsymbol{u} = \sum_i u_i \boldsymbol{e}_i$, $\boldsymbol{v} = \sum_i v_i \boldsymbol{e}_i$ then $u_i v_j$ is the i, j'the entry of the matrix $\boldsymbol{u}\boldsymbol{v}^T$.

In the Hilbert space, which has a complex vector space, we can understand that $|\phi\rangle|\varphi\rangle^\dagger = |\phi\rangle\langle\varphi|$ is a matrix. Indeed $\langle\varphi|$ is a linear functional on $\mathcal{H}$ and map $\mathcal{H}$ to $\mathbb{C}$, and so $\phi\rangle\langle\varphi|$ is a linear map from $\mathcal{H}$ to $\mathcal{H}$. Expanding $|\phi\rangle = \sum_i \phi_i|e_i\rangle$ and $|\varphi\rangle = \sum_i \varphi_i|e_i\rangle$ then the components of $|\phi\rangle\langle\varphi|$ are

$$\begin{aligned}\langle e_i|\phi\rangle\langle\varphi|e_j\rangle &= \langle e_i|\sum_k \phi_k|e_k\rangle \sum_l \langle e_l|\varphi_l^*|e_j\rangle \\ &= \sum_k \phi_k \langle e_i|e_k\rangle \sum_l \varphi_l^* \langle e_l|e_j\rangle \\ &= \phi_i \varphi_j^*.\end{aligned} \tag{4.76}$$

Finally let us return to the idea of an identity operator $\hat{\mathcal{I}}$ its components, as a matrix, are

$\boldsymbol{I}_{ij} = \delta_{ij}$, that is

$$\begin{aligned}\delta_{ij} &= \langle e_i|e_j\rangle \\ &= \langle e_i|\hat{\mathcal{I}}|e_j\rangle \\ &= \langle e_i|(\sum_k |e_k\rangle\langle e_k|)|e_j\rangle \end{aligned} \tag{4.77}$$

and as we observed previously $\sum_k |e_k\rangle\langle e_k|$ is the identity operator for a complete orthonormal basis $\{|e_1\rangle, |e_2\rangle, \ldots\}$ for $\mathcal{H}$. The statement that $\hat{\mathcal{I}} = \sum_k |e_k\rangle\langle e_k|$ is the identity operator is often referred to as the completeness relation for the space, as if the $|e_k\rangle$ did not form a complete basis for $\mathcal{H}$ then some part of $\mathcal{H}$ would be projected out under the action of the operator $\hat{\mathcal{I}}$ (i.e. suppose that $\{|e_1\rangle, |e_2\rangle, \ldots\} \cup \{|f\rangle\}$ formed a complete orthonormal basis for $\mathcal{H}$ then a state $|\Psi\rangle = \sum_i \Psi_i|e_i\rangle + \Psi'|f\rangle \in \mathcal{H}$ would be mapped by $\hat{\mathcal{I}}$ to

$$\hat{\mathcal{I}}|\Psi\rangle = \sum_k |e_k\rangle\langle e_k| \left(\sum_i \Psi_i|e_i\rangle + \Psi'|f\rangle\right) = \sum_k \Psi_k|e_k\rangle \neq |\Psi\rangle. \tag{4.78}$$

Given $\hat{\mathcal{I}} = \sum_i |e_i\rangle\langle e_i|$, it is frequently a very useful algebraic move to insert the identity operator in an expression, especially bearing in mind that the identity in any complete orthonormal basis for the Hilbert space. Indeed looking back on some of our expressions developed earlier we can reinterpret some of the development as the insertion of the identity operator. A simple example of its use is

$$\langle\phi|\varphi\rangle = \langle\phi|\hat{\mathcal{I}}|\varphi\rangle = \langle\phi|(\sum_k |e_k\rangle\langle e_k|)|\varphi\rangle = \sum_k \langle\phi|e_k\rangle\langle e_k|\varphi\rangle = \sum_k \phi_k^*\varphi_k \tag{4.79}$$

where $\phi_k = \langle e_k|\phi\rangle$ and $\varphi_k = \langle e_k|\varphi\rangle$. A more involved example of its use would involve two different bases for a Hilbert space. Since we have a complete basis for any Hermitian operator, we understand there are multiple bases. Suppose we have two different complete bases: $\{|e_1\rangle, |e_2\rangle, \ldots\}$ and $\{|f_1\rangle, |f_2\rangle, \ldots\}$, then suppose we wished to compute a wave-function expressed as $\Psi = \sum_i \Psi_i|e_i\rangle$ in the $|f_i\rangle$ basis, then we can write $\hat{\mathcal{I}} = \sum_k |f_k\rangle\langle f_k|$ to find

$$|\Psi\rangle = \sum_i \Psi_i|e_i\rangle = \sum_i \Psi_i \sum_k |f_k\rangle\langle f_k|e_i\rangle = \sum_k \sum_i \Psi_i\langle f_k|e_i\rangle|f_k\rangle \tag{4.80}$$

i.e. the components of $|\Psi\rangle$ in the $|f_i\rangle$ basis are $\sum_i \Psi_i\langle f_k|e_i\rangle$. This idea will be very useful if, for example, we know the expansion of a wave-function in, say, the position basis but wish to measure an energy. When measuring a quantity it will be very useful to expand a wave-function in a basis of eigenstates of the observable (Hermitian operator), and if we wish to measure several quantities, then we would like to be able to change the basis of eigenstates used to express the wave-function.

5. The Statistical Interpretation of Quantum Mechanics

The key axiom of quantum theory is that if a state $|\Psi\rangle$ is an eigenstate of an operator $\hat{\mathcal{O}}$ with eigenvalue λ,

$$\hat{\mathcal{O}}|\Psi\rangle = \lambda|\Psi\rangle, \tag{5.1}$$

then measuring the observable described by $\hat{\mathcal{O}}$ in the state $|\Psi\rangle$ will always give the result λ. We say that $\hat{\mathcal{O}}$ has the definite value[1] λ in the state $|\Psi\rangle$.

We would now like to consider the measurement problem in a more general state $|\Phi\rangle$ which is not an eigenstate of $\hat{\mathcal{O}}$. Suppose that we measure $\hat{\mathcal{O}}$ in such a $|\Phi\rangle$. This always involves some interaction between the physical system (e.g. the electron) with the measuring device, and unlike the case of classical physics, the effect of this measurement on the physical system cannot be ignored. In general, the measurement will cause the system to transition from $|\Phi\rangle$ to some new state $|\Psi\rangle$.

Suppose that the outcome of our measurement is some number λ. We now employ the following physical intuition: if *immediately* after performing our measurement of $\hat{\mathcal{O}}$ we perform a second measurement of $\hat{\mathcal{O}}$, we should always get the same result λ. That is, the result λ of the first measurement may be random, but the second measurement should always give the same result as the first one.

The first measurement happens in the state $|\Phi\rangle$, which transitions into a state $|\Psi\rangle$. The second measurement happens therefore in the state $|\Psi\rangle$. By the above, we expect that $\hat{\mathcal{O}}$ has a definite value in the state $|\Psi\rangle$. This means that we expect $|\Psi\rangle$ to be an eigenstate of $\hat{\mathcal{O}}$ with eigenvalue λ,

$$\hat{\mathcal{O}}|\Psi\rangle = \lambda|\Psi\rangle. \tag{5.2}$$

[1]Recall that an observable is a Hermitian operator, and as we saw in section (4.2.2), the uncertainty in measuring a system in an eigenstate of a Hermitian operator is zero: the result is certain, there is no variation over a series of measurements. This is what we mean by a definite value.

This leads us to an important conclusion about quantum theory: the only possible outcomes of measuring $\hat{\mathcal{O}}$ are the eigenvalues of $\hat{\mathcal{O}}$; if we measure $\hat{\mathcal{O}}$ in a state $|\Phi\rangle$ and get a result λ, the system will have transitioned to a state $|\Psi\rangle$ which is an eigenstate of $\hat{\mathcal{O}}$ with the eigenvalue λ in the process of the measurement.

As we have seen previously, the condition that measurements give real numbers is the condition that the spectrum of an operator $\hat{\mathcal{O}}$ is real, which is implied by $\mathcal{O}$ being an Hermitian operator. Eigenstates of Hermitian operators form a basis for the Hilbert space, and so any wave-function can be expressed in terms of eigenstates of an observable. Suppose we have an observable $\hat{\mathcal{O}}$ with a discrete spectrum λ_i, then we may denote the eigenstates by $|\lambda_i\rangle$, so that $\hat{\mathcal{O}}|\lambda_i\rangle = \lambda_i|\lambda_i\rangle$. For a state,

$$|\Phi\rangle = \sum_i \Phi_i|\lambda_i\rangle \tag{5.3}$$

the statistical interpretation of quantum mechanics is that outcome λ_i of a measurement of $\hat{\mathcal{O}}$ on $|\Phi\rangle$ will occur with probability $|\Phi_i|^2$ where $\Phi_i = \langle\lambda_i|\Phi\rangle$ (as $\langle\lambda_i|\lambda_j\rangle = \delta_{ij}$).

If the spectrum of the observable $\hat{\mathcal{O}}$ is continuous and given by λ with eigenfunctions $|\lambda\rangle$ then

$$|\Phi\rangle = \int_\lambda d\lambda\, \Phi_\lambda|\lambda\rangle \tag{5.4}$$

then the probability density is $|\Phi_\lambda|^2$ (i.e. the probability of measuring $\hat{\mathcal{O}}$ between λ_1 and λ_2 for $\lambda_1 < \lambda_2$ is $\int_{\lambda_1}^{\lambda_2} d\lambda\, |\Phi_\lambda|^2$) and $\Phi_\lambda = \langle\lambda|\Phi\rangle$ (as $\langle\lambda|\lambda'\rangle = \delta(\lambda - \lambda')$).

Once the wave-function has been measured to give an outcome, it then collapses to guarantee that subsequent measurements are consistent with the initial measurement. You might say that $|\Phi\rangle$ initially contains information about each possible outcome of a measurement (the case when we have no information), but after a measurement is known the state of the system is restricted so as to agree with the observation that has been made.

In the case of discrete eigenvalues, if we measure $\hat{\mathcal{O}}$ on $\Phi\rangle$ and find the outcome is λ_i then, upon measurement, the wave-function collapses to the eigenstate $|\lambda_i\rangle$. Subsequent measurement of the eigenstate always give the definite outcome λ_i. In the case of continuous eigenvalues, then the wave-function $|\Phi\rangle$ collapses to a narrow range about the measurement - the more refined the measurement, the narrower the range of eigenfunctions about the measurement. Subsequent measurements may give slight variations of the initial measurement, but always within the range of error in the working of the measuring device.

The statistical interpretation is really Born's probabilistic interpretation of the wave-function and we can confirm for a measurement of an operator with a discrete spectrum[2] that normalisation of the wave-function guarantees that the probabilities of possible outcomes of a measure

[2]We will leave the continuous spectra to the examples.

sum to one:

$$\begin{aligned} 1 &= \langle\Phi|\Phi\rangle \\ &= \sum_i \langle\lambda_i|\Phi_i^* \sum_j \Phi_j|\lambda_j\rangle \\ &= \sum_i \sum_j \Phi_i^*\Phi_j\langle\lambda_i|\lambda_j\rangle \\ &= \sum_i \Phi_i^*\Phi_i \\ &= \sum_i |\Phi_i|^2. \end{aligned} \tag{5.5}$$

Further the expectation value of an outcome of measuring $\hat{\mathcal{O}}$ is

$$\begin{aligned} \langle\hat{\mathcal{O}}\rangle &= \langle\Phi|\hat{\mathcal{O}}|\Phi\rangle \\ &= \left(\sum_j \langle\lambda_j|\Phi_j^*\right)\hat{\mathcal{O}}\left(\sum_i \Phi_i|\lambda_i\rangle\right) \\ &= \sum_j \langle\lambda_j|\Phi_j^* \sum_i \Phi_i\lambda_i|\lambda_i\rangle \\ &= \sum_j \sum_i \lambda_i\Phi_j^*\Phi_i\langle\lambda_j|\lambda_i\rangle \\ &= \sum_j \lambda_j\Phi_j^*\Phi_j \\ &= \sum_j \lambda_j|\Phi_j|^2. \end{aligned} \tag{5.6}$$

5.1 Examples of measurement for continuous spectra

In these two examples we will look at the consistency of our ideas about measurement and probability for operators which have a continuous spectrum. Our chief concern will be that we can find a consistent set of eigenfunctions which we can work with.

5.1.1 Measuring position

We will restrict ourselves to one dimension and consider $\hat{\mathcal{O}} = \hat{x}$. We start by looking for the eigenvalues and the eigenstates. In terms of the wave-functions, the action of $\hat{x}$ is

$$(\hat{x}\Psi)(x) = x\Psi(x). \tag{5.7}$$

We would like to solve the eigenvalue equation

$$(\hat{x}\Psi)(x) = \lambda\Psi(x). \tag{5.8}$$

Plugging in the action of $\hat{x}$, we get

$$x\Psi(x) = \lambda\Psi(x), \tag{5.9}$$

equivalently

$$(x - \lambda)\Psi(x) = 0. \tag{5.10}$$

This implies that $\Psi(x) = 0$ if $x \neq \lambda$, so only $\Psi(\lambda)$ can be non-zero. Setting $\Psi(\lambda)$ to a finite value would yield a function equivalent to zero element of $L^2(\mathbb{R})$. To get a non-trivial result we need to admit a more general type of wave-function, which is not really in $L^2(\mathbb{R})$,

$$\Psi(x) = \delta(x - \lambda), \tag{5.11}$$

where δ is the Dirac delta-function. Denoting the corresponding vector by $|\lambda\rangle$, we can compute

$$\langle\lambda|\lambda'\rangle = \int_{-\infty}^{+\infty} dx\, \delta(x - \lambda)\delta(x - \lambda') = \delta(\lambda - \lambda'), \tag{5.12}$$

where we have used the defining property of the delta-function. We therefore see that $|\lambda\rangle$ forms a delta-normalised basis. Since we could choose λ to be any real number, the spectrum of $\hat{x}$ is $\mathbb{R}$.

Let us see how a general state $|\Phi\rangle$ is expanded in this basis. We write

$$|\Phi\rangle = \int_{-\infty}^{+\infty} d\lambda\, \Phi_\lambda|\lambda\rangle \tag{5.13}$$

and the coefficients $\Phi_\lambda = \Phi_\lambda(\lambda)$ are functions of λ. In terms of the wave-functions this becomes

$$\Phi(x) = \int_{-\infty}^{+\infty} d\lambda\, \Phi_\lambda(\lambda)\delta(x - \lambda). \tag{5.14}$$

We can compute the last integral using the defining property of the delta-function,

$$\Phi(x) = \int_{-\infty}^{+\infty} d\lambda\, \Phi_\lambda(\lambda)\delta(x - \lambda) = \Phi_\lambda(x). \tag{5.15}$$

Therefore, the original expression in terms of vectors can be written as

$$|\Phi\rangle = \int_{-\infty}^{+\infty} d\lambda\, \Phi(\lambda)|\lambda\rangle. \tag{5.16}$$

It is common to rename the integration variable $\lambda \to x$, so that we get

$$|\Phi\rangle = \int_{-\infty}^{+\infty} dx\, \Phi(x)|x\rangle. \tag{5.17}$$

Recall that here $|x\rangle$ is the eigenstate of $\hat{x}$ with the eigenvalue x.

We see that the coefficients $\Phi(x)$ of the expansion of $|\Phi\rangle$ into the basis of the eigenstates of $\hat{x}$ are given by the wave-function, $\Phi(x)$. Our measurement postulate tells us then that $|\Phi(x)|^2$ gives the probability density of obtaining the result x when measuring position $\hat{x}$, which is indeed how we interpreted the wave-function initially.

5.1.2 Measuring momentum

The momentum operator in one dimension is $\hat{p} = -i\hbar\frac{d}{dx}$, and its "normalised" eigenfunctions are[3] given by

$$|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar}. \tag{5.18}$$

Now we can confirm that

$$\langle p'|p\rangle = \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} dx\, e^{i(p-p')x/\hbar} = \delta(p-p'). \tag{5.19}$$

For a wave-function $\Phi(x)$ expanded in terms of eigenstates of $\hat{p}$ we have

$$\begin{aligned}\Phi(x) &= \int_{p=-\infty}^{\infty} dp\, \Phi(p)|p\rangle \\ &= \frac{1}{\sqrt{2\pi\hbar}}\int_{p=-\infty}^{\infty} dp\, \Phi(p) e^{ipx/\hbar}\end{aligned} \tag{5.20}$$

but we also know that its coefficients in the eigenfunction expansion are given by

$$\begin{aligned}\Phi(p) &= \langle p|\Phi\rangle \\ &= \frac{1}{\sqrt{2\pi\hbar}}\int_{x=-\infty}^{\infty} dx\, e^{-ipx/\hbar}\Phi(x)\end{aligned} \tag{5.21}$$

where we have used $\langle p| = (|p\rangle)^\dagger$. What we see in the two expressions in (5.19) and (5.21) are the inverse Fourier transform and the Fourier transform again, but now in the setting of Hilbert space and the bra-ket notation.

[3] In the sense that $\langle p'|p\rangle = \delta(p-p')$, as derived in equation (4.37).

5.2 Simultaneous Measurement and Commutativity

Making a measurement of $\hat{\mathcal{O}}$ on a system causes the wave-function of the system to collapse to an eigenstate of the Hermitian operator $\mathcal{O}$. We may wonder if we can make measurements of different observables $\hat{A}$ and $\hat{B}$ and get definite (i.e. unchanging) outcomes for both measurements. If this is the case then $\hat{A}$ and $\hat{B}$ must have their eigenfunctions in common. Suppose this is the case and that Ψ is a common eigenfunction with $\hat{A}\Psi = a\Psi$, and $\hat{B}\Psi = b\Psi$, consequently

$$\hat{B}\hat{A}\Psi = a\hat{B}\Psi = ab\Psi \tag{5.22}$$

$$\hat{A}\hat{B}\Psi = b\hat{A}\Psi = ba\Psi \tag{5.23}$$

hence

$$[\hat{A}, \hat{B}]\Psi = (\hat{A}\hat{B} - \hat{B}\hat{A})\Psi = 0. \tag{5.24}$$

As this must hold for any eigenvalues, a and b, and so consequently any simultaneous eigenfunction of $\hat{A}$ and $\hat{B}$, then we deduce that if $\hat{A}$ commutes with $\hat{B}$, i.e. if $[\hat{A}, \hat{B}] = 0$ then they have common eigenfunctions and they can be measured in any order, and yield the same consistent outcomes.

We can also run this argument in the other direction: suppose $|a_i\rangle$ are an orthonormal basis of eigenstates for $\hat{A}$ with $\hat{A}|a_i\rangle = a_i|a_i\rangle$ and that $\hat{B}$ commutes with $\hat{A}$, then,

$$\begin{aligned} 0 &= \langle a_i|[\hat{A}, \hat{B}]|a_j\rangle \\ &= \langle a_i|\hat{A}\hat{B}|a_j\rangle - \langle a_i|\hat{B}\hat{A}|a_j\rangle \\ &= a_i\langle a_i|\hat{B}|a_j\rangle - a_j\langle a_i|\hat{B}|a_j\rangle \\ &= (a_i - a_j)\langle a_i|\hat{B}|a_j\rangle \end{aligned} \tag{5.25}$$

where we developed $\langle a_i|\hat{A}\hat{B}|a_j\rangle = a_i\langle a_i|\hat{B}|a_j\rangle$, e.g. by considering $\langle a_i|\hat{A} = (\hat{A}|a_i\rangle)^\dagger = (a_i|a_i\rangle)^\dagger = a_i\langle a_i|$ as $\hat{A}$ is Hermitian and so $a_i \in \mathbb{R}$. Then for $a_i \neq a_j$ we must have $\langle a_i|\hat{B}|a_j\rangle = 0$ and so $\langle a_i|\hat{B}|a_j\rangle = C\delta_{ij} = C\langle a_i|a_j\rangle$ and so $\hat{B}|a_j\rangle = C|a_j\rangle$ and the eigenvectors of $\hat{A}$ are also eigenvectors $\hat{B}$ (if $[\hat{A}, \hat{B}] = 0$). When we can measure operators one after the other without disturbing the system we say the operators are compatible. We also say that we can measure the operators simultaneously, because although we imagined sequential measurements, we are not considering that the system is evolving in any way, i.e. it is only being affected by our measurements.

Suppose now that $\hat{A}$ and $\hat{B}$ are two observables that do not commute. When we measure $\hat{B}$ on Ψ the wave-function collapses to an eigenfunction of $\hat{B}$, say $|b_i\rangle$, consistent with the outcome of the measurement. If we now measure $\hat{A}$ on the state $|b_i\rangle$ the wave-function $|b_i\rangle$ collapses to an eigenstate of $\hat{A}$, say $|a_i\rangle$, necessarily distinct from the eigenfunction $|b_i\rangle$ as only commuting

observables have a shared set of eigenfunctions. Now if we measure $\hat{B}$ again on the system, now in state $|a_i\rangle = \sum_{k=1}^{\infty} \beta_k |b_k\rangle$ we only find the outcome b_i with probability $|\beta_i|^2$, i.e. we are not certain to measure the same outcome as we did previously for $\hat{B}$.

For example consider $\hat{A} = \hat{x}$ and $\hat{B} = \hat{p}$. We compute

$$(\hat{x}\hat{p}\Psi)(x) = (\hat{x}(\hat{p}\Psi))(x) = x(\hat{p}\Psi)(x) = -i\hbar x \frac{\partial \Psi(x)}{\partial x} \tag{5.26}$$

$$(\hat{p}\hat{x}\Psi)(x) = (\hat{p}(\hat{x}\Psi))(x) = -i\hbar \frac{\partial(\hat{x}\Psi)(x)}{\partial x} = -i\hbar \frac{\partial(x\Psi(x))}{\partial x} = -i\hbar x \frac{\partial \Psi(x)}{\partial x} - i\hbar\Psi(x). \tag{5.27}$$

Therefore

$$(([\hat{x}, \hat{p}])\Psi)(x) = (\hat{x}\hat{p}\Psi)(x) - (\hat{p}\hat{x}\Psi)(x) = i\hbar\Psi(x). \tag{5.28}$$

Since $\Psi(x)$ was arbitrary we get

$$[\,\hat{x}, \hat{p}\,] = i\hbar. \tag{5.29}$$

Thus there doesn't exist a single state in which both $\hat{x}$ and $\hat{p}$ have a definite value. This is clearly very different from the classical theory in which x and p can be measured simultaneously. We might say that classically x and p commute, and this is why we can measure them simultaneously. That said, the quantum mechanical commutator $[\hat{x}, \hat{p}]$ while non-zero, is proportional to $\hbar$ which is a small number. In this sense, the commutator may be thought of as a small quantum correction to the classical theory.

Since $\hat{H} = \hat{p}^2/2m + V(\hat{x})$ then whether $\hat{H}$ commutes with $\hat{p}$ or $\hat{x}$ depends on the form of $V(\hat{x})$ and it won't generally be possible to measure either of position or momentum and energy definitely for a state. There are special cases of course, for example the free particle with $V(\hat{x}) = 0$, then $[\hat{H}, \hat{p}] = 0$ and one can determine both energy and momentum simultaneously. This is perhaps not so surprising as for the free particle $E = p^2/2m$.

5.3 The Generalised Uncertainty Principle

We can further quantify our observations about simultaneous measurement in the generalised uncertainty principle. Recall the definition of the uncertainty $\Delta\hat{A}$ of measuring an observable $\hat{A}$ on system in a normalised state $|\Psi\rangle$. Recall that the expectation value of $\hat{A}$ is

$$\langle \hat{A} \rangle = \langle \Psi | \hat{A} | \Psi \rangle. \tag{5.30}$$

Of course, $\langle \hat{A} \rangle$ describes not a single measurement of $\hat{A}$, but the average of many measurements, each performed on the same wave-function $|\Psi\rangle$. Each individual measurement will be different unless $|\Psi\rangle$ is an eigenstate of $\hat{A}$. We can quantify this variation by computing the expectation value of $(\hat{A} - \langle \hat{A} \rangle)^2$, i.e. the expected squared deviation of $\hat{A}$ from its average value,

$$(\Delta\hat{A})^2 = \langle (\hat{A} - \langle \hat{A} \rangle)^2 \rangle = \langle \Psi | (\hat{A} - \langle \hat{A} \rangle)^2 | \Psi \rangle. \tag{5.31}$$

The uncertainty $\Delta\hat{A}$ is defined by this equation. There is a very useful way to rewrite this formula for uncertainty:

$$\begin{aligned}(\Delta\hat{A})^2 &= \langle(\hat{A}-\langle\hat{A}\rangle)^2\rangle \\ &= \langle\hat{A}^2 - 2\hat{A}\langle\hat{A}\rangle + \langle\hat{A}\rangle^2|\Psi\rangle \\ &= \langle\hat{A}^2\rangle - 2\langle\hat{A}\rangle\langle\hat{A}\rangle + \langle\hat{A}\rangle^2 \\ &= \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2 \end{aligned} \tag{5.32}$$

and this version of the uncertainty $\Delta\hat{A}$ is frequently the simplest way to make a calculation of the the uncertainty for an operator for a system in a given state.

Let us show that $(\Delta\hat{A})^2 \geq 0$, so that the uncertainty is a real quantity. We have

$$\begin{aligned}(\Delta\hat{A})^2 &= \langle\Psi|(\hat{A}-\langle\hat{A}\rangle)^2|\Psi\rangle \\ &= \langle\Psi|(\hat{A}-\langle\hat{A}\rangle)(\hat{A}-\langle\hat{A}\rangle)|\Psi\rangle \\ &= \langle\Psi|(\hat{A}-\langle\hat{A}\rangle)^\dagger(\hat{A}-\langle\hat{A}\rangle)|\Psi\rangle \\ &= \Big((\hat{A}-\langle\hat{A}\rangle)|\Psi\rangle\Big)^\dagger(\hat{A}-\langle\hat{A}\rangle)|\Psi\rangle \\ &= \Big\|(\hat{A}-\langle\hat{A}\rangle)|\Psi\rangle\Big\|^2 \\ &\geq 0. \end{aligned} \tag{5.33}$$

We used that $\hat{A}^\dagger = \hat{A}$ by the assumption that $\hat{A}$ is Hermitian and so also $\langle\hat{A}\rangle^\dagger = \langle\hat{A}\rangle$.

We will now prove the following theorem

Theorem 5.1. *Let $\hat{A}, \hat{B}$ be Hermitian operators, with $\Delta\hat{A}, \Delta\hat{B}$ their uncertainties in a state $|\Psi\rangle$, then*

$$\Delta\hat{A}\,\Delta\hat{B} \geq \frac{1}{2}\left|\langle[\hat{A},\hat{B}]\rangle\right|. \tag{5.34}$$

Proof. Let us define $\hat{A}_0 = \hat{A} - \langle\hat{A}\rangle$ and $\hat{B}_0 = \hat{B} - \langle\hat{B}\rangle$. We first compute

$$\begin{aligned}[\hat{A}_0,\hat{B}_0] &= [\hat{A},\hat{B}] - [\langle\hat{A}\rangle,\hat{B}\,] - [\hat{A},\langle\hat{B}\rangle] + [\langle\hat{A}\rangle,\langle\hat{B}\rangle] \\ &= [\hat{A},\hat{B}\,], \end{aligned} \tag{5.35}$$

where we have used the fact that $\langle\hat{A}\rangle$ and $\langle\hat{B}\rangle$ are ordinary numbers and thus their commutators with any operators are zero. Furthermore, $\hat{A}_0$ and $\hat{B}_0$ are Hermitian since $\langle\hat{A}\rangle$ and $\langle\hat{B}\rangle$ are real (prove this). Then, we have, from the definition of $\Delta\hat{A}$,

$$\begin{aligned}(\Delta\hat{A})^2 &= \langle\Psi|(\hat{A}-\langle\hat{A}\rangle)^2|\Psi\rangle \\ &= \langle\Psi|\hat{A}_0^2|\Psi\rangle \end{aligned} \tag{5.36}$$

and, similarly, $(\Delta\hat{B})^2 = \langle\Psi|\hat{B}_0^2|\Psi\rangle$, which implies

$$(\Delta\hat{A})^2(\Delta\hat{B})^2 = \langle\Psi|\hat{A}_0^2|\Psi\rangle\langle\Psi|\hat{B}_0^2|\Psi\rangle = \|\hat{A}_0|\Psi\rangle\|^2\|\hat{B}_0|\Psi\rangle\|^2. \tag{5.37}$$

In the last equality we used the same computation that we used when proving $(\Delta\hat{A})^2 \geq 0$, $\langle\Psi|\hat{A}_0^2|\Psi\rangle = \langle\Psi|\hat{A}_0^\dagger A_0|\Psi\rangle = (A_0|\Psi\rangle)^\dagger A_0|\Psi\rangle$.

Recall now the Cauchy-Schwartz inequality

$$|\langle u|v\rangle|^2 \leq \||u\rangle\|^2\||v\rangle\|^2, \tag{5.38}$$

which is the generalisation of the idea that $|\cos\theta| \leq 1$. Using this above, we have

$$\begin{aligned}(\Delta\hat{A})^2(\Delta\hat{B})^2 &= \|\hat{A}_0|\Psi\rangle\|^2\|\hat{B}_0|\Psi\rangle\|^2 \\ &\geq \left|(\hat{A}_0|\Psi\rangle)^\dagger\hat{B}_0|\Psi\rangle\right|^2 \\ &= \left|\langle\Psi|\hat{A}_0\hat{B}_0|\Psi\rangle\right|^2.\end{aligned} \tag{5.39}$$

Using the fact that for any complex number $z \in \mathbb{C}$ we have $|z|^2 \geq |\mathrm{Im}\, z|^2 = \frac{1}{4}|z - z^*|^2$, we have

$$\begin{aligned}(\Delta\hat{A})^2(\Delta\hat{B})^2 &\geq \left|\langle\Psi|\hat{A}_0\hat{B}_0|\Psi\rangle\right|^2 \\ &\geq \frac{1}{4}\left|\langle\Psi|\hat{A}_0\hat{B}_0|\Psi\rangle - (\langle\Psi|\hat{A}_0\hat{B}_0|\Psi\rangle)^*\right|^2\end{aligned} \tag{5.40}$$

Finally, using $(\langle\Psi|\hat{A}_0\hat{B}_0|\Psi\rangle)^* = (\langle\Psi|\hat{A}_0\hat{B}_0|\Psi\rangle)^\dagger = |\Psi\rangle^\dagger\hat{B}_0^\dagger\hat{A}_0^\dagger\langle\Psi|^\dagger = \langle\Psi|\hat{B}_0\hat{A}_0|\Psi\rangle$, we find

$$\begin{aligned}(\Delta\hat{A})^2(\Delta\hat{B})^2 &\geq \frac{1}{4}\left|\langle\Psi|\hat{A}_0\hat{B}_0|\Psi\rangle - \langle\Psi|\hat{B}_0\hat{A}_0|\Psi\rangle\right|^2 \\ &= \frac{1}{4}\left|\langle\Psi|[\hat{A}_0, \hat{B}_0]|\Psi\rangle\right|^2 \\ &= \frac{1}{4}\left|\langle\Psi|[\hat{A}, \hat{B}]|\Psi\rangle\right|^2.\end{aligned} \tag{5.41}$$

Taking the square root we find

$$\Delta\hat{A}\ \Delta\hat{B} \geq \frac{1}{2}\left|\langle[\hat{A}, \hat{B}]\rangle\right|. \tag{5.42}$$

□

This relation for any pair of operators is called the generalised uncertainty principle. When we take $\hat{A} = \hat{x}$ and $\hat{B} = \hat{p}$ we get

$$\Delta\hat{x}\ \Delta\hat{p} \geq \frac{1}{2}\left|\langle[\hat{x}, \hat{p}]\rangle\right| = \frac{1}{2}\left|\langle i\hbar\rangle\right| = \frac{\hbar}{2}, \tag{5.43}$$

which is called the Heisenberg uncertainty principle. It tells us that the uncertainty in simultaneous measurements of $\hat{x}$ and $\hat{p}$ is not due to error in our measurements, rather that it is inherent in quantum mechanics. The better we can determine the position, the worse we know the momentum and vice-versa. Heisenberg used the example of determining the location of a particle very precisely by using radiation of short wavelength, the shorter the wavelength the better the resolution one can determine. However short wavelength radiation is of higher energy than long wavelength radiation, and the energy transferred to the particle in measuring its position changes its momentum. So one cannot measure both precisely, and moreover there is no clever experiment that can get around this idea, uncertainty is inherent in quantum mechanics.

5.3.1 The Energy Time Uncertainty Relation

Let us assume that the observable $\hat{\mathcal{O}}$ is explicitly independent of time, then[4]

$$\begin{aligned}\frac{d}{dt}\langle\hat{\mathcal{O}}\rangle &= \frac{d}{dt}\langle\Psi|\hat{\mathcal{O}}|\Psi\rangle \\ &= \frac{d}{dt}\langle\Psi|\hat{\mathcal{O}}\Psi\rangle \\ &= \langle\frac{\partial\Psi}{\partial t}|\hat{\mathcal{O}}\Psi\rangle + \langle\Psi|\hat{\mathcal{O}}\frac{\partial\Psi}{\partial t}\rangle. \end{aligned} \tag{5.44}$$

As $i\hbar\frac{\partial\Psi}{\partial t} = \hat{H}\Psi$ then $\frac{\partial\Psi}{\partial t} = -\frac{i}{\hbar}\hat{H}\Psi$ and

$$\begin{aligned}\frac{d}{dt}\langle\hat{\mathcal{O}}\rangle &= \frac{i}{\hbar}\langle\hat{H}\Psi|\hat{\mathcal{O}}\Psi\rangle - \frac{i}{\hbar}\langle\Psi|\hat{\mathcal{O}}\hat{H}\Psi\rangle \\ &= \frac{i}{\hbar}\langle\Psi|\hat{H}\hat{\mathcal{O}}\Psi\rangle - \frac{i}{\hbar}\langle\Psi|\hat{\mathcal{O}}\hat{H}\Psi\rangle \\ &= \frac{i}{\hbar}\langle\Psi|\hat{H}\hat{\mathcal{O}} - \hat{\mathcal{O}}\hat{H}|\Psi\rangle \\ &= \frac{i}{\hbar}\langle\Psi|[\hat{H},\hat{\mathcal{O}}]|\Psi\rangle. \end{aligned} \tag{5.45}$$

Using the generalised uncertainty relation with $\hat{A} = \hat{H}$ and $\hat{B} = \hat{\mathcal{O}}$ we have,

$$\begin{aligned}\Delta\hat{H}\Delta\hat{\mathcal{O}} &\geq \frac{1}{2}|\langle[\hat{H},\hat{\mathcal{O}}]\rangle| \\ &= \frac{1}{2}|-i\hbar\frac{d}{dt}\langle\hat{\mathcal{O}}\rangle.| \end{aligned} \tag{5.46}$$

[4]We already did this computation before when we developed Ehrenfest's theorems in section 2.8, but we repeat it here as we covered this a long time ago now, and we can develop our confidence in using bra-ket notation.

Writing $\Delta\hat{H} = \Delta E$ and $\Delta\hat{\mathcal{O}} = \Delta t|\frac{d}{dt}\langle\hat{\mathcal{O}}\rangle|$ then we have

$$\Delta E \Delta t \geq \frac{\hbar}{2}. \tag{5.47}$$

This is called the energy time uncertainty relation, but it's a little different to our generalised uncertainty relation, because t is not an operator in quantum mechanics, it is just a parameter. So what is Δt really representing? We have $\Delta t = \Delta\hat{\mathcal{O}}/|\frac{d}{dt}\langle\hat{\mathcal{O}}\rangle|$ and so it is the time for the expectation value of an operator $\hat{\mathcal{O}}$ to change by one standard deviation (or uncertainty), $\Delta\hat{O}$. To make this more concrete, consider a Gaussian wavepacket for a free particle, which at time $t = 0$ has $|\Psi|^2 = |A|^2 \exp(-(x - x_0)^2/(2\Delta x)^2)$. The quantity Δx in the Gaussian wavepacket is the uncertainty in measuring $\hat{x}$, the notation was chose on purpose, and then one standard deviation in x from x_0 corresponds to $|x - x_0| = \Delta x$ and so it corresponds to the point on the Gaussian wavepacket of height $|A|^2/\sqrt{e}$ (approximately 68% of the Gaussian wavepacket is within a standard deviation of x_0). Now consider an experiment measuring the time it takes a free particle to move past a point in space, the energy time uncertainty relation indicates that the uncertainty in the measurement of time will be $\Delta t = \frac{\hbar}{2\Delta E}$. It is an equality because we are considering a Gaussian wavepacket, for which the uncertainty relation becomes an equality (see the problem set for the problem minimising uncertainty). But since $\Delta E = \Delta(p^2/2m) = p(\Delta p)/m$ and for the Gaussian wave packet we saw that[5] $\frac{d\langle\hat{x}\rangle}{dt} = v$, the classical velocity, hence here $\Delta t = \Delta\hat{x}/|\frac{d}{dt}\langle\hat{x}\rangle| = \Delta\hat{x}/v$ and so, as we claimed

$$\Delta E \Delta t = \frac{p}{mv}\Delta x \Delta p = \frac{\hbar}{2} \tag{5.48}$$

but we note that we have really applied the position uncertainty relation. The energy time uncertainty relation is often used in a cavalier way to motivate correct statements, without a full computation. For example you may come across the energy time uncertainty relation being cited to justify the idea large amounts of energy can be borrowed for short amounts of time, e.g. when discussing virtual particle in quantum field theory. In this example one should do a full quantum field theory calculation to find the result. For the energy time uncertainty condition one should take care to pin down the meaning precisely and recall it is not saying that energy is not conserved, rather that there is a spread of energies that can be measured, and further that it is not time that is borrowed against, rather it is the time period it would take for a standard deviations worth of change in the expectation value of an operator. This is not as exciting, but it has the benefit of being correct.

[5]See Ehrenfest's theorems in section 2.8.

5.4 The Time Evolution Operator

Time evolution of a wave-function $\Psi(x,t)$ is governed by Schrödinger's equation:

$$i\hbar\frac{\partial\Psi(x,t)}{\partial t} = \hat{H}\Psi(x,t). \tag{5.49}$$

We would like to interpret this equation in the Hilbert space language. However, so far, we have thought of the state vectors of the Hilbert space as being time-independent[6]. In the Schrödinger equation the wave-functions are time-dependent[7]. To construct a time-dependent wave-function in Hilbert space, we can allow the components of the state vector in Hilbert space to become time-dependent and write

$$|\Psi(t)\rangle = \sum_i \psi_i(t)|\psi_i\rangle \tag{5.50}$$

for orthonormal basis vectors $|\psi_i\rangle \in \mathcal{H}$ where $i \in \mathbb{N}$. As time passes, the wave-function sweeps out a path in Hilbert space.

How does Schrödinger's equation look in the bra-ket notation on Hilbert space? It becomes

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle \tag{5.51}$$

and we may solve this, formally, to find,

$$|\Psi(t)\rangle = Ae^{-i\hat{H}t/\hbar}. \tag{5.52}$$

Since at $t = 0$ we have $|\Psi(0)\rangle = A$ then

$$|\Psi(t)\rangle = e^{-i\hat{H}t/\hbar}|\Psi(0)\rangle. \tag{5.53}$$

We understand the exponential of an operator $\hat{\mathcal{O}}$ to be defined by a power series as

$$e^{\alpha\hat{\mathcal{O}}} = \hat{\mathcal{I}} + \alpha\hat{\mathcal{O}} + \frac{(\alpha\hat{\mathcal{O}})^2}{2!} + \frac{(\alpha\hat{\mathcal{O}})^3}{3!} + \cdots, \tag{5.54}$$

where $\hat{\mathcal{I}}$ denotes the identity operator. For a finite-dimensional Hilbert space, the operator $\hat{\mathcal{O}}$ would be a matrix, and each term would be defined by the repeated action of the matrix on itself by matrix-multiplication. Similarly for an operator in infinite-dimensional Hilbert space, the operator would act repeatedly on itself.

[6]Such a view, as we will define, is part of the Heisenberg picture of quantum mechanics.

[7]This is a feature of the so-called Schrödinger picture of quantum mechanics where the state depends on the time.

The Hamiltonian operator $\hat{H}$ has a simple action on its normalised, eigenfunctions φ_n, the stationary states. If we label these states by their eigenvalue E_n i.e. $\varphi_n \to |n\rangle$ with $\hat{H}|n\rangle = E_n|n\rangle$, then, writing $\alpha = -it/\hbar$, we have

$$\begin{aligned} e^{-i\hat{H}t/\hbar}|n\rangle &= \left(\hat{\mathcal{I}} + \alpha\hat{H} + \frac{\alpha^2}{2!}\hat{H}^2 + \frac{\alpha^3}{3!}\hat{H}^3 + \cdots\right)|n\rangle \\ &= \left(\hat{\mathcal{I}} + \alpha E_n + \frac{\alpha^2}{2!}E_n^2 + \frac{\alpha^3}{3!}E_n^3 + \cdots\right)|n\rangle \\ &= e^{-iE_nt/\hbar}|n\rangle \end{aligned} \tag{5.55}$$

which is a well-defined vector in the Hilbert space. One should compare the expression above to the result we found in section 3.1 where we assumed that we could separating the dependence on the variables x and t in $\Psi(x,t)$ i.e. by writing $\Psi(x,t) = \phi(t)\varphi(x)$. This led us to the time-independent Schrödinger equation which we could solve to find an energy eigenstates of the system at energy E so that $\Psi(x,t) = e^{-iEt/\hbar}\varphi(x)$ where $\varphi(x)$ is the eigenfunction of $\hat{H}$ with eigenvalue E.

As each basis vector evolves under the action of the operator $e^{-i\hat{H}t/\hbar}$ then its action on any vector in the Hilbert space is also well-defined. For example suppose $|\Psi\rangle = \sum_n \psi_n|n\rangle$, where ψ_n are constants, in the energy eigenstate basis of $\mathcal{H}$, then

$$\begin{aligned} e^{-i\hat{H}t/\hbar}|\Psi\rangle &= e^{-i\hat{H}t/\hbar}\sum_n \psi_n|n\rangle \\ &= \sum_n \psi_n e^{-i\hat{H}t/\hbar}|n\rangle \\ &= \sum_n \psi_n e^{-iE_nt/\hbar}|n\rangle. \end{aligned} \tag{5.56}$$

To summarise, the time evolution of the wave-function given by Schrödinger's equation results in the motion of a vector $|\Psi(0)\rangle$ to $|\Psi(t)\rangle$ in the Hilbert space, given by

$$|\Psi(t)\rangle = \hat{\mathcal{U}}(t)|\Psi(0)\rangle \tag{5.57}$$

where

$$\hat{\mathcal{U}}(t) \equiv e^{-i\hat{H}t/\hbar} \tag{5.58}$$

is called the time evolution operator. It is a unitary operator as

$$\hat{\mathcal{U}}(t)^\dagger = e^{i\hat{H}^\dagger t/\hbar} = e^{i\hat{H}t/\hbar} = \hat{\mathcal{U}}(-t) = \hat{\mathcal{U}}(t)^{-1} \tag{5.59}$$

as the Hamiltonian operator is Hermitian. The existence of an inverse means that the $\hat{\mathcal{U}}(t)$ may form a group. Indeed

$$\hat{\mathcal{U}}(t_1)\hat{\mathcal{U}}(t_2) = \hat{\mathcal{U}}(t_1 + t_2) \tag{5.60}$$

and the identity element is $\hat{\mathcal{U}}(0)$, and one can show that the multiplicative action is associative[8]. The group is the one-dimensional unitary group $U(1) \cong SO(2)$. The fact that the time evolution operator is periodic is, at first, intriguing. Notice that on a single energy eigenstate $|n\rangle$, $\hat{\mathcal{U}}(t)$ is evaluated as $e^{-iE_nt/\hbar}$ which indicates that after a time $\tau = 2\pi\hbar/E_n$ the eigenstate returns to its initial vector in the Hilbert space, indeed on an eigenstate $e^{-iE_nt/\hbar}$ is a phase factor, having no physical meaning. However on a general state, which will be a sum of energy eigenstates, each eigenstate will evolve under $\hat{\mathcal{U}}(t)$ at a different rate determined by the energy E_n of the eigenstate. However, if all the energies of a system differ by a multiple of some fundamental amount of energy[9], then the wave-function of the system will return to its original state in a finite time[10].

The ideas here may be summarised in the statement that a Hermitian Hamiltonian leads to a unitary time evolution of the wave-function[11]. Further as

$$\begin{aligned}\langle\Psi(t)|\Psi(t)\rangle &= \langle\hat{\mathcal{U}}(t)|\Psi(0)\rangle,\hat{\mathcal{U}}(t)|\Psi(0)\rangle\rangle \\ &= \langle\hat{\mathcal{U}}(t)^\dagger\hat{\mathcal{U}}(t)|\Psi(0)\rangle, |\Psi(0)\rangle\rangle \\ &= \langle|\Psi(0)\rangle, |\Psi(0)\rangle\rangle \\ &= \langle\Psi(0)|\Psi(0)\rangle \end{aligned} \tag{5.61}$$

then the unitary evolution does not change the normalisation of the wave-function.

Example 5.1. *Suppose that*

$$|0\rangle = \begin{pmatrix}1\\0\end{pmatrix}, \quad \textit{and} \quad |1\rangle = \begin{pmatrix}0\\1\end{pmatrix}$$

form an orthonormal basis for $\mathcal{H}$ and that the Hamiltonian of a system in this basis is the Hermitian matrix,

$$\hat{H} = \begin{pmatrix}a & b\\ b & a\end{pmatrix}.$$

where a and b are real constants. What are the energies allowed by the system? Identify the unit-length eigenvectors of the Hamiltonian $\hat{H}$. If at $t = 0$ the system is in state $|0\rangle$, what is its state at some later time t?

---

[8] If in doubt about any of these claims, try the out by acting on an energy eigenstate $|n\rangle$.

[9] This will be the case for the quantum harmonic oscillator which we will study in the next chapter.

[10] For a simple two state case see the following example (example 5.1) where the wave function returens to an earlier state periodically.

[11] One can also argue in the other direction, that a unitary evolution of a wave-function in Hilbert space, leads to Schrödinger's equation. See, for example, chapter 16 of Mastering Quantum Mechanics by Barton Zwiebach.

We first seek to solve the time-independent Schrödinger equation $\hat{H}\varphi(x) = E\varphi(x)$, i.e. we must find the eigenvectors of $\hat{H}$. We find these by solving the equation

$$\det(\hat{H} - E_\pm) = (a - E)^2 - b^2 = 0$$

so $(a - E_\pm) = \pm b$, i.e. $E_\pm = a \mp b$. These are the two energies of the system. We next determine the energy eigenvectors for each eigenvalue. First we note that a general state in the system is given by $|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$ for $\alpha, \beta \in \mathbb{C}$. For each energy E_+ or E_-, we seek α and β such that

$$\begin{pmatrix} a & b \\ b & a \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E_\pm \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

For $E_\pm = a \mp b$ we have

$$\begin{aligned} a\alpha + b\beta &= (a \mp b)\alpha \\ b\alpha + a\beta &= (a \mp b)\beta \end{aligned}$$

and we read off that $\beta = \mp\alpha$ hence the two normalised energy eigenstates are

$$|E_-\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \textit{and} \quad |E_+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Now we can express the basis states $|0\rangle$ and $|1\rangle$ in terms of the energy eigenvectors. We do this because we know that the energy eigenstates evolve according to $\exp(-iE_\pm t/\hbar)|E_\pm\rangle$. We have

$$\begin{aligned} |0\rangle &= \frac{1}{\sqrt{2}}(|E_-\rangle + |E_+\rangle) \\ |1\rangle &= \frac{1}{\sqrt{2}}(|E_-\rangle - |E_+\rangle). \end{aligned}$$

For a system starting with $|\Psi(0)\rangle = |0\rangle$ then we have

$$\begin{aligned}
|\Psi(t)\rangle &= \hat{\mathcal{U}}(t)|\Psi(0)\rangle \\
&= e^{-i\hat{H}t/\hbar}\frac{1}{\sqrt{2}}(|E_-\rangle + |E_+\rangle) \\
&= \frac{1}{\sqrt{2}}(e^{-i\hat{E}_-t/\hbar}|E_-\rangle + e^{-i\hat{E}_+t/\hbar}|E_+\rangle) \\
&= \frac{1}{\sqrt{2}}(e^{-i(a+b)t/\hbar}|E_-\rangle + e^{-i(a-b)t/\hbar}|E_+\rangle) \\
&= \frac{1}{2}(e^{-i(a+b)t/\hbar}(|0\rangle + |1\rangle) + e^{-i(a-b)t/\hbar}(|0\rangle - |1\rangle)) \\
&= \frac{1}{2}e^{-iat/\hbar}\begin{pmatrix} e^{-ibt/\hbar} + e^{ibt/\hbar} \\ e^{-ibt/\hbar} - e^{ibt/\hbar} \end{pmatrix} \\
&= e^{-iat/\hbar}\begin{pmatrix} \cos(bt/\hbar) \\ -i\sin(bt/\hbar). \end{pmatrix}
\end{aligned}$$

Hence the wave-function oscillates between state $|0\rangle$ and state $|1\rangle$ and at times $t = n\hbar\pi/b$ it returns to having probability one of being in state $|0\rangle$.

5.5 The Heisenberg Picture

We have been developing a version of quantum mechanics in which the states $|\Psi(t)\rangle$ depend on time, while the operators such as $\hat{x}$, $\hat{p}$ and $\hat{H}$ are time-independent. This is called the Schrödinger picture of quantum mechanics. We will indicate states and operators in the Schrödinger picture with a subscript S, e.g. $|\Psi_S(t)\rangle$ will be a time-dependent state and $\hat{\mathcal{O}}_S$ will be a time-independent operator.

We have seen in the time evolution operator $\hat{\mathcal{U}}(t)$ an example of a time-dependent operator, which we used to act on time-independent states. This invites the complementary approach to quantum mechanics, in which the states are time-independent but the operators are time-dependent. This framework, is called the Heisenberg picture, and we will develop it in this section. The Heisenberg picture will give the same predictions for all experiments as the Schrödinger picture, but in the Heisenberg picture the states $|\Psi_H\rangle$ do not depend on time, while the operators $\hat{\mathcal{O}}_H(t)$ do (here we are have added a subscript H to indicate the Heisenberg picture).

For example, in the Heisenberg picture the expectation value of an observable $\hat{\mathcal{O}}_H(t)$ is

$$\langle\Psi_H|\hat{\mathcal{O}}_H(t)|\Psi_H\rangle \tag{5.62}$$

and it depends on time only due to the time-dependence of $\hat{\mathcal{O}}_H(t)$. The expectation values in both pictures should agree,

$$\langle\Psi_H|\hat{\mathcal{O}}_H(t)|\Psi_H\rangle = \langle\Psi_S(t)|\hat{\mathcal{O}}_S|\Psi_S(t)\rangle. \tag{5.63}$$

Using the time evolution operator $\hat{\mathcal{U}}(t)$, we have $|\Psi_S(t)\rangle = \hat{\mathcal{U}}(t)|\Psi_S(0)\rangle$, which gives

$$\langle\Psi_H|\hat{\mathcal{O}}_H(t)|\Psi_H\rangle = \langle\Psi_S(0)|\hat{U}(t)^\dagger\hat{\mathcal{O}}_S\hat{U}(t)|\Psi_S(0)\rangle. \tag{5.64}$$

We see that the two expressions agree if we identify

$$\boxed{\begin{aligned} |\Psi_H\rangle &= |\Psi_S(0)\rangle, \qquad (5.65)\\ \hat{\mathcal{O}}_H(t) &= \hat{\mathcal{U}}(t)^\dagger\hat{\mathcal{O}}_S\hat{\mathcal{U}}(t). \qquad (5.66)\end{aligned}}$$

These formulae define the correspondence between the Heisenberg and Schrödinger pictures.

Consider the eigenstates of $\hat{\mathcal{O}}$ in the Schrödinger picture, we have

$$\hat{\mathcal{O}}_S|\lambda_S\rangle = \lambda_S|\lambda_S\rangle. \tag{5.67}$$

In the Heisenberg picture, the operators $\hat{\mathcal{O}}_H(t)$ depend on time, consequently their eigenfunctions will also depend on time, since they have to satisfy the equation

$$\hat{\mathcal{O}}_H(t)|\lambda_H(t)\rangle = \lambda|\lambda_H(t)\rangle. \tag{5.68}$$

While it is not obvious, the eigenvalues, on the other hand, do not depend on time. Let us try to translate the equation above to the Schrödinger picture. Using (5.66) we have

$$\hat{\mathcal{O}}_H(t)|\lambda_H(t)\rangle = (\hat{\mathcal{U}}(t)^\dagger\hat{\mathcal{O}}_S\hat{\mathcal{U}}(t))|\lambda_H(t)\rangle = \lambda|\lambda_H(t)\rangle. \tag{5.69}$$

and from (5.67) we see that we can solve the equation above if

$$|\lambda_H(t)\rangle = \hat{\mathcal{U}}(t)^\dagger|\lambda_S\rangle \tag{5.70}$$

as we then see that

$$\begin{aligned}\hat{\mathcal{O}}_H(t)|\lambda_H(t)\rangle &= \hat{\mathcal{U}}(t)^\dagger\hat{\mathcal{O}}_S\hat{\mathcal{U}}(t)\hat{\mathcal{U}}(t)^\dagger|\lambda_S\rangle\\ &= \hat{\mathcal{U}}(t)^\dagger\hat{\mathcal{O}}_S|\lambda_S\rangle\\ &= \hat{\mathcal{U}}(t)^\dagger\lambda|\lambda_S\rangle\\ &= \lambda\hat{\mathcal{U}}(t)^\dagger|\lambda_S\rangle\\ &= \lambda|\lambda_H(t)\rangle, \qquad (5.71)\end{aligned}$$

where we used $\hat{\mathcal{U}}(t)\hat{\mathcal{U}}(t)^\dagger = \hat{\mathcal{I}}$ which is true since $\hat{\mathcal{U}}(t)$ is unitary. So, (5.69) is indeed satisfied with a time-independent eigenvalue (which is the same as in the Schrödinger picture). This is natural because the eigenvalues of an operator are a prediction of the theory – they tell us the possible outcomes of measurements – so they should be the same in both pictures.

The probabilities, in the Heisenberg picture, for the outcome of measuring $\hat{\mathcal{O}}_H$ are found by using the orthonormal basis eigenstates $|\lambda_H(t)\rangle$. Indeed for a system in the state $|\Psi_H\rangle$, the probability that λ is measured is given by $|\langle\lambda_H(t)|\Psi_H\rangle|^2$ and it depends on time in the Heisenberg picture because the eigenstates do. It is easy to see that this agrees with the Schrödinger picture, as

$$\begin{aligned}\langle\lambda_H(t)|\Psi_H\rangle &= \langle(\hat{\mathcal{U}}(t)^\dagger|\lambda_S\rangle)^\dagger, |\Psi_S(0)\rangle\rangle \\ &= \langle\lambda_S|\hat{\mathcal{U}}(t)|\Psi_S(0)\rangle \\ &= \langle\lambda_S|\Psi_S(t)\rangle. \end{aligned} \tag{5.72}$$

5.5.1 Equations of Motion In The Heisenberg Picture

In the Schrödinger picture, the time-evolution of a state is determined by the time-dependent Schrödinger equation. If in the Heisenberg picture the states are time-independent, what replaces the time-dependent Schrödinger equation?

Recall that $\hat{\mathcal{U}}(t) = e^{-i\hat{H}t/\hbar}$, which implies, for a time-independent Hamiltonian $\hat{H}$, that [12]

$$\begin{aligned}\frac{\partial}{\partial t}\hat{\mathcal{U}}(t) &= \frac{\partial}{\partial t}e^{-i\hat{H}t/\hbar} \\ &= -\frac{i\hat{H}}{\hbar}e^{-i\hat{H}t/\hbar} \\ &= -\frac{i\hat{H}}{\hbar}\hat{\mathcal{U}}(t) \\ &= -\hat{\mathcal{U}}(t)\frac{i\hat{H}}{\hbar}. \end{aligned} \tag{5.73}$$

[12] As discussed in section 5.4

Let us apply this to the Heisenberg picture operator $\hat{\mathcal{O}}_H(t)$ defined in equation (5.66),

$$\begin{aligned}
\frac{\partial}{\partial t}\hat{\mathcal{O}}_H(t) &= \frac{\partial}{\partial t}\left(\hat{\mathcal{U}}(t)^\dagger \hat{\mathcal{O}}_S \hat{\mathcal{U}}(t)\right) \\
&= \left(\frac{\partial}{\partial t}\hat{\mathcal{U}}(t)\right)^\dagger \hat{\mathcal{O}}_S \hat{\mathcal{U}}(t) + \hat{\mathcal{U}}(t)^\dagger \hat{\mathcal{O}}_S \frac{\partial}{\partial t}\hat{\mathcal{U}}(t) \\
&= \left(-\hat{\mathcal{U}}(t)\frac{i\hat{H}}{\hbar}\right)^\dagger \hat{\mathcal{O}}_S \hat{\mathcal{U}}(t) - \hat{\mathcal{U}}(t)^\dagger \hat{\mathcal{O}}_S \hat{\mathcal{U}}(t)\frac{i\hat{H}}{\hbar} \\
&= \frac{i\hat{H}}{\hbar}\hat{\mathcal{U}}(t)^\dagger \hat{\mathcal{O}}_S \hat{\mathcal{U}}(t) - \hat{\mathcal{U}}(t)^\dagger \hat{\mathcal{O}}_S \hat{\mathcal{U}}(t)\frac{i\hat{H}}{\hbar} \\
&= \frac{i\hat{H}}{\hbar}\hat{\mathcal{O}}_H(t) - \hat{\mathcal{O}}_H(t)\frac{i\hat{H}}{\hbar} \\
&= \frac{i}{\hbar}[\hat{H}, \hat{\mathcal{O}}_H(t)]. \qquad (5.74)
\end{aligned}$$

Therefore, we conclude that

$$\boxed{\frac{\partial}{\partial t}\hat{\mathcal{O}}_H(t) = \frac{i}{\hbar}[\hat{H}, \hat{\mathcal{O}}_H(t)],} \qquad (5.75)$$

which is known as the Heisenberg equation of motion. In the Heisenberg picture, it replaces the time-dependent Schrödinger equation and gives a differential equation that all Heisenberg picture operators must satisfy.

In classical mechanics, the equation of motion for a function on phase space, or observable, $\mathcal{O} = \mathcal{O}(x, p)$ is given by

$$\frac{d}{dt}\mathcal{O} = -\{H, \mathcal{O}\}. \qquad (5.76)$$

Comparing this to the Heisenberg equation of motion, we see the analogy

$$[\cdot, \cdot] \sim i\hbar\{\cdot, \cdot\}. \qquad (5.77)$$

In other words, the commutator becomes a generalisation of the classical Poisson bracket[13]. We remarked on this previously when we first looked at the time evolution of expectation values in theorem 2.1.

[13]In fact, in an appropriate sense, $[\cdot, \cdot] = i\hbar\{\cdot, \cdot\} + O(\hbar^2)$. Those interested to appreciate better the motivation for terms nonlinear in $\hbar$, you should read about the Moyal product, for example in 'Quantum Mechanics in Phase Space' by Thomas L. Curtright and Cosmas K. Zachos available here: https://arxiv.org/abs/1104.5269

6. The Harmonic Oscillator

"The career of a young theoretical physicist consists of treating the harmonic oscillator in ever-increasing levels of abstraction."

Sidney Coleman

6.1 The Classical Solution

The standard example of a harmonic oscillator is that of a mass, m attached to a spring, of spring constant k, so that the force on the mass is

$$F = -kx. \tag{6.1}$$

This is a one-dimensional system, and Newton's second law of motion applied to this forces gives the second-order differential equation:

$$m\frac{d^2x}{dt^2} = -kx. \tag{6.2}$$

This equation is solved to give

$$\begin{aligned} x(t) &= A\cos(\sqrt{\frac{k}{m}}t) + B\sin(\sqrt{\frac{k}{m}}t) \\ &= A\cos(\omega t) + B\sin(\omega t) \end{aligned} \tag{6.3}$$

where A, B are real constants, and $\omega = \sqrt{\frac{k}{m}}$. The constants are fixed by if two initial conditions are known. We see that the solution oscillates with a period $\tau = 2\pi/\omega = 2\pi\sqrt{\frac{m}{k}}$.

The force is derived from a potential

$$V(x) = \frac{1}{2}kx^2 = \frac{1}{2}m\omega^2x^2 \tag{6.4}$$

so that $F = -\frac{\partial V}{\partial x} = -kx$, and the classical Hamiltonian is

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2. \tag{6.5}$$

The quadratic potential $V = \frac{1}{2}m\omega^2 x^2$ in the harmonic oscillator is a very useful potential to study, as it may be used to approximate a local minimum in any continuous potential. Suppose that $U(x)$ is a potential function, having a local minimum at $x = x_0$, then, so long as $U(x)$ admits a Taylor expansion in the vicinity of x_0, we have

$$\begin{aligned} U(x) &= U(x_0) + (x - x_0)\frac{\partial U(x)}{\partial x}\bigg|_{x=x_0} + \frac{1}{2}(x - x_0)^2\frac{\partial^2 U(x)}{\partial x^2}\bigg|_{x=x_0} + \dots \\ &= U(x_0) + \frac{1}{2}(x - x_0)^2\frac{\partial^2 U(x)}{\partial x^2}\bigg|_{x=x_0} + \dots \end{aligned} \tag{6.6}$$

where we used the fact that $x = x_0$ is a minimum of $U(x)$ and so $\frac{\partial U(x)}{\partial x}\big|_{x=x_0} = 0$. The first term in the expansion is a constant, and so plays no role in the force $F = -\frac{\partial}{\partial x}(U(x))$, and the leading variable term is a term quadratic in $(x - x_0)$, the distance from the minimum of $U(x)$. Further, since we are studying the minimum of a potential, then $\frac{\partial^2 U(x)}{\partial x^2}\big|_{x=x_0}$ is positive. For small $(x - x_0)$, we may approximate $U(x)$ by neglecting the remaining terms of cubic order and higher in $(x - x_0)$. The quadratic potential then is of great use in investigating dynamics near the minimum of any potential, which is, perhaps, why the harmonic oscillator is of such interest.

6.2 The Quantum Solution

The quantum harmonic oscillator is the system with the Hamiltonian operator found by promoting $x \mapsto \hat{x}$ and $p \mapsto \hat{p}$ in the Hamiltonian of the classical harmonic oscillator:

$$\hat{H} = \frac{1}{2}m\omega^2\left(\hat{x}^2 + \frac{\hat{p}^2}{m^2\omega^2}\right). \tag{6.7}$$

There are two standard approaches to solving the quantum harmonic oscillator problem, a chiefly algebraic approach and a more analytical approach.

6.2.1 The Algebraic Approach

Given the expression for $\hat{H}$ above in (6.7), we may be tempted to try to factorise $\hat{H}$ over the complex numbers, however we must take care to remember that $\hat{x}$ and $\hat{p}$ do not commute, as

ordinary numbers or functions do. We can develop the factorisation we are thinking of, and carefully extract the Hamiltonian operator as:

$$\begin{aligned}\frac{1}{2}m\omega^2\left(\hat{x}-\frac{i\hat{p}}{m\omega}\right)\left(\hat{x}+\frac{i\hat{p}}{m\omega}\right) &= \frac{1}{2}m\omega^2\left[\hat{x}^2+\frac{\hat{p}^2}{m^2\omega^2}+\frac{i}{m\omega}\left(\hat{x}\hat{p}-\hat{p}\hat{x}\right)\right] \\ &= \hat{H}+\frac{1}{2}i\omega[\hat{x},\hat{p}] \\ &= \hat{H}-\frac{1}{2}\hbar\omega \end{aligned} \tag{6.8}$$

where we made use of $[\hat{x},\hat{p}]=i\hbar$ to find the final expression. Defining the operator $\hat{V}=\hat{x}+\frac{i\hat{p}}{m\omega}$ then, we have,

$$\hat{H}=\frac{1}{2}\hbar\omega+\frac{1}{2}m\omega^2\hat{V}^\dagger\hat{V}. \tag{6.9}$$

Notice that the order $\hat{V}^\dagger\hat{V}$ matters as:

$$\begin{aligned}[\hat{V}^\dagger,\hat{V}] &= [\hat{x}-\frac{i\hat{p}}{m\omega},\hat{x}+\frac{i\hat{p}}{m\omega}] \\ &= \frac{2i}{m\omega}[\hat{x},\hat{p}] \\ &= -\frac{2\hbar}{m\omega}. \end{aligned} \tag{6.10}$$

Suppose we define $\hat{a}=\sqrt{\frac{m\omega}{2\hbar}}\hat{V}$, then we would have,

$$\boxed{[\hat{a},\hat{a}^\dagger]=1} \tag{6.11}$$

which will be a very useful commutator for our subsequent work. The operator $\hat{a}$ is called the annihilation (or lowering) operator, while $\hat{a}^\dagger$ is called the creation (or raising) operator. Together they are also sometimes called the ladder operators. Explicitly we have,

$$\boxed{\begin{aligned}\hat{a} &= \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x}+\frac{i\hat{p}}{m\omega}\right) \\ \hat{a}^\dagger &= \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x}-\frac{i\hat{p}}{m\omega}\right).\end{aligned}} \tag{6.12}$$

These expression can be inverted to give the position and momentum operators in terms of the creation and annihilation operators:

$$\hat{x}=\sqrt{\frac{\hbar}{2m\omega}}\left(\hat{a}+\hat{a}^\dagger\right) \tag{6.13}$$

$$\hat{p}=-i\sqrt{\frac{m\hbar\omega}{2}}\left(\hat{a}-\hat{a}^\dagger\right) \tag{6.14}$$

and since, by definition, $\hat{V}^\dagger\hat{V} = \frac{2\hbar}{m\omega}\hat{a}^\dagger\hat{a}$ we have

$$\begin{aligned}\hat{H} &= \frac{1}{2}\hbar\omega + \frac{1}{2}m\omega^2(\frac{2\hbar}{m\omega}\hat{a}^\dagger\hat{a}) \\ &= \hbar\omega\left(\frac{1}{2} + \hat{a}^\dagger\hat{a}\right) \\ &\equiv \hbar\omega\left(\frac{1}{2} + \hat{N}\right)\end{aligned} \tag{6.15}$$

where we have defined $\hat{N} \equiv \hat{a}^\dagger\hat{a}$ which is called the number operator.

The number operator is Hermitian as $\hat{N}^\dagger = (\hat{a}^\dagger\hat{a})^\dagger = \hat{N}$, and differs from the Hamiltonian operator by a constant and a coefficient. Therefore the eigenvalues of $\hat{N}$ will determine those of $\hat{H}$, as if $\hat{N}|n\rangle = n|n\rangle$ then

$$\begin{aligned}\hat{H}|n\rangle &= \hbar\omega\left(\frac{1}{2} + \hat{N}\right)|n\rangle \\ &= \hbar\omega(\frac{1}{2} + n)|n\rangle\end{aligned} \tag{6.16}$$

The eigenstates $|n\rangle$ for $\hat{N}$ with $\hat{N}|n\rangle = n|n\rangle$ imply that

$$\begin{aligned}\langle n|\hat{N}|n\rangle &= \langle n|\hat{a}^\dagger\hat{a}|n\rangle \\ &= \|\hat{a}|n\rangle\|^2 \\ &\geq 0\end{aligned} \tag{6.17}$$

and as $\langle n|\hat{N}|n\rangle = n\langle n|n\rangle$ then we have $n\langle n|n\rangle \geq 0$ and further because $\langle n|n\rangle = \||n\rangle\|^2 \geq 0$ then we must have that the eigenvalues of the number operator are positive i.e. $n \geq 0$. Consequently the eigenvalues of $\hat{H}$ which are the energies $\hbar\omega(\frac{1}{2} + n)$ satisfy $E \geq \frac{1}{2}\hbar\omega$. The state with the energy which saturates this bound (which we will soon show exists) is called the ground state, or sometimes the vacuum state, of the harmonic oscillator and has $E = \frac{1}{2}\hbar\omega$.

To find states with energy $E > \frac{1}{2}\hbar\omega$, which will be called excited states, first consider the commutators:

$$\begin{aligned}[\hat{N}, \hat{a}] &= [\hat{a}^\dagger\hat{a}, \hat{a}] \\ &= \hat{a}^\dagger[\hat{a}, \hat{a}] + [\hat{a}^\dagger, \hat{a}]\hat{a} \\ &= -\hat{a}\end{aligned} \tag{6.18}$$

(where we have used $[\hat{a}, \hat{a}^\dagger] = 1$ and that $[\hat{a}, \hat{a}] = 0$) and,

$$\begin{aligned}[\hat{N}, \hat{a}^\dagger] &= [\hat{a}^\dagger\hat{a}, \hat{a}^\dagger] \\ &= \hat{a}^\dagger[\hat{a}, \hat{a}^\dagger] + [\hat{a}^\dagger, \hat{a}^\dagger]\hat{a} \\ &= \hat{a}^\dagger.\end{aligned} \tag{6.19}$$

Now we can prove the more general identities:

$$[\hat{N}, (\hat{a})^k] = -k(\hat{a})^k \tag{6.20}$$
$$[\hat{N}, (\hat{a}^\dagger)^k] = k(\hat{a}^\dagger)^k. \tag{6.21}$$

For the first case, in equation (6.20), we first assume the statement is true for some $k > 1$ and consider

$$\begin{aligned}[\hat{N}, (\hat{a})^{k+1}] &= [\hat{N}, (\hat{a})^k]\hat{a} + (\hat{a})^k[\hat{N}, \hat{a}] \\ &= -k(\hat{a})^k\hat{a} + (\hat{a})^k(-\hat{a}) \\ &= -(k+1)(\hat{a})^{k+1}. \end{aligned} \tag{6.22}$$

The development above shows the inductive step of an induction proof, and the base case (for $k = 1$) is the statement we showed above, namely, $[\hat{N}, \hat{a}] = -\hat{a}$. One can prove the second statement in equation (6.21) in a similar manner.

Now supposing that $|n\rangle$ is an eigenstate of $\hat{N}$ then we have that,

$$\begin{aligned}\hat{N}(\hat{a})^k|n\rangle &= [\hat{N}, (\hat{a})^k]|n\rangle + (\hat{a})^k\hat{N}|n\rangle \\ &= -k(\hat{a})^k|n\rangle + (\hat{a})^k n|n\rangle \\ &= (n-k)(\hat{a})^k|n\rangle \end{aligned} \tag{6.23}$$

and we see that $(\hat{a})^k|n\rangle$ is another eigenstate of $\hat{N}$ with eigenvalue $n-k$, i.e. $(\hat{a})^k|n\rangle \propto |n-k\rangle$. Therefore

$$\hat{H}(\hat{a})^k|n\rangle = \hbar\omega\left(\frac{1}{2} + (n-k)\right)(\hat{a})^k|n\rangle. \tag{6.24}$$

The annihilation operator $\hat{a}$ has lowered the energy of the harmonic oscillator, and each action of $\hat{a}$ has lowered the energy by $\hbar\omega$. This is why $\hat{a}$ is also called the lowering operator.

Similarly we also have,

$$\begin{aligned}\hat{N}(\hat{a}^\dagger)^k|n\rangle &= [\hat{N}, (\hat{a}^\dagger)^k]|n\rangle + (\hat{a}^\dagger)^k\hat{N}|n\rangle \\ &= k(\hat{a}^\dagger)^k]|n\rangle + (\hat{a}^\dagger)^k n|n\rangle \\ &= (n+k)(\hat{a}^\dagger)^k]|n\rangle \end{aligned} \tag{6.25}$$

and so $(\hat{a}^\dagger)^k|n\rangle \propto |n+k\rangle$ and

$$\hat{H}(\hat{a}^\dagger)^k|n\rangle = \hbar\omega\left(\frac{1}{2} + (n+k)\right)(\hat{a}^\dagger)^k|n\rangle. \tag{6.26}$$

The creation operator $\hat{a}^\dagger$ has raised the energy of the harmonic oscillator. Each action of $\hat{a}^\dagger$ has raised the energy by $\hbar\omega$, hence it is also called the raising operator. Together $\hat{a}$ and $\hat{a}^\dagger$ are called ladder operators, and they act to map energy eigenstates onto each other.

There seems to be an infinite tower of eigenstates of the Hamiltonian, but we have seen that the eigenvalue n of the states $|n\rangle$ must satisfy $n \geq 0$. So consider acting repeatedly with $\hat{a}$ on a state $|n\rangle$ with $n > 0$ to generate the sequence of eigenvectors $(\hat{a})^k|n\rangle \propto |n-k\rangle$ with decreasing eigenvalues. At some point the sequence must terminate (when the eigenvalue $n-k \geq 0$ and $n-k-1 < 0$). Let this lowest energy state of energy E_0 be denoted $|\psi_0\rangle$, then since it has the lowest energy state we must have $\hat{a}|\psi_0\rangle = 0$, the zero vector in the Hilbert space. I.e. $\hat{a}$ acts to annihilate the lowest energy state, and this is why it is called the annihilator. Let's expand the equation $\hat{a}|\psi_0\rangle = 0$:

$$\hat{a}\psi_0(x) = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} + \frac{i\hat{p}}{m\omega}\right)\psi_0(x) = \sqrt{\frac{m\omega}{2\hbar}}\left(x + \frac{\hbar}{m\omega}\frac{d}{dx}\right)\psi_0(x) = 0 \tag{6.27}$$

or, equivalently,

$$\frac{d}{dx}\psi_0(x) = -\frac{m\omega x}{\hbar}\psi_0(x). \tag{6.28}$$

We can solve this equation to find

$$\psi_0(x) = Ae^{-\frac{1}{2}m\omega x^2/\hbar} \tag{6.29}$$

for some constant $A \in \mathbb{C}$. We can find A by normalising $\psi_0(x)$, as

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} dx\, |A|^2 e^{-m\omega x^2/\hbar} \\ &= |A|^2\sqrt{\frac{\pi\hbar}{m\omega}} \end{aligned} \tag{6.30}$$

and therefore,

$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} e^{-\frac{1}{2}m\omega x^2/\hbar}. \tag{6.31}$$

We emphasise that the solution is unique, there is only a single eigenstate for the ground state. Having constructed the ground state $\psi_0(x)$ we can now compute its energy E_0 using $\hat{H}|\psi_0\rangle = E_0|\psi_0\rangle$,

$$\begin{aligned} \hat{H}|\psi_0\rangle &= \hbar\omega\left(\frac{1}{2} + \hat{a}^\dagger\hat{a}\right)|\psi_0\rangle \\ &= \frac{1}{2}\hbar\omega|\psi_0\rangle \end{aligned} \tag{6.32}$$

as $\hat{a}|\psi_0\rangle = 0$ by construction. Consequently we read off that the ground state energy is $E_0 = \frac{1}{2}\hbar\omega$, which is the energy of the eigenstate $|n\rangle$ of $\hat{N}$ for $n = 0$, that is $|\psi_0\rangle = |0\rangle$. Note that $|0\rangle$ is not the zero vector of the Hilbert space, it is simply the eigenstate of $\hat{N}$ with eigenvalue 0, which as we see in $|\psi_0\rangle$ is a non-trivial function of x (and not the zero function).

We have identified the unique ground state of the harmonic oscillator, hence we can now list all the eigenstates of $\hat{H}$ by acting with the creation operator to find an infinite set of eigenstates:

$$\boxed{\begin{aligned} |n\rangle &= A_n(\hat{a}^\dagger)^n|0\rangle \quad \text{with} \\ E_n &= \hbar\omega(\tfrac{1}{2}+n) \quad \forall n \in \mathbb{Z}^+, \end{aligned}} \tag{6.33}$$

$$\tag{6.34}$$

where A_n are coefficients determined by normalising the states, and the ground state is

$$\boxed{|0\rangle = \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} e^{-\frac{1}{2}m\omega x^2/\hbar}.} \tag{6.35}$$

In order to normalise the states, and determine A_n, we first focus on the coefficients of proportionality between states of consecutive energies and see what this implies when the states $|n\rangle$ are normalised, in particular let us suppose that

$$\begin{aligned} |n\rangle &= C\hat{a}^\dagger|n-1\rangle \\ &= D\hat{a}|n+1\rangle \end{aligned} \tag{6.36}$$

where C and D are constants. Now we have

$$\begin{aligned} \langle n|n\rangle &= \langle n-1|\hat{a}C^*C\hat{a}^\dagger|n-1\rangle \\ &= |C|^2\langle n-1|\hat{a}\hat{a}^\dagger|n-1\rangle \\ &= |C|^2\langle n-1|\hat{a}\hat{a}^\dagger + (\hat{a}^\dagger\hat{a} - \hat{a}^\dagger\hat{a})|n-1\rangle \\ &= |C|^2\langle n-1|\hat{N} + [\hat{a},\hat{a}^\dagger]|n-1\rangle \\ &= |C|^2\langle n-1|(n-1)+1|n-1\rangle \\ &= |C|^2 n\langle n-1|n-1\rangle. \end{aligned} \tag{6.37}$$

Now suppose that each eigenstate $|n\rangle$ is normalised, then $\langle n|n\rangle = 1 = \langle n-1|n-1\rangle$ and we find $C = 1/\sqrt{n}$, in other words,

$$|n\rangle = \frac{1}{\sqrt{n}}\hat{a}^\dagger|n-1\rangle. \tag{6.38}$$

We can deduce D in a similar way:

$$\begin{aligned} \langle n|n\rangle &= \langle n+1|\hat{a}^\dagger D^* D\hat{a}|n+1\rangle \\ &= |D|^2\langle n-1|\hat{N}|n-1\rangle \\ &= |D|^2(n-1)\langle n-1|n-1\rangle \end{aligned} \tag{6.39}$$

hence, neglecting phase factors, if $|n\rangle$ and $|n-1\rangle$ are both normalised states then we have that $D = 1/\sqrt{n+1}$, i.e.

$$|n\rangle = \frac{1}{\sqrt{n+1}}\hat{a}|n+1\rangle. \tag{6.40}$$

Consequently, by using equation (6.38) repeatedly to reach the normalised ground state, we find

$$\begin{aligned} |n\rangle &= \frac{1}{\sqrt{n}}\hat{a}^\dagger|n-1\rangle \\ &= \frac{1}{\sqrt{n}}\frac{1}{\sqrt{n-1}}\hat{a}^\dagger|n-2\rangle \\ &\vdots \\ &= \frac{1}{\sqrt{n!}}(\hat{a}^\dagger)^n|0\rangle \end{aligned} \tag{6.41}$$

and all the states are normalised.

We have constructed all the normalised states of the quantum harmonic oscillator. There is a single ground state $|0\rangle$ with energy $\frac{1}{2}\hbar\omega$ which is destroyed by the annihilation operator: $\hat{a}|0\rangle = 0$. The action of the creation operator on the ground state creates an infinite tower of excited states. Each action raises the energy by $\hbar\omega$ and generates a set of orthogonal eigenstates (which we have shown how to normalise). In this way the harmonic oscillator gives a concrete way to see that $L^2(\mathbb{R})$ admits a countable, orthonormal basis: in this example they are the bound states in the potential of the harmonic oscillator.

6.2.2 The Analytical Approach

Let us begin again, and see how the same ideas appear from a more analytical approach to solving the Schrödinger equation for the harmonic oscillator; of course we will find all the same results, but they way the details appear will be interesting. Starting from the time-independent Schrödinger equation for a stationary state $\varphi(x)$ with energy $E > 0$, we have

$$-\frac{\hbar^2}{2m}\frac{d^2\varphi}{dx^2} + \frac{1}{2}m\omega^2x^2\varphi = E\varphi \tag{6.42}$$

where we have dropped the notation showing the explicit dependence of φ on x. Let us make a change of variables replacing x with $u = \sqrt{\frac{m\omega}{\hbar}}x$, as then

$$\frac{d}{dx} = \frac{du}{dx}\frac{d}{du} = \sqrt{\frac{m\omega}{\hbar}}\frac{d}{du} \tag{6.43}$$

and so the time-independent Schrödinger equation becomes, after rearrangement,

$$\begin{aligned}\frac{d^2\varphi}{du^2} &= u^2\varphi - \frac{2}{\hbar\omega}E\varphi \\ &\equiv (u^2-\lambda)\varphi \end{aligned} \tag{6.44}$$

where $\lambda \equiv \frac{2E}{\hbar\omega}$. Then, for large u we have,

$$\frac{d^2\varphi}{du^2} \propto u^2\varphi. \tag{6.45}$$

Approximately then, for large u,

$$\varphi(u) \approx Ae^{-u^2/2} + Be^{+u^2/2}, \tag{6.46}$$

as

$$\begin{aligned}\frac{d\varphi(u)}{du} &= -uAe^{-u^2/2} + uBe^{+u^2/2} \\ \frac{d^2\varphi(u)}{du^2} &= -Ae^{-u^2/2} + Be^{+u^2/2} + u^2Ae^{-u^2/2} + u^2Be^{+u^2/2} \\ &= -Ae^{-u^2/2} + Be^{+u^2/2} + u^2\varphi(u) \\ &\approx u^2\varphi(u) \quad \text{for large } u. \end{aligned} \tag{6.47}$$

Further if $\varphi(u) \to 0$ as $u \to \pm\infty$ then we must set $B = 0$, so that asymptotically we anticipate that $\varphi(u) \propto Ae^{-u^2/2}$.

How can we determine a more accurate solution? Let us try

$$\varphi(u) = h(u)e^{-u^2/2} \tag{6.48}$$

where $h(u)$ is a polynomial function of u, so that for large u $e^{-u^2/2}$ dominates in $\varphi(u)$. If this is the case then using

$$\begin{aligned}\frac{d\varphi(u)}{du} &= \frac{dh(u)}{du}e^{-u^2/2} - uh(u)e^{-u^2/2} \\ \frac{d^2\varphi(u)}{du^2} &= \frac{d^2h(u)}{du^2}e^{-u^2/2} - 2u\frac{dh(u)}{du}e^{-u^2/2} - h(u)e^{-u^2/2} + u^2h(u)e^{-u^2/2} \end{aligned} \tag{6.49}$$

in (6.44) gives the requirement that $f(u)$ must satisfy

$$\frac{d^2h(u)}{du^2} - 2u\frac{dh(u)}{du} + (\lambda-1)h(u) = 0. \tag{6.50}$$

In order to think about solving this, let us first see what would result if $h(u)$ were not a polynomial of finite degree, but instead was a power series. We suppose that

$$h(u) = \sum_{k=0}^{\infty} a_k u^k \tag{6.51}$$

giving

$$\frac{dh(u)}{du} = \sum_{k=0}^{\infty} k a_k u^{k-1}$$
$$\frac{d^2h(u)}{du^2} = \sum_{k=0}^{\infty} k(k-1) a_k u^{k-2} \tag{6.52}$$

so that (6.50) becomes

$$\sum_{k=0}^{\infty} k(k-1) a_k u^{k-2} - 2u \sum_{k=0}^{\infty} k a_k u^{k-1} + (\lambda - 1) \sum_{k=0}^{\infty} a_k u^k = 0. \tag{6.53}$$

We can bring all the terms together under a single summation, by redefinition of the summation label in the first term to give

$$\sum_{k=0}^{\infty} \left((k+2)(k+1) a_{k+2} - 2k a_k + (\lambda - 1) a_k \right) u^k = 0. \tag{6.54}$$

This equation is solved if the coefficients a_k satisfy the recursive relation:

$$(k+2)(k+1) a_{k+2} = (2k + 1 - \lambda) a_k. \tag{6.55}$$

This relation implies that

$$\begin{aligned} \lim_{k\to\infty} \left[\frac{a_{k+2}}{a_k} \right] &= \lim_{k\to\infty} \left[\frac{2k+1-\lambda}{(k+2)(k+1)} \right] \\ &= \lim_{k\to\infty} \left[\frac{2 + \frac{1}{k} - \frac{\lambda}{k}}{k + 3 + \frac{2}{k}} \right] \\ &= \frac{2}{k}. \end{aligned} \tag{6.56}$$

Such a behaviour in coefficients is seen in the power series for the function e^{u^2}:

$$\begin{aligned} e^{u^2} &= \sum_{n=0}^{\infty} \frac{1}{n!} u^{2n} \\ &= \sum_{k,\text{even}}^{\infty} \frac{1}{(\frac{k}{2})!} u^k \\ &\equiv \sum_{k,\text{even}}^{\infty} c_k u^k \end{aligned} \tag{6.57}$$

where we defined $k = 2n$ in the second line. This series has coefficients $c_{k+2} = \frac{1}{\left(\frac{k+2}{2}\right)!}$ and $c_k = \frac{1}{\left(\frac{k}{2}\right)!}$ which satisfy

$$\frac{c_{k+2}}{c_k} = \frac{(k/2)!}{((k+2)/2)!} = \frac{m!}{(m+1)!} = \frac{1}{m+1} \tag{6.58}$$

where we have used $k = 2m$, and for large k the ratio of the coefficients above is approximately $1/m = 2/k$, which was the same as for the coefficients in $h(u)$. That is, if $h(u)$ is a power series then it is approximated by e^{u^2} for large u, in which case we would have $\varphi(u) \approx e^{u^2}e^{-u^2/2} = e^{u^2}2$. Such asymptotic behaviour is ruled out by the requirement that $\varphi(u)$ should vanish as $u \to \pm\infty$. Instead it must be that $h(u)$ is a polynomial function of finite degree. In this case the sequence of coefficients must eventually become zero, i.e. for some k we must have that

$$a_{k+2} = \frac{2k+1-\lambda}{(k+1)(k+2)}a_k = 0 \tag{6.59}$$

while $a_k \neq 0$. This is the case if $\lambda = 2k+1$, and then writing $\varphi_k(u) = h_k(u)e^{-u^2/2}$, the functions in $h_k(u)$ in the state of energy E_k is a polynomial of degree k. Recalling that $\lambda = \frac{2E}{\hbar\omega}$ then

$$E = \frac{1}{2}\hbar\omega\lambda = \hbar\omega(k + \frac{1}{2}) \tag{6.60}$$

or if we relabel,

$$E_n \equiv \hbar\omega(n + \frac{1}{2}) \tag{6.61}$$

as we found previously in the algebraic approach to the quantum harmonic oscillator. Now the state with energy E_n is denoted $\varphi_n(u) = h_n(u)e^{-u^2/2}$ and the function $h_n(u)$, a polynomial of degree n in u has coefficients a_j satisfying

$$\begin{aligned} a_{j+2} &= \frac{2j+1-2n+1}{(j+1)(j+2)}a_j = \frac{2(j-n)}{(j+1)(j+2)}a_j, && 0 \leq j \leq n-2 \\ a_{j+2} &= 0 && j > n-2. \end{aligned} \tag{6.62}$$

In particular the first three lowest energy wave-functions are listed in table 6.1. The wave-functions are only determined up to a_0 or a_1 using the recursion relations for the state with energy E_n in (6.62), and a_0 and a_1 may be determined by normalising the wavefunctions. Up to a factor the functions $h_n(u)$ are called the Hermite polynomials, and denoted $H_n(u)$. The

n	Energy E_n	$\varphi_n(u)$	$\varphi_n(x)$
0	$\frac{1}{2}\hbar\omega$	$a_0 e^{-u^2/2}$	$a_0 e^{-\frac{1}{2}m\omega x^2/\hbar}$
1	$\frac{3}{2}\hbar\omega$	$a_1 u e^{-u^2/2}$	$a_1\sqrt{\frac{m\omega}{\hbar}}x e^{-\frac{1}{2}m\omega x^2/\hbar}$
2	$\frac{5}{2}\hbar\omega$	$a_0(-2u^2+1)e^{-u^2/2}$	$a_0(1-\frac{m\omega}{\hbar}x^2)e^{-\frac{1}{2}m\omega x^2/\hbar}$
3	$\frac{7}{2}\hbar\omega$	$a_1(-\frac{2}{3}u^3+u)e^{-u^2/2}$	$a_1\sqrt{\frac{m\omega}{\hbar}}(x-\frac{2}{3}\frac{m\omega}{\hbar}x^3)e^{-\frac{1}{2}m\omega x^2/\hbar}$
4	$\frac{9}{2}\hbar\omega$	$a_0(\frac{4}{3}u^4-4u^2+1)e^{-u^2/2}$	$a_0(1-4\frac{m\omega}{\hbar}x^2+\frac{4}{3}\left(\frac{m\omega}{\hbar}\right)^2 x^4)e^{-\frac{1}{2}m\omega x^2/\hbar}$
5	$\frac{11}{2}\hbar\omega$	$a_1(\frac{4}{15}u^5-\frac{4}{3}u^3+u)e^{-u^2/2}$	$a_1\sqrt{\frac{m\omega}{\hbar}}(x-\frac{4}{3}\frac{m\omega}{\hbar}x^2+\frac{4}{15}\left(\frac{m\omega}{\hbar}\right)^2 x^5)e^{-\frac{1}{2}m\omega x^2/\hbar}$

Table 6.1: The ground state φ_0 and the first few excited states for the quantum harmonic oscillator. The functions are $\varphi_n(u) = h_n(u)e^{-u^2/2}$ and the coefficients in $h_n(u)$ are related to a_0 or a_1 by the recursion relation in equation (6.62).

lowest order Hermite polynomials[1] are:

$$\begin{aligned} H_0(u) &= 1 \\ H_1(u) &= 2u \\ H_2(u) &= 4u^2 - 2 \\ H_3(u) &= 8u^3 - 12u \\ H_4(u) &= 16u^4 - 48u^2 + 12 \\ H_5(u) &= 32u^5 - 160u^3 + 120u \end{aligned} \tag{6.63}$$

and we recognise the functions $h_n(u)$ upto a factor as the Hermite polynomials. There are an infinite set of Hermite polynomial functions, and each one corresponds to an energy eigenstate of the quantum harmonic oscillator. As such we know that they form an orthonormal basis[2] of the Hilbert space $\mathcal{H} = L^2(\mathbb{R})$, which, by construction is countable, implying that $L^2(\mathbb{R})$ is a separable space.

[1] This is the physicist's definition of the Hermite polynomials. Mathematicians often use an alternative, related, definition: $H_n^{\text{math}}(u) = 2^{-n/2}H_n(u/\sqrt{2}) = (-1)^n e^{u^2/2}\partial_u^n(e^{-u^2/2})$ be careful in your reading beyond this course to check which definition is being used.

[2] Such a set of orthogonal polynomial functions is of great interest mathematically, and the interested reader may consult `https://en.wikipedia.org/wiki/Orthogonal_polynomials` where you will see the Hermite polynomials are a class of orthogonal polynomials. The Legendre polynomials are another class of orthogonal polynomials which are of great use in mathematics and mathematical physics.

7. Further Ideas

The topics in the chapter are meant to be an interesting addendum to this first course in quantum mechanics. It is supposed that the interested reader has many natural questions and some objections by now to the ideas of quantum mechanics, and the material here is aimed at solidifying some of these next ideas. Within this chapter you will also find some more explicit discussions of spin and spin states, which properly fits within the set of ideas discussed in the course. All the other ideas, however, should be considered as food for thought.

7.1 Tensor Products

One natural question at the end of a first course in quantum mechanics, typically focussed on problems about a single particle or a single quantum system, is 'what do we do if we have more than one particle or quantum system?' This line of thought opens up a number of interesting ideas. Some which we will not focus on here include the density matrix, and the idea of measuring correspondence between two states. These interesting ideas all rely on an underlying structure called the tensor product.

Suppose we have two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, where the states of two particles/quantum systems reside. Let $\mathcal{H}_S$ denote the Hilbert space of the full system compromised of just the two particles. We would like to develop some notation for a state in $\mathcal{H}_S$ given the states $\Psi_A \in \mathcal{H}_A$ and $\Psi_B \in \mathcal{H}_B$. The following three notations are commonplace[1]

$$|\Psi\rangle_A|\Psi\rangle_B = |\Psi_A\rangle|\Psi_B\rangle \equiv |\Psi_A\rangle \otimes |\Psi_B\rangle \equiv |\Psi_A, \Psi_B\rangle \in \mathcal{H}_S. \tag{7.1}$$

The notation $\otimes$ denotes the tensor product, and Hilbert space $\mathcal{H}_S$ is called the tensor product of $\mathcal{H}_A$ and $\mathcal{H}_B$ and is denoted $\mathcal{H}_S = \mathcal{H}_A \otimes \mathcal{H}_B$. The tensor product space inherits properties

[1]Given time we could probably come up with more ways to write the same thing! The aim is not to bamboozle, just to make sure the reader is wary that most notation will require some interpretation from the reader. The essence is that we are trying to keep track of a pair of states, and each of these notations is sensible.

from $\mathcal{H}_A$ and $\mathcal{H}_B$, and is itself a Hilbert space. We may add elements as

$$(|\Psi_1\rangle_A + |\Psi_2\rangle_A) \otimes |\Phi\rangle_B = |\Psi_1\rangle_A \otimes |\Phi\rangle_B + |\Psi_2\rangle_A \otimes |\Phi\rangle_B$$
$$|\Psi\rangle_A \otimes (|\Phi_1\rangle_B + |\Phi_2\rangle_B) = |\Psi\rangle_A \otimes |\Phi_1\rangle_B + |\Psi\rangle_A \otimes |\Phi_2\rangle_B$$

and scale states as

$$(\alpha|\Psi\rangle_A) \otimes |\Phi\rangle_B = \alpha(|\Psi\rangle_A) \otimes |\Phi\rangle_B) = |\Psi\rangle_A \otimes (\alpha|\Phi\rangle_B) \tag{7.2}$$

for $\alpha \in \mathbb{C}$. Note that this implies that $0 \otimes |\Phi\rangle_B = 0 = |\Psi\rangle_A$. Let us suppose that $|a_i\rangle$ form a basis on $\mathcal{H}_A$ and $|b_j\rangle$ are a basis for $\mathcal{H}_B$. The inner product on $\mathcal{H}_S$ is

$$\langle |a_i\rangle \otimes |b_j\rangle, |a_k\rangle \otimes |b_l\rangle = \langle a_i|a_k\rangle\langle b_j|b_l\rangle. \tag{7.3}$$

Operators $\hat{\mathcal{O}}_A \otimes \hat{\mathcal{O}}_B$ act on states as

$$(\hat{\mathcal{O}}_A \otimes \hat{\mathcal{O}}_B)(|a_i\rangle \otimes |b_j\rangle) = \hat{\mathcal{O}}_A|a_i\rangle \otimes \hat{\mathcal{O}}_B|b_j\rangle \tag{7.4}$$

and matrix components for the operator are

$$\langle a_i| \otimes \langle b_j|\hat{\mathcal{O}}_A \otimes \hat{\mathcal{O}}_B|a_k\rangle \otimes |b_l\rangle \equiv (\hat{\mathcal{O}}_{AB})_{ij,kl}. \tag{7.5}$$

We have defined the basis states

$$|a_i\rangle \otimes |b_j\rangle \tag{7.6}$$

and a general state is

$$\Psi_{AB} = \sum_{i,j} C_{ij}|a_i\rangle \otimes |b_j\rangle \tag{7.7}$$

where, as usual, the probability of $|a_i\rangle \otimes |b_j\rangle$ being the measured state is $|C_{ij}|^2$.

Counting the basis states we deduce $\dim \mathcal{H}_S = \dim \mathcal{H}_A \times \dim \mathcal{H}_B$, and we realise that we can create more states in $\mathcal{H}_S$ than are found from taking a product of a state in $\mathcal{H}_A$ with a state in $\mathcal{H}_B$. For simplicity let us suppose that $\dim \mathcal{H}_A = 2 = \dim \mathcal{H}_B$, and consider the superposition of two states:

$$\alpha\big(|a_1\rangle \otimes |b_1\rangle\big) + \beta\big(|a_2\rangle \otimes |b_2\rangle\big) \in \mathcal{H}_S \tag{7.8}$$

where $\alpha, \beta \in \mathbb{C}$ and are both non-zero, then, we cannot factorise this state into a product of states of the form $|\Psi\rangle_A \otimes |\Phi\rangle_B$, where $|\Psi\rangle_A \in \mathcal{H}_A$ and $|\Phi\rangle_B \in \mathcal{H}_B$. We can see this by first supposing that $|\Psi\rangle_A$ and $|\Phi\rangle_B$ are generic states with

$$|\Psi\rangle_A = \alpha_1|a_1\rangle + \alpha_2|a_2\rangle$$
$$|\Phi\rangle_B = \beta_1|b_1\rangle + \beta_2|b_2\rangle \tag{7.9}$$

then we can compute the form a general product state and write it in the basis of $\mathcal{H}_S$

$$\begin{aligned}|\Psi\rangle_A \otimes |\Phi\rangle_B &= (\alpha_1|a_1\rangle + \alpha_2|a_2\rangle) \otimes (\beta_1|b_1\rangle + \beta_2|b_2\rangle) \\ &= \alpha_1\beta_1|a_1\rangle \otimes |b_1\rangle + \alpha_1\beta_2|a_1\rangle \otimes |b_2\rangle + \alpha_2\beta_1|a_2\rangle \otimes |b_1\rangle + \alpha_2\beta_2|a_2\rangle \otimes |b_2\rangle \\ &= C_{11}|a_1\rangle \otimes |b_1\rangle + C_{12}a_1\rangle \otimes |b_2\rangle + C_{21}|a_2\rangle \otimes |b_1\rangle + C_{22}|a_2\rangle \otimes |b_2\rangle. \end{aligned} \tag{7.10}$$

Hence we can say that a general state $\Psi = \sum_{i,j} C_{ij}|a_i\rangle \otimes |b_j\rangle$ in $\mathcal{H}_S$ factorises if $C_{ij} = \alpha_i\beta_j$. This set of coefficients may be organised in a matrix:

$$C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \tag{7.11}$$

and if Ψ is a product state then $\det(C) = \alpha_1\beta_1\alpha_2\beta_2 - \alpha_2\beta_1\alpha_1\beta_2 = 0$. We do not show it here, but it is the case that if $\det C = 0$ then Ψ is a product state. With this test in mind we can see that the state in (7.8) does not factorise into a product of states.

A product state such as

$$|\Psi\rangle_A \otimes |\Psi\rangle_B = \sum_i \alpha_i|a_i\rangle \otimes \sum_j \beta_j|b_j\rangle \tag{7.12}$$

where $|a_i\rangle$ and $|b_j\rangle$ are the eigenstates of $\hat{\mathcal{O}}_A$ and $\hat{\mathcal{O}}_B$ respectively, allows an independent measurement on each part of the state. If we measure $\hat{\mathcal{O}}_A$, the state tells us that with probability $|\alpha_i|^2$ we will obtain the measurement a_i. When we make the measurement of $\hat{\mathcal{O}}_A$[2] and find the outcome a_i then the product wave-function collapses to $|a_i\rangle \otimes \sum_j \beta_j|b_j\rangle$. A subsequent measurement of $\hat{\mathcal{O}}_B$ gives the result b_j with probability $|\beta_j|^2$, and the wave-function collapses to $|a_i\rangle \otimes |b_j\rangle$, a single basis state in $\mathcal{H}_S$. Each set of outcomes (a_i, b_j) (for any i, j) is possible in this state, and the probability of measuring the specific outcome (a_i, b_j) (for fixed i, j) is $P(a_i, b_j) = P(a_i)P(b_j)$. The measurement of $\hat{\mathcal{O}}_A$ and $\hat{\mathcal{O}}_B$ are independent in the product state. For an entangled state, such as,

$$\alpha\big(|a_1\rangle \otimes |b_1\rangle\big) + \beta\big(|a_2\rangle \otimes |b_2\rangle\big) \tag{7.13}$$

the situation is different. A measurement of a_1 for $\hat{\mathcal{O}}_A$ collapses the entangled wave-function to $|a_1\rangle \otimes |b_1\rangle$, and now there is only one outcome possible is $\hat{\mathcal{O}}_B$ is measured, the outcome must be b_1. Alternatively if we had found the outcome a_2 from measuring $\hat{\mathcal{O}}_A$ then the wave-function collapses to $|a_2\rangle \otimes |b_2\rangle$, and it is now only possible to measure b_2 for $\hat{\mathcal{O}}_B$. A measurement of one operator in $\mathcal{H}_A$ in the entangled state, affects the measurement in $\mathcal{H}_B$. The two wavefunctions

[2]Or, more properly of $\hat{\mathcal{O}}_A \otimes \hat{\mathcal{I}}_B$ on $\mathcal{H}_S$, where $\hat{\mathcal{I}}_B$ is the identity operator on $\mathcal{H}_B$, however we will keep writing the less proper 'measurement of $\hat{\mathcal{O}}_A$' or similar where it is decipherable as it is less cumbersome.

are not independent in an entangled state. Entangled states were first introduced by Einstein, Podolsky and Rosen[3] to exhibit what they interpreted to be the incompleteness of quantum mechanics.

7.2 Spin States

In order to exhibit in detail the ideas of the Einstein, Podolsky and Rosen (or EPR) thought experiment, we will look at one of the simplest examples of a tensor product, that encoding the spin states of two spin $\frac{1}{2}$ particles. The spin operators (as seen in the problem set) are

$$\hat{S}_1 = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{S}_2 = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{S}_3 = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{7.14}$$

and these are related to the Pauli matrices σ_i (of $SU(2)$) as $\hat{S}_i = \frac{\hbar}{2}\sigma_i$, and the labelling $1, 2, 3$ corresponding to the directions associated to the coordinates x, y, z. The eigenvectors of $\hat{S}_i$ are denoted χ_i^+ and χ_i^-, and have eigenvalues $\pm\hbar/2$. It is customary to drop the $\hbar$ and speak of spin plus or minus one-half particles.[4] The explicit eigenvectors are

$$\chi_1^+ = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{and} \qquad \chi_1^- = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}; \tag{7.15}$$

$$\chi_2^+ = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix} \qquad \text{and} \qquad \chi_2^- = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}; \quad \text{and} \tag{7.16}$$

$$\chi_3^+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \chi_3^- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{7.17}$$

Each spin operator is Hermitian, and any of the pairs of eigenstates form a basis for the two-dimensional Hilbert state of the spin states. The pairs χ_i^+ and χ_i^- are orthonormal, which we can confirm using the inner product of $\mathbb{C}^2 = \mathcal{H}$. The simplicity of the eigenstates of $\hat{S}_3$ mean that they are normally taken as a basis for the Hilbert space. The outcome of measuring $\hat{S}_3$, i.e. the spin in the z-direction is always one of the eigenvalues of $\hat{S}_3$, so it is either $\hbar/2$ or $-\hbar/2$, and it is standard practise to call these outcomes spin-up and spin-down respectively, and also to write

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{7.18}$$

[3]'Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?' by A. Einstein, B. Podolsky and N. Rosen, Phys. Rev. 47, 777 (1935).

[4]Examples of spin half particles include electrons, protons, neutrons and quarks are spin 1/2 particles.

Suppose that we have $\Psi = \chi_3^+ = |+\rangle$ and we measure $\hat{S}_1$ for Ψ. We would first need to expand[5] Ψ in terms of the eigenstates of $\hat{S}_1$, to find

$$\Psi = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \left(\chi_1^+ + \chi_1^- \right) \tag{7.19}$$

and we see that in the spin up state that there is a probability of one-half of measuring the spin in the x-direction to be either $\hbar/2$ or $-\hbar/2$, both outcomes are equally likely. Similarly, if we were to measure $\hat{S}_2$ on the state $\Psi = |+\rangle$ we would expand Ψ in the basis χ_2^+ and χ_2^- to find

$$\Psi = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \left(\chi_2^+ + \chi_2^- \right) \tag{7.20}$$

so that again there is a fifty-fifty outcome of measure the spin in the y-direction to be $\hbar/2$ or $-\hbar/2$.

7.2.1 Tensor Products of Spin States

Consider two copies of the spin-state Hilbert space $\mathcal{H}_A \cong \mathbb{C}^2$ and $\mathcal{H}_B \cong \mathbb{C}^2$. In the setting of $\mathcal{H}_A \otimes \mathcal{H}_B$ we have the basis spin up and spin down basis states for each particle:

$$|+\rangle_A|+\rangle_B, \quad |+\rangle_A|-\rangle_B, \quad |-\rangle_A|+\rangle_B, \quad |-\rangle_A|-\rangle_B. \tag{7.21}$$

Each of these basis states is a product state, but, of course, linear sums of them give entangled states. We can, in fact, create a set of basis states which are each entangled. To do this, commence with a normalised entangled state

$$|\Phi_0\rangle_{AB} = \frac{1}{\sqrt{2}} \Big(|+\rangle_A|+\rangle_B + |-\rangle_A|-\rangle_B \Big). \tag{7.22}$$

From this we can create three more, independent, entangled states given by

$$|\Phi_i\rangle_{AB} = (\hat{\mathcal{I}}_A \otimes \sigma_i)|\Phi_0\rangle_{AB} \tag{7.23}$$

where we recall that the Pauli matrices are

$$\hat{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{7.24}$$

[5]We can, of course, make use of the inner products of spin eigenvectors $\langle \chi_1^+ | \chi_2^+ \rangle = \frac{1}{2}(1+i)$, $\langle \chi_1^+ | \chi_2^- \rangle = \frac{1}{2}(1-i)$, $\langle \chi_1^+ | \chi_3^+ \rangle = \frac{1}{\sqrt{2}}$, $\langle \chi_1^+ | \chi_3^- \rangle = \frac{1}{\sqrt{2}}$ to expand the states discussed here. E.g. $\chi_3^+ = \langle \chi_1^+ | \chi_3^+ \rangle \chi_1^+ + \langle \chi_1^- | \chi_3^+ \rangle \chi_1^-$.

which satisfy $\hat{\sigma}_i^2 = \hat{\mathcal{I}}$ for $i \in \{1,2,3\}$, so that, for example,

$$\begin{aligned}
|\Phi_1\rangle_{AB} &= (\hat{\mathcal{I}}_A \otimes \sigma_1)|\Phi_0\rangle_{AB} \\
&= \frac{1}{\sqrt{2}}(\hat{\mathcal{I}}_A \otimes \sigma_1)\Big(|+\rangle_A|+\rangle_B + |-\rangle_A|-\rangle_B\Big) \\
&= \frac{1}{\sqrt{2}}(\hat{\mathcal{I}}_A \otimes \sigma_1)\Big(|+\rangle_A|+\rangle_B\Big) + \frac{1}{\sqrt{2}}(\hat{\mathcal{I}}_A \otimes \sigma_1)\Big(|-\rangle_A|-\rangle_B\Big) \\
&= \frac{1}{\sqrt{2}}\Big(\hat{\mathcal{I}}_A|+\rangle_A\,\sigma_1|+\rangle_B\Big) + \frac{1}{\sqrt{2}}\Big(\hat{\mathcal{I}}_A|-\rangle_A\,\sigma_1|-\rangle_B\Big) \\
&= \frac{1}{\sqrt{2}}\Big(|+\rangle_A|-\rangle_B + |-\rangle_A|+\rangle_B\Big). \qquad (7.25)
\end{aligned}$$

In a similar way we can construct the remaining entangled basis states, altogether we have

$$|\Phi_0\rangle_{AB} = \frac{1}{\sqrt{2}}\big(|+\rangle_A|+\rangle_B + |-\rangle_A|-\rangle_B\big) \qquad (7.26)$$

$$|\Phi_1\rangle_{AB} = \frac{1}{\sqrt{2}}\big(|+\rangle_A|-\rangle_B + |-\rangle_A|+\rangle_B\big) \qquad (7.27)$$

$$|\Phi_2\rangle_{AB} = \frac{i}{\sqrt{2}}\big(|+\rangle_A|-\rangle_B + |-\rangle_A|+\rangle_B\big) \qquad (7.28)$$

$$|\Phi_3\rangle_{AB} = \frac{1}{\sqrt{2}}\big(|+\rangle_A|+\rangle_B - |-\rangle_A|-\rangle_B\big) \qquad (7.29)$$

and these are called the Bell states. One can check that the states are orthogonal, normalised[6] and entangled and so form a set of entangled basis states for the four-dimensional Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$.

7.3 Quantum Teleportation

Let us try to use the collapse of a wave-function upon measurement to "teleport" a state $|\Psi\rangle_Q$ from Alice (A) to Bob (B). A qubit is a two state quantum system, we will take the example of a spin $\frac{1}{2}$ particle (e.g. an electron) as such a qubit. Our aim will be to pass the qubit:

$$|\Psi\rangle_Q = \alpha|+\rangle_Q + \beta|-\rangle_Q \qquad (7.30)$$

for some $\alpha, \beta \in \mathbb{C}$ describing the spin state of the particle, from A to B. To do this we will first create[7] a separate, entangled pair of spin-$\frac{1}{2}$ particles. One of the entangled pair will be

[6] Since σ_1 are unitary operators, they must act to preserve the norm of $|\Phi_0\rangle$, which was normalised.

[7] For example, in the decay of a neutral π meson, or pion, into an electron and positron, it happens that the spin of the particle's involved is conserved. The pion is spin zero, and the spins of the electron and positron

in Alice's possession and the other will be with Bob. The idea is that Alice and Bob will be able to interact with the entangled state, and perhaps we can make use of the collapse of its wave-function to transfer the information in the qubit's wave-function. Let us suppose we can arrange for the entangled pair of particles to be in the Bell state

$$|\Phi_0\rangle_{AB} = \frac{1}{\sqrt{2}}\big(|+\rangle_A|+\rangle_B + |-\rangle_A|-\rangle_B\big). \tag{7.31}$$

The total state, including the qubit is in the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_Q$, and is

$$\begin{aligned}|\Phi_0\rangle_{AB} \otimes |\Psi\rangle_Q &= \frac{1}{\sqrt{2}}\Big(|+\rangle_A|+\rangle_B + |-\rangle_A|-\rangle_B\Big) \otimes \Big(\alpha|+\rangle_Q + \beta|-\rangle_Q\Big) \\ &= \frac{1}{\sqrt{2}}\Big(\alpha|+\rangle_A|+\rangle_Q|+\rangle_B + \beta|+\rangle_A|-\rangle_Q|+\rangle_B \\ &\qquad + \alpha|-\rangle_A|+\rangle_Q|-\rangle_B + \beta|-\rangle_A|-\rangle_Q|-\rangle_B\Big).\end{aligned} \tag{7.32}$$

We can write this state in terms of the entangled Bell basis states, but now for the AQ system - we will see why this will be useful in a moment. A little bit of linear algebra on the Bell states in (7.29), give us,

$$\begin{aligned}|\Phi_0\rangle_{AB} \otimes |\Psi\rangle_Q &= \frac{1}{\sqrt{2}}\Big[\frac{\alpha}{\sqrt{2}}\Big(|\Phi_0\rangle_{AQ} + |\Phi_3\rangle_{AQ}\Big)|+\rangle_B + \frac{\beta}{\sqrt{2}}\Big(|\Phi_1\rangle_{AQ} - i|\Phi_2\rangle_{AQ}\Big)|+\rangle_B \\ &\qquad + \frac{\alpha}{\sqrt{2}}\Big(|\Phi_1\rangle_{AQ} + i|\Phi_2\rangle_{AQ}\Big)|-\rangle_B + \frac{\beta}{\sqrt{2}}\Big(|\Phi_0\rangle_{AQ} - |\Phi_3\rangle_{AQ}\Big)|-\rangle_B\Big] \\ &= \frac{1}{2}|\Phi_0\rangle_{AQ}\Big(\alpha|+\rangle_B + \beta|-\rangle_B\Big) + \frac{1}{2}|\Phi_1\rangle_{AQ}\Big(\beta|+\rangle_B + \alpha|-\rangle_B\Big) \\ &\qquad + \frac{1}{2}|\Phi_2\rangle_{AQ}\Big(-i\beta|+\rangle_B + i\alpha|-\rangle_B\Big) + \frac{1}{2}|\Phi_3\rangle_{AQ}\Big(\alpha|+\rangle_B - \beta|-\rangle_B\Big) \\ &= \frac{1}{2}|\Phi_0\rangle_{AQ}|\Psi\rangle_B + \frac{1}{2}|\Phi_1\rangle_{AQ}\sigma_1|\Psi\rangle_B + \frac{1}{2}|\Phi_2\rangle_{AQ}\sigma_2|\Psi\rangle_B + \frac{1}{2}|\Phi_3\rangle_{AQ}\sigma_3|\Psi\rangle_B \\ &= \frac{1}{2}\sum_{i=0}^{3}|\Phi_i\rangle_{AQ}\sigma_i|\Psi\rangle_B\end{aligned} \tag{7.33}$$

where we have defined $|\Psi\rangle_B = \alpha|+\rangle_B + \beta|-\rangle_B$ in the penultimate line, and $\sigma_0 \equiv \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$, and σ_1, σ_2, σ_3 are the Pauli matrices in the last line.

must be opposite. Spin is not always a conserved quantity in particle decays, but angular momentum is always conserved, and angular momentum is contributed to by both spin (intrinsic angular momentum) and orbital angular momentum.

If A can now measure $|\Psi\rangle_{AQ}$ in the entangled basis and suppose the outcome is that the wave-function must be in state $|\Phi_i\rangle_{AQ}$ for some $i \in \{0, 1, 2, 3\}$, then the collapse of the wave-function means that the part of the wavefunction remaining in $\mathcal{H}_B$ is $\sigma_i|\Psi\rangle_B$. Finally if Bob is informed of the outcome of Alice's measurement, they can deduce the form of the wave-function and if they then act on with the unitary operator σ_i, then they find the state

$$\sigma_i^2|\Psi\rangle_B = |\Psi\rangle_B = \alpha|+\rangle_B + \beta|-\rangle_B. \tag{7.34}$$

In other words, the qubit whose information is encoded in α and β has been teleported to Bob. Bob can then interrogate the state to deduce α and β. Notice that Alice does not at any point measure the qubit or deduce α or β. If Alice had interrogated $\Psi\rangle_Q$ then the wave-function would have collapsed. In principle A and B can be any distance apart, and the wave-function collapse would seem to offer the opportunity to communicate information at a speed greater than that of light. However for B to decode the qubit, they needed to know the outcome of A's measurement of the entangled state, and A then had to communicate that information to B (at best at the speed of light). You might now reasonably point out, that A could instead have interrogated $|\Psi\rangle_Q$ directly and sent the results of their information to B, the process would have been just as quick. However, in the transfer of information, especially secret messages, sending a message is open to interception at any point along its journey. The teleportation of information, allows A to send a neutral message (the result of measuring $|\Psi\rangle_{AQ}$ in the entangled basis) whose interception will not reveal anything about the information in the qubit to anyone who does not have access to B's entangled particle. So it is a more secure method to share information.

This discussion hopefully highlights the difference between the instantaneous collapse of the wave-function, and the ability to extract meaningful information from the collapse without a key. This means that wave-function collapse cannot be used to send instantaneous signal, such a possibility would lead to causality paradoxes. For example suppose A and B share their entangled pair of particles whose spins are opposite, but otherwise unknown. If A wishes to communicate information to B, via the entangled particles, then A must also tell B what the outcome of their measurement is. For example suppose A wishes to securely pass a message to B, first A has to measure the spin of their particle, if it is up they know B will find their particle is spin down and vice versa. Then, informed of the outcome B will find when they measure the spin of their particle, A can send a key to B: if you find spin up do Y, if find spin down do X. Now A will know how B should act, but until B receives the key from A the result of the measurement is meaningless for them. Note that A can't agree the key in advance with B, because until they do the measurement they do not know what outcome B will find. So despite the instantaneous wave-function collapse, it does not offer a way to send messages

superluminally.

Quantum teleportation was successfully demonstrated in 1997 by two research groups led by Sandu Popescu and Anton Zeillinger, and in 2017 Jian-Wei Pan's team successfully teleported a photon's quantum state over a distance of 1400km from Tibet to a satellite.

7.4 The No-Cloning Theorem

Given a wave-function $|\Psi\rangle$ suppose that it was possible to instantly clone it as many times as we would like. We could then proceed to make a series of different measurements on each copy of $|\Psi\rangle$, and each time the state would collapse to an eigenstate of whatever observable we are measuring. With sufficient measurement we may aim to discover significant detail of the original wave-function prior to its collapse. This is already a worrying scenario, as we would be able to obtain a situation in which we could measure simultaneously two non-commuting variables to sufficient accuracy (with enough sampling of enough copies of $|\Psi\rangle$) to violate the uncertainty principle. Indeed it is also reasonable, should cloning be possible, to use accurate knowledge of a state, to be able to agree on a key between Alice and Bob so that they may effect instantaneous communication. In this scenario, Alice and Bob would build up a reasonably likely picture of their cloned, entangled state, it would not be perfect, all that would be needed is that to a certain confidence level Bob can deduce if a measurement has caused the wave-function to collapse. In such a scheme the wave-function's collapse would signal to Bob that Alice has made a measurement - in order to be able to do this Bob has to be able to deduce a change in the wave-function, and with sufficient copies of the wave-function he would be able to make enough measurements to be reasonably confident that a collapse of the wave-function has occurred. The collapse of the wave-function over a period would be used in the same way that morse code to communicate a message faster than the speed of light. Such a scenario would admit violations of causality. Let us imagine the extreme scenario in which it is possible to send instant messages between two distant locations. Suppose A sends a birthday card to B, and B instantly replies with a thank-you for the card message to A, and each communication is instantaneous. To someone, Charlie, who is moving towards B at any speed, they see B receive the card and send their thank-you before A sent the birthday card.

There is a simple proof that no universal cloning operation exists, this result is the so-called no-cloning theorem.[8] Suppose $\hat{\mathcal{U}}$ is a unitary map on $\mathcal{H} \otimes \mathcal{H}$ that can act as

$$\hat{\mathcal{U}} : |\Psi_1\rangle \otimes |\Phi\rangle \rightarrow e^{i\theta_1}|\Psi_1\rangle \otimes |\Psi_1\rangle \tag{7.35}$$

[8]Originally proved in the following two papers both published in 1982: 'A Single Quantum Cannot be Cloned' by W. Wootters and W. Zurek; and 'Communication by EPR devices' by D. Dieks.

for normalised states $|\Psi_1\rangle$ and $|\Phi\rangle$. The operator $\hat{\mathcal{U}}$ is acting to make a clone of $|\Psi_1\rangle$ on the by-stander state $|\Phi\rangle$ up to a phase $e^{i\theta_1}$. Now we ask if $\hat{\mathcal{U}}$ can also clone another state $|\Psi_2\rangle$. For the sake of argument, let's suppose it can and we also have:

$$\hat{\mathcal{U}} : |\Psi_2\rangle \otimes |\Phi\rangle \to e^{i\theta_2}|\Psi_2\rangle \otimes |\Psi_2\rangle. \tag{7.36}$$

Since unitary operators preserve the inner product, we have,

$$\begin{aligned} \langle |\Psi_1\rangle \otimes |\Phi\rangle, |\Psi_2\rangle \otimes |\Phi\rangle \rangle &= \langle \Psi_1|\Psi_2\rangle \\ &= \langle e^{i\theta_1}|\Psi_1\rangle \otimes |\Psi_1\rangle, e^{i\theta_2}|\Psi_2\rangle \otimes |\Psi_2\rangle \rangle \\ &= \langle \Psi_1|\Psi_2\rangle^2 e^{i(\theta_2-\theta_1)} \end{aligned} \tag{7.37}$$

where we have used that $\langle\Phi|\Phi\rangle = 1$. Consequently, either $\langle\Psi_1|\Psi_2\rangle = 0$ or $\langle\Psi_1|\Psi_2\rangle = e^{-i(\theta_2-\theta_1)}$. If the second possibility holds then the Cauchy-Schwarz inequality implies

$$|\langle\Psi_1|\Psi_2\rangle|^2 \leq \||\Psi_1\rangle\|^2 \||\Psi_2\rangle\|^2 = 1 \tag{7.38}$$

because $\langle\Psi_1|\Psi_1\rangle = 1 = \langle\Psi_2|\Psi_2\rangle$, and since the saturation of the inequality implies that the two vectors involved are proportional, then we have $|\Psi_2\rangle = e^{-i(\theta_2-\theta_1)}|\Psi_1\rangle$. In this case, $|\Psi_1\rangle$ and $|\Psi_2\rangle$ differ only by a phase and so are physically identical. As a result we have that $\hat{\mathcal{U}}$ can only clone states, representing different physical systems, which are orthogonal. For a finite-dimensional Hilbert space, $\hat{\mathcal{U}}$ can only clone $\dim \mathcal{H}$ states: there does not exist a universal cloning operator which can clone every state. Even prior to the argument above one realises that there are technical problems: one cannot probe the state to be cloned without collapsing the wave-function. There may seem to be some wiggle room in an infinite-dimensional Hilbert space, however, since most states are not mutually orthogonal the same issues arise. We may wonder if we cannot allow some errors up to a set of measure zero in the wave-function, but since the inner product is not sensitive to such a difference, there is no scope for such a loophole. However, if we concede partial defeat, we can still aim for approximate cloning[9], in which imperfect copies of a state are made.

7.5 The EPR Thought Experiment and Bell's Inequality

Einstein, Podolsky and Rosen (EPR) argued in 1935 that quantum mechanics must be incomplete. In making their argument they introduced the idea of an entangled state[10], and in

[9] See for example the pioneering work 'Quantum copying: Beyond the no-cloning theorem' by V. Bužek and M. Hillery, published in 1996.

[10] Although it was Schrödinger who coined the term 'entanglement', and called it the characteristic trait of quantum mechanics.

their original thought experiment they considered a pair of particles entangled in position and momentum. Sometime later David Bohm proposed a version of the EPR thought experiment using two spin-$\frac{1}{2}$ particles. In Bohm's variant the entangled state is the singlet state[11]:

$$|\Phi_2\rangle_{AB} = \frac{1}{\sqrt{2}}\big(|+\rangle_A|-\rangle_B - |-\rangle_A|+\rangle_B\big) \tag{7.39}$$

so that if A measures spin $\hbar/2$ then it is known that B will subsequently measure $-\hbar/2$ as the spin of their particle in the entangled pair, and vice-versa. The results are correlated.

The Copenhagen interpretation of quantum mechanics says that the state instantaneously collapses as a result of a measurement, and go further in saying that the physical quantities do not exist prior to the measurement. The probabilistic essence of quantum mechanics is not to be regarded, say the proponents of the Copenhagen interpretation, as evidence of our ignorance of the system, it is simply inherent, and this is how the world is. The word reality is frequently brought to these interpretations, and Einstein, Podolsky and Rosen wrote that

> "... if, without in any way disturbing a system, we can predict with certainty the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity "

whereby the word reality is used to mean that there exists some unequivocal description of a state which can be used to infer measurements prior to experiment. The opposing view, which in this context may be described as saying there is no such underlying reality, is the orthodox quantum view, that the wave-function is the fundamental description of a system and it only informs us of the probabilities of an outcome[12]. Part of the willingness to accept the orthodox view is that one can understand that in making a measurement one is interfering with the physical system, as in Heisenberg's example of measuring the location of a particle where to be determine the location photons of higher energies are used, and this higher energy effects the momentum of the particle whose position is being measured. But EPR had prepared a quantum state in which a measurement at one. location, A, affects the system at some other location B. Indeed if B is located sufficiently far away, experiments can be done which prove the quantum correlation of results which could not be communicated from A to B by any mechanism at the speed of light. This affecting of distant experiments without a means to send information, at light speed or slower, to correlate outcomes between the points is called

[11]So called, because in the representation theory of $SU(2)$ the three entangled states $|\Phi_0\rangle$, $|\Phi_1\rangle$ and $|\Phi_3\rangle$ are mapped to each other under a simple action of the algebra, and so are called the triplet, while $|\Phi_2\rangle$ remains in a set of its own under the same action, and so is called the singlet.

[12]For example, Heisenberg, a proponent of this view, wrote of an electron's motion that 'the path comes into existence only when we observe it'.

non-locality. EPR did not think that quantum mechanics was incorrect, after all, it had proven so successful already by the time of their paper, just that it was incomplete, and that there was. a possibility that with further structure such issues as non-localism and a lack of reality could be resolved. To put it another way, they wonder whether a local reality might somehow be possible, and perhaps better than the alternative.

The extra structure imagined by EPR is simply called some hidden variable of the system (some feature of reality that we do not yet know about) and its existence allows for an underlying local reality in physics. If the full theory existed and were known, one could abandon working with probabilities - we would understand that the statistics were simply governing our ignorance of the hidden variable in the theory. Armed with such a variable we would be able to take a measurement of state with definite properties prior to measurement, and we would find definite outcomes.

In 1964, John S. Bell, working with only the minimal set of assumptions about a hidden variable theory, came up with an inequality that must hold in such a theory, but which is violated by quantum mechanics. The inequality presented a way to differentiate between a hidden variable theory with local reality and a theory such as quantum mechanics that we have presented in these lectures. The inequalities were put to experimental test and found that they were violated, meaning that no hidden variable theory was possible as a completion of quantum mechanics. Either the assumptions of a reality or of locality in physics have to be discarded. In 2022 the lead scientists in the various experiments testing Bell's inequality, Alain Aspect, John Clauser, and Anton Zeilinger, were awarded the Nobel prize for physics for "establishing the violation of Bell's inequalities."

To describe Bell's inequalities we should develop a little our thoughts about measurements of the spin of a spin-$\frac{1}{2}$ particle. We have thought about the outcome of measuring spin along the x, y and z axes, but Bell proposed the use of a measurement of spin along arbitrary axes: Alice would measure spin along the axis parallel to $\boldsymbol{a}$ and Bob would measure spin along the axis parallel to $\boldsymbol{b}$. In earlier versions of the thought experiment Alice and Bob would measure spins along the z-axis, and due to the entangled state (given in equation (7.39)) then if Alice measured spin up, Bob would measure spin down and vice-versa: the measurements are correlated. If Alice and Bob take measurements along different axes, while they will still measure either $+\hbar/2$ or $-\hbar/2$, their results will no longer be correlated. For example is A measures $\hbar/2$ along the z-axis, then the state (7.39) collapses:

$$|\Phi_2\rangle_{AB} \to |-\rangle_B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}_B = \frac{1}{\sqrt{2}}(\chi_2^+ + \chi_2^-)_B \tag{7.40}$$

and if B then measures the spin of their particle along the y-axis (using the $\hat{S}_2$ operator) then

he has a probability of a half of finding the outcome is $+\hbar/2$ and a half of finding $-\hbar/2$. A hidden variable theory should not disagree with quantum mechanics but it should reproduce its predictions. In the quantum mechanical picture the spins in the entangled state are correlated when $\boldsymbol{a}$ and $\boldsymbol{b}$ are parallel, but then the correlation weakens as the angle between the two measurements increases from zero to $\pi/2$ (as we will see below). A hidden variable theory has to reproduce these results: in this setting any correlation between spins in the two particles must be fixed before they separate and travel to A and B, there is no room for the outcomes measured by B to change due to the measurements made by A.

Bell's proposal was to calculate the expectation value of the product of the outcomes for measurements along directions $\boldsymbol{a}$ and $\boldsymbol{b}$ for a hidden variable theory. The expectation value for measurement of spin along these directions will be denoted

$$E(\boldsymbol{a},\boldsymbol{b}). \tag{7.41}$$

The quantum mechanical expectation value is calculated using the entangled state $|\Phi_2\rangle$. For example if $\boldsymbol{a}=\boldsymbol{b}=\boldsymbol{e}_z$ then as $|\Phi_2\rangle$ is entangled the distinct outcomes are always of the opposite sign (we will take minus one to denote spin down, and plus one to denote spin up). Consequently $E(\boldsymbol{e}_z,\boldsymbol{e}_z)=-1$, and so also $E(\boldsymbol{e}_z,-\boldsymbol{e}_z)=1$. In general, $E(\boldsymbol{a},\boldsymbol{b})=-\boldsymbol{a}\cdot\boldsymbol{b}$, and it depends only on the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$. To see this, let us take $\boldsymbol{a}=\boldsymbol{e}_z$, and so $\hat{S}^A_{\boldsymbol{a}}=\hat{S}^A_3$ while if $\boldsymbol{b}$ is at angle θ to the z-axis,

$$\hat{S}^B_{\boldsymbol{b}}=(\cos\theta)\hat{S}^B_3+(\sin\theta)\hat{S}^B_1. \tag{7.42}$$

Then,

$$\begin{aligned}
\hat{S}^A_{\boldsymbol{a}}\otimes\hat{S}^B_{\boldsymbol{b}}|\Phi_2\rangle_{AB} &= \frac{1}{\sqrt{2}}\hat{S}^A_3((\cos\theta)\hat{S}^B_3+(\sin\theta)\hat{S}^B_1)\big(|+\rangle_A|-\rangle_B-|-\rangle_A|+\rangle_B\big)\\
&= \frac{\hbar}{2\sqrt{2}}\Big(\cos\theta|+\rangle_A\hat{S}^B_3|-\rangle_B+\sin\theta|+\rangle_A\hat{S}^B_1|-\rangle_B\\
&\qquad +\cos\theta|-\rangle_A\hat{S}^B_3|+\rangle_B+\sin\theta|-\rangle_A\hat{S}^B_1|+\rangle_B\Big)\\
&= \frac{\hbar^2}{4\sqrt{2}}\Big(-\cos\theta|+\rangle_A|-\rangle_B+\sin\theta|+\rangle_A|+\rangle_B\\
&\qquad +\cos\theta|-\rangle_A|+\rangle_B+\sin\theta|-\rangle_A|-\rangle_B\Big)\\
&= \frac{\hbar^2}{4\sqrt{2}}\Big(\cos\theta\big(-|+\rangle_A|-\rangle_B+|-\rangle_A|+\rangle_B\big)\\
&\qquad +\sin\theta\big(|+\rangle_A|+\rangle_B+|-\rangle_A|-\rangle_B\big)\Big)\\
&= \frac{\hbar^2}{4\sqrt{2}}\big(-\cos\theta|\Phi_2\rangle_{AB}+\sin\theta|\Phi_0\rangle_{AB}\big)
\end{aligned} \tag{7.43}$$

where we used $\hat{S}_1|+\rangle = \frac{\hbar}{2}|-\rangle$ and $\hat{S}_1|-\rangle = \frac{\hbar}{2}|+\rangle$. Consequently we can compute the expectation value of the two measurements for the entangled state $|\Phi_2\rangle_{AB}$. For simplicity we have adopted to measure spin in this discussion as either $+1$ or -1, this means we should scale our spin operators by $2/\hbar$ (so that our spin operators become the Pauli spin matrices) and so, recalling that the entangled Bell states are orthonormal, we find

$$\begin{aligned} E(\boldsymbol{a},\boldsymbol{b}) &= \frac{4}{\hbar^2}\langle\Phi_2|_{AB}\hat{S}^A_{\boldsymbol{a}} \otimes \hat{S}^B_{\boldsymbol{b}}|\Phi_2\rangle_{AB} \\ &= -\cos\theta \\ &= -\boldsymbol{a}\cdot\boldsymbol{b} \end{aligned} \tag{7.44}$$

as claimed above.

Now we turn to how we would consider the same computation for a completion of quantum theory using a hidden variable λ. In this case we define the outcomes measured by Alice and Bob, along directions $\boldsymbol{a}$ and $\boldsymbol{b}$ to be functions dependent on λ:

$$A(\boldsymbol{a},\lambda) \in \{+1,-1\} \qquad \text{and} \qquad B(\boldsymbol{b},\lambda) \in \{+1,-1\}. \tag{7.45}$$

Then our assertion that the two measurements are independent means that the probability that Alice measures A and Bob measures B can be factorised as

$$P(A,B,\boldsymbol{a},\boldsymbol{b},\lambda) = P(A,\boldsymbol{a},\lambda)P(B,\boldsymbol{b},\lambda). \tag{7.46}$$

Now since λ determines the specific outcome, then for example $P(1,-1,\boldsymbol{a},\boldsymbol{b})$ will either equal 1 if λ informs us that this will be the outcome or zero if not. Now λ is determined according to a probability distribution $\rho(\lambda)$ set at the point where the pair of particles became entangled and gives, under the integral, the probabilities of outcomes A and B. Hence the expectation value is

$$E(\boldsymbol{a},\boldsymbol{b}) \equiv \int d\lambda\rho(\lambda)A(\boldsymbol{a},\lambda)B(\boldsymbol{b},\lambda) \tag{7.47}$$

and $\int d\lambda\rho(\lambda) = 1$ as ρ is a probability density. Now we know that if Alice and Bob makes measurements along the same direction then $A(\boldsymbol{a},\lambda) = -B(\boldsymbol{a},\lambda)$, hence we can write

$$E(\boldsymbol{a},\boldsymbol{b}) \equiv -\int d\lambda\rho(\lambda)A(\boldsymbol{a},\lambda)A(\boldsymbol{b},\lambda). \tag{7.48}$$

Now for any third direction $\boldsymbol{c}$, we have

$$\begin{aligned} E(\boldsymbol{a},\boldsymbol{b}) - E(\boldsymbol{a},\boldsymbol{c}) &= -\int d\lambda\rho(\lambda)\left(A(\boldsymbol{a},\lambda)A(\boldsymbol{b},\lambda) - A(\boldsymbol{a},\lambda)A(\boldsymbol{c},\lambda)\right) \\ &= -\int d\lambda\rho(\lambda)\left(1 - A(\boldsymbol{c},\lambda)A(\boldsymbol{b},\lambda)\right)A(\boldsymbol{b},\lambda)A(\boldsymbol{a},\lambda) \end{aligned} \tag{7.49}$$

where we have used that since $A(\boldsymbol{a},\lambda) \in \{-1,+1\}$ then $(A(\boldsymbol{a},\lambda))^2 = 1$. Furthermore we know that $-1 \leq A(\boldsymbol{a},\lambda)A(\boldsymbol{b},\lambda) \leq 1$ and also $(1 - A(\boldsymbol{c},\lambda)A(\boldsymbol{b},\lambda)) \geq 0$, giving,

$$\begin{aligned} |E(\boldsymbol{a},\boldsymbol{b}) - E(\boldsymbol{a},\boldsymbol{c})| &\leq \int d\lambda\rho(\lambda)\,(1 - A(\boldsymbol{c},\lambda)A(\boldsymbol{b},\lambda)) \\ &= 1 + E(\boldsymbol{c},\boldsymbol{b}). \end{aligned} \tag{7.50}$$

This is Bell's famous inequality:

$$\boxed{|E(\boldsymbol{a},\boldsymbol{b}) - E(\boldsymbol{a},\boldsymbol{c})| \leq 1 + E(\boldsymbol{c},\boldsymbol{b}).} \tag{7.51}$$

We can see that it is not consistent with the expectation value from quantum mechanics $E(\boldsymbol{a},\boldsymbol{b}) = -\boldsymbol{a}\cdot\boldsymbol{b}$ which we found above for the entangled state $|\Phi_2\rangle_{AB}$, by a simple choice for the three vectors. Suppose that

$$\boldsymbol{a} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \text{ and } \boldsymbol{c} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \tag{7.52}$$

then $E(\boldsymbol{a},\boldsymbol{b}) = -\boldsymbol{a}\cdot\boldsymbol{b} = 0$, $E(\boldsymbol{a},\boldsymbol{c}) = -\frac{1}{\sqrt{2}} = E(\boldsymbol{c},\boldsymbol{b})$, and Bell's inequality is clearly violated. In (7.51) the left-hand side becomes $\frac{1}{\sqrt{2}} \sim 0.707$ while the right-hand side is $1 - \frac{1}{\sqrt{2}} \sim 0.293$, and the inequality is violated.

The experimental verification that Bell's inequality is violated means, in the words of John S. Bell that

> "if a hidden variable theory is local it will not agree with quantum mechanics, and if it agrees with quantum mechanics it will not be local."

Einstein didn't live long enough to have to confront this vivid choice, to have to accept either what he termed 'spooky action at a distance' or to give up reality, or, perhaps worse, to have to surrender both. There are other interpretations of quantum mechanics besides the Copenhagen interpretation, for example the de Broglie-Bohm pilot wave theories (which are non-local), but the orthodox picture of quantum mechanics has proven incredibly successful and durable, and its ideas are at the heart of the modern computer-based world. Perhaps one day there will be a resolution to the parts of quantum mechanics that we may find unsettling, but until then perhaps it is worth taking the quote of N. David Mermin about quantum mechanics to heart

> "Shut up and calculate!"

For there is much to learn, revere and to marvel at in carrying out the computations of quantum mechanics.

Appendices

A. Waves and Fourier Transforms

A.1 Travelling Waves and Wave Equations

A wave equation is any equation that admits waves as solutions! In one dimension the most famous wave equation is:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \tag{A.1}$$

where $u = u(x,t)$ is a time-dependent function that encodes the amplitude of a wave at x and at time t, and c is a constant. It is a linear, homogeneous equation, and so any two solutions may be added to give a third solution. Any function of $(x - ct)$ or $(x + ct)$ solves (A.1), i.e. let

$$u(x,t) = f(x - ct) + g(x + ct) \tag{A.2}$$

where f and g are at least twice-differentiable functions, then we may check that

$$\frac{\partial^2 u}{\partial t^2} = c^2(f''(x - ct) + g''(x + ct)) \tag{A.3}$$

$$c^2 \frac{\partial^2 u}{\partial x^2} = c^2(f''(x - ct) + g''(x + ct)) \tag{A.4}$$

and so the wave equation is solved for a very general set of functions.

Suppose that $f(x - ct)$ is a smooth function, and consider its profile at $t = 0$. The initial profile $f(x, 0)$ is maintained under the evolution encoded in the wave equation, i.e. it does not change shape - it is called a travelling wave (the same holds for $g(x + ct)$). In time t the travelling wave $f(x, 0)$ is translated by ct. Fixing a spot at $x = x_0$ on the wave profile $f(x_0, 0)$ we see that at a later time $t = t_1$ the point we aer watching on f is now at x where $x - ct_1 = x_0$, i.e. $x = x_0 + ct_1$ and we can deduce that the wave profile has moved at constant speed c in the positive x-direction. That is, $f(x - ct)$ is a travelling wave, of fixed profile moving at speed c in the positive x-direction. If had made the same investigation for the wave profile $g(x + ct)$ we would deduce that $g(x + ct)$ is a travelling wave moving in the negative x-direction at speed c. We say that $f(x - ct)$ and $g(x + ct)$ are right and left-moving travelling waves.

Let us focus on waves moving to the right: $u(x,t) = f(x-ct)$. The function, so long as it is twice differentiable, may take any form, when we choose the function to sine or cosine then the travelling waves are called harmonic waves. Let us write

$$u(x,t) = f(x-ct) \tag{A.5}$$
$$= A\cos(k(x-ct)) \tag{A.6}$$
$$\equiv A\cos(kx-\omega t) \tag{A.7}$$

where we have introduced parameters A to control the amplitude and k, the wavenumber to allow us to specify the wavelength (as we will understand momentarily). We have also defined $ck = \omega$ and ω is called the angular frequency of the wave. Why do k and ω inform us about the wavelength and angular frequency of the harmonic wave? For fixed t, the wave is stationary, and we may measure its wavelength, e.g. when $t = 0$ we read off that the wavelength is

$$\lambda = \frac{2\pi}{k}. \tag{A.8}$$

The period, τ, of the wave (the time taken for one wavelength to travel past a fixed point) is determined from the wavespeed c to be

$$\tau = \frac{\lambda}{c} \tag{A.9}$$
$$= \frac{2\pi}{kc} \tag{A.10}$$
$$= \frac{2\pi}{\omega} \tag{A.11}$$

where we used $\omega = kc$ and from the last expression we discern why ω is called the angular frequency. It is also called angular frequency, to distinguish it from the frequency ν which is the number of wavelengths passing a fixed point in one second, i.e.

$$\nu = \frac{1}{\tau} \tag{A.12}$$
$$= \frac{\omega}{2\pi}. \tag{A.13}$$

We can also see that

$$\nu = \frac{1}{\tau} \tag{A.14}$$
$$= \frac{c}{\lambda}. \tag{A.15}$$

The expressions and formulae we have derived above will be used frequently throughout the course.

What we have described in (A.7) is a wave in one dimension of fixed wavelength k (and so fixed angular frequency $\omega = ck$ as c is a constant determined by the wave equation). The relation $\omega(k) = ck$ is called the dispersion relation and is determined from the wave equation by substituting a travelling wave solution $u(x,t) = f(kx - \omega t)$. For the wave equation (A.1), the dispersion relation $\omega = ck$ means that all waves of any wavelength or amplitude travel at the same speed $c = \frac{\omega}{k}$. Other wave equations have more complicated dispersion relations, for example the Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} + V(x)\Psi \tag{A.16}$$

for $\Psi = f(kx - \omega t)$ becomes

$$-i\hbar\omega f' = -\frac{\hbar^2 k^2}{2m}f'' + Vf \tag{A.17}$$

this has solutions for $f(kx - \omega t) = \exp(i(kx - \omega t))$ (so that $f' = if$, $f'' = -f$) which are complex travelling waves if

$$\hbar\omega = \frac{\hbar^2 k^2}{2m} + V \tag{A.18}$$

and in this case the dispersion relation gives wavespeeds

$$c = \frac{\omega}{k} = \frac{\hbar k}{2m} + \frac{V}{\hbar k} \tag{A.19}$$

which depends on wavelength as $k = \frac{2\pi}{\lambda}$. Solutions to the Schrödinger equation in one dimension are complex waves whose wavespeed depends on the wavelength. Consequently a wave which is constructed of a sum of wave solutions of different wavelengths will not normally keep the same shape, as the constituent waves of different wavelengths travel at different speeds, and the wave disperses over time.

The wave equation and the Schrödinger equation are both linear and homogenous differential equations, so we may add solutions to find new solutions. Different complex waves may be characterised by their wavelength vis k and their amplitude A, and a general wave solution in one dimension is

$$\Psi(x,t) = \int_{k=-\infty}^{\infty} A(k)\exp(i(kx - \omega t)) \tag{A.20}$$

(since k is a continuous variable, we have an integral) where $A(k)$ is the amplitude of the wave of wavenumber k. Note that integrating over k is equivalent to integrating over wavelengths. The meaning of negative values of k is that is encodes a wave travelling to the left, while the positive values of k describe waves travelling to the right.

In three dimensions the general wave solution is

$$\Psi(\boldsymbol{x},t) = \int_{k=-\infty}^{\infty} A(\boldsymbol{k})\exp(i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)) \tag{A.21}$$

where the wavenumber has become a vector of three wavenumbers k_x, k_y and k_z. The direction of propagation of the wave is now determined by the direction of $\boldsymbol{k}$, and wavespeed is given by $c = \frac{\omega}{|\boldsymbol{k}|}$.

A.2 The Fourier Transform

The general complex wave in one dimension in (A.20) is made up of waves with harmonic real and imaginary parts:

$$\exp(i(kx - \omega t)) = \cos(kx - \omega t) + i \sin(kx - \omega t). \tag{A.22}$$

The complex waves at $t = 0$ are e^{ikx} and these functions can be used to form a basis of periodic functions. To understand the meaning of basis, we define an inner product on functions periodic and determined from their values on $[-\pi, \pi]$ by

$$\langle f(x), g(x) \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)^* g(x) dx. \tag{A.23}$$

We now observe that for $k \neq 0$,

$$\frac{1}{2\pi} \int_{x=-\pi}^{\pi} dx\, e^{ikx} = \frac{1}{2\pi} \int_{x=-\pi}^{\pi} dx\, \frac{1}{ik} \frac{d}{dx}(e^{ikx}) \tag{A.24}$$

$$= \frac{1}{2\pi} \left[\frac{1}{ik} e^{ikx} \right]_{x=-\pi}^{\pi} \tag{A.25}$$

$$= 0 \tag{A.26}$$

while if $k = 0$ we have

$$\frac{1}{2\pi} \int_{x=-\pi}^{\pi} dx \; = 1 \tag{A.27}$$

consequently

$$\frac{1}{2\pi} \int_{x=-\pi}^{\pi} dx\, e^{ikx} = \delta_{k,0} \tag{A.28}$$

where

$$\delta_{n,m} = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{otherwise.} \end{cases} \tag{A.29}$$

Hence we may deduce that

$$\langle e^{ikx}, e^{ik'x} \rangle = \delta_{k,k'} \tag{A.30}$$

where $k, k' \in \mathbb{Z}$. This idea is the basis of the Fourier series. Given a series

$$f(x) = \sum_{k=-\infty}^{\infty} A_k e^{ikx} \tag{A.31}$$

we can extract the coefficients A_k as

$$\langle f(x), e^{ik'x} \rangle = \sum_{k=-\infty}^{\infty} A_k \delta_{k,k'} = A_{k'}. \tag{A.32}$$

that is,

$$A_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} dx\, f(x)^* e^{-ikx}. \tag{A.33}$$

How do we develop these methods so that they are viable when a function $f(x)$ is not necessarily periodic and k in any real number rather than just an integer?

First we note that we could change the periodic interval of $[-\pi, \pi]$ to $[-L, L]$ noting that while e^{ix} is periodic with wavelength 2π, $e^{i\pi x/L}$ is periodic with wavelength $2L$ and may be defined from the interval $[-L, L]$ and consequently[1] we have

$$\langle e^{i(\frac{k\pi}{L}x)}, e^{i(\frac{k'\pi}{L}x)} \rangle = \frac{1}{2L} \int_{-L}^{L} dx\, e^{-i(\frac{k\pi}{L}x)} e^{i(\frac{k'\pi}{L}x)} = \delta_{k,k'}. \tag{A.34}$$

Then we have

$$\Psi(x) = \sum_{k=-\infty}^{\infty} A_k e^{i(\frac{k\pi}{L}x)} \tag{A.35}$$

$$A_k = \frac{1}{2L} \int_{-L}^{L} dx\, \Psi(x) e^{-i(\frac{k\pi}{L}x)} \tag{A.36}$$

and $\Psi(x)$ is a periodic function which can be defined by its values in the interval $[-L, L]$ with wavelength $2L$. Our aim is to attempt to extend $L \to \infty$.

We will use the notation $\alpha_k = k\pi/L$ and since $k \in \mathbb{Z}$, we can also define the change in α for subsequent integers k as $\Delta\alpha_k = \pi/L$. This gives us

$$A_k = \frac{\Delta\alpha_k}{2\pi} \int_{-L}^{L} dx\, \Psi(x) e^{-i\alpha_k x} \tag{A.37}$$

and, so,

$$\Psi(x) = \sum_{k=-\infty}^{\infty} \left(\frac{\Delta\alpha_k}{2\pi} \int_{-L}^{L} dx'\, \Psi(x') e^{-i\alpha_k x'} e^{i\alpha_k x} \right) \tag{A.38}$$

$$\equiv \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \Delta\alpha_k \widetilde{\Psi}(\alpha_k) \tag{A.39}$$

[1] One can also see this by making a substitution $x = y\pi/L$ in (A.30).

where

$$\widetilde{\Psi}(\alpha_k) = \int_{-L}^{L} dx'\, \Psi(x') e^{i\alpha_k(x-x')}. \tag{A.40}$$

Taking the limit $L \to \infty$, turns α_k into a continuous variable dependent on continuous k, i.e. $\alpha_k \to \alpha$, $\sum \Delta\alpha_k \to \int d\alpha$ to give

$$\Psi(x) = \frac{1}{2\pi} \int_{\alpha=-\infty}^{\infty} d\alpha\, \widetilde{\Psi}(\alpha) \tag{A.41}$$

$$\widetilde{\Psi}(\alpha) = \int_{-\infty}^{\infty} dx'\, \Psi(x') e^{i\alpha(x-x')} \tag{A.42}$$

i.e.,

$$\Psi(x) = \frac{1}{2\pi} \int_{\alpha=-\infty}^{\infty} d\alpha \int_{x'=-\infty}^{\infty} dx'\, \Psi(x') e^{i\alpha(x-x')} \tag{A.43}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} d\alpha\, e^{i\alpha x} \int_{-\infty}^{\infty} dx'\, \Psi(x') e^{-i\alpha x'} \tag{A.44}$$

$$= \int_{-\infty}^{\infty} d\alpha\, \widetilde{\Psi}(\alpha). \tag{A.45}$$

Notice that the first line above implies that

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} d\alpha\, e^{i\alpha(x-x')} = \delta(x - x'). \tag{A.46}$$

This is a continuous version of (A.30), and the object on the right hand side above is called the Dirac delta function. It is an example of a distribution or generalised function and it satisfies

$$\int_{-\infty}^{\infty} dx\, \delta(x) = 1 \tag{A.47}$$

with

$$\delta(x) = \begin{cases} 0 & x \neq 0 \\ \infty & x = 0. \end{cases} \tag{A.48}$$

As a consequence,

$$\int_{-\infty}^{\infty} dx \delta(x - x') f(x) = f(x') \int_{-\infty}^{\infty} dx\, \delta(x - x') \tag{A.49}$$

$$= f(x'). \tag{A.50}$$

We are interested in general complex wave solutions, which can be expressed in the form of the continuous sum (integral) over k of complex waves weighted by $\widetilde{\Psi}(k)$ at $t = 0$). For this

case we have, and for convention, we have

$$\Psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk\, \widetilde{\Psi}(k) e^{ikx} \tag{A.51}$$

$$\widetilde{\Psi}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx\, \Psi(x) e^{-ikx} \tag{A.52}$$

where $\widetilde{\Psi}(k)$ is the Fourier transform of $\Psi(x)$ and the first expression is called the inverse Fourier transform of $\widetilde{\Psi}(k)$. We can check the conventions are sensible and that the two transforms are indeed inverse to each other explicitly by substituing the Fourier transform into the inverse Fourier transform and using the Dirac delta function:

$$\Psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk\, \widetilde{\Psi}(k) e^{ikx} \tag{A.53}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk\, \left(\frac{1}{\sqrt{2\pi}} \int dx'\, \Psi(x') e^{-ikx'} \right) e^{ikx} \tag{A.54}$$

$$= \int dx'\, \Psi(x') \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{ik(x-x')} \right) \tag{A.55}$$

$$= \int dx'\, \Psi(x') \delta(x - x') \tag{A.56}$$

$$= \Psi(x). \tag{A.57}$$

As the transform is invertible, no information is lost, and in transforming a wave-function from a function of space to a function of wavenumber (or equivalently momentum) no information about the system contained in the wave-function is lost. The Fourier transform gives an alternative setting called momentum space in which to think about quantum physics.

B. Tutorial Exercises

INTRODUCTORY QUANTUM THEORY PROBLEMS I

In this problem set you will think about the ideas in the old quantum theory, that motivated the thought that light could be thought of both as a wave and as a particle. Planck suggested the quantised transfer of energy to avoid the ultraviolet catastrophe in blackbody radiation, while it was Compton's effect which persuaded physicists to accept Einstein's 1905 proposal of the quantum of light, the photon.

(1) (Planck's law) Consider a some physical object at a temperature T. Given a small piece of its surface of area dS, the total energy dE radiated over time dt into wavelengths between λ and $\lambda + d\lambda$, and into a unit solid angle $d\Omega$ (see figure B.1) is characterised by the quantity

$$B(\lambda, T) \equiv \frac{dE}{dS\, dt\, d\lambda\, d\Omega} \tag{B.1}$$

which is known as radiance. The spectral energy density u, discussed in the lectures is related to B by $u = \frac{4\pi}{c} B$.

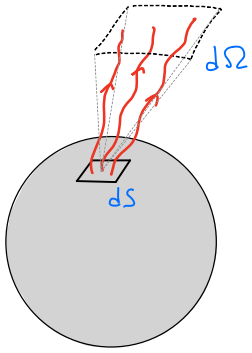


Figure B.1: A cartoon of electromagnetic radiation emitted from a surface element dS into a solid angle $d\Omega$.

(a) What are the units in which the radiance B is measured, expressed in terms of kg, m, s? Recall that it is defined as

$$B = \frac{dE}{dS\, dt\, d\lambda\, d\Omega},$$

where dE is energy, dS is area, dt is time, and $d\lambda$ is length, while $d\Omega$ is dimensionless.

(b) The laws of statistical physics and electromagnetism are governed by equations which contain only two fundamental constants: the Boltzmann constant k_B and the speed of light c. The units of k_B are such that k_BT has units of energy ($kg{\cdot}m^2/s^2$) where T is temperature measured in degrees Kelvin K. Show that the only possible formula for $B(\lambda, T)$ in terms of T, c, k_B, λ that is consistent with the dimensional analysis is

$$B(\lambda, T) = A\frac{ck_BT}{\lambda^4}, \tag{B.2}$$

where A is a dimensionless constant. The classical theory predicts $A = 2$.

Show that the total power radiated over all wavelengths

$$\int_0^\infty d\lambda\, B(\lambda, T)$$

is infinite if (B.2) is true. This clearly does not make sense, and this issue is known as the ultraviolet catastrophe since the divergence comes from very short (ultraviolet) wavelengths.

(c) In the lectures we considered $u(\nu, T)$, a function of ν. Use $\nu = \frac{c}{\lambda}$ to show that

$$u(\lambda, T) = u(\nu, T)\frac{c}{\lambda^2}.$$

[Hint: recall that u is a density and consider $\int u(\lambda)d\lambda$ and $\int u(\nu, T)d\nu$.] Hence use $u = \frac{4\pi}{c}B$ and (B.2) to argue that:

$$u(\nu, T) = \frac{N\nu^2}{c^3}k_BT$$

where N is some number.

(d) Show that Planck's law in terms of the radiance

$$B(\lambda, T) = \frac{2hc^2}{\lambda^5}\frac{1}{e^{\frac{hc}{\lambda k_BT}} - 1},$$

reduces to the Rayleigh-Jeans law

$$B(\lambda, T) = \frac{2ck_BT}{\lambda^4}$$

for λ large such that $\frac{hc}{\lambda} \ll k_BT$.

(2) (Compton scattering) Consider a scattering process where initially we have an electron at rest and an incoming photon with wavelength λ, and in the end we have electron moving with momentum $\boldsymbol{p}$ and a photon with wavelength λ' traveling at angle θ to the direction of the original photon, see figure 1.4.

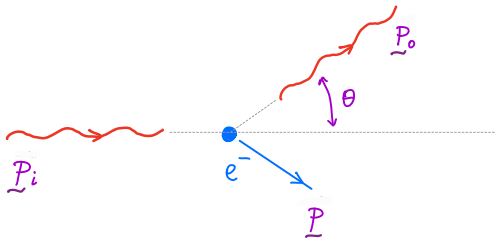


Figure B.2: Kinematics of Compton effect. The incoming radiation is coming from the left, with the wavelength λ, and the radiation scattered at angle θ has the wavelength λ'.

(a) Using conservation of energy and momentum, compute λ' in terms of θ and λ. For this, recall that the energy of the photon is $E = h\nu$ ($\nu = c/\lambda$) and use the fact that the energy E of a relativistic particle with mass m is related to its momentum $\boldsymbol{p}$ by the relation

$$E^2 = m^2c^4 + \boldsymbol{p}^2c^2.$$

You can treat photons as massless relativistic particles ($m = 0$).

(b) Show that it is impossible (i.e. inconsistent with energy and momentum conservation) for the electron to simply absorb the initial positive-energy photon and not emit a final photon.

INTRODUCTORY QUANTUM THEORY PROBLEMS II

In this problem set you will compute explicitly the time evolution of the free particle wave-function which was used in the lectures, and you will consider whether certain wave-functions are normalisable.

(3) (The free particle) Recall the free Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi\ . \tag{B.3}$$

As we discussed in the lecture, this is a time evolution equation: it tells us how $\Psi(\boldsymbol{x},t)$ evolves in t given some initial $\Psi(\boldsymbol{x},t_0)$ at $t=t_0$ (for example, $t_0=0$). Show that if at $t=0$ the wave-function Ψ takes the form

$$\Psi(\boldsymbol{x},0) = \frac{1}{(2\pi)^{3/4}(\Delta x)^{3/2}}\exp\left(-\frac{(\boldsymbol{x}-\boldsymbol{x}_0)^2}{4\Delta x^2} + i\boldsymbol{p}\cdot\boldsymbol{x}/\hbar\right), \tag{B.4}$$

then at other times (B.3) implies that the wave-function is

$$\Psi(\boldsymbol{x},t) = \frac{1}{(2\pi)^{3/4}(\Delta x)^{3/2}\left(1+\frac{i\hbar t}{2m\Delta x^2}\right)^{\frac{3}{2}}}\exp\left(-\frac{(\boldsymbol{x}-\boldsymbol{x}_0-\boldsymbol{p}t/m)^2}{4(\Delta x^2+i\hbar t/2m)} + i\boldsymbol{p}\cdot\boldsymbol{x}/\hbar - iEt/\hbar\right). \tag{B.5}$$

You can assume that it is known that the solution of (B.3) is unique. You can try to solve this problem in your own way, or you can follow the steps below.

(a) Show that (B.5) is also equal to

$$\Psi(\boldsymbol{x},t) = \frac{1}{(2\pi)^{3/4}(\Delta x)^{3/2}\left(1+\frac{i\hbar t}{2m\Delta x^2}\right)^{\frac{3}{2}}}\exp\left(-\frac{\boldsymbol{y}^2}{4(\Delta x^2+i\hbar t/2m)} + i\boldsymbol{p}\cdot\boldsymbol{x}_0/\hbar - \Delta x^2p^2/\hbar^2\right), \tag{B.6}$$

where $\boldsymbol{y} = \boldsymbol{x}-\boldsymbol{x}_0-\frac{2i\Delta x^2\boldsymbol{p}}{\hbar}$. By changing variables from $\boldsymbol{x}$ to $\boldsymbol{y}$ and factoring out constants (terms independent of $\boldsymbol{x}$ and t), show that Ψ solves the Schrödinger equation (B.3) if and only if the function

$$\Phi(\boldsymbol{y},t) = \frac{1}{\left(\Delta x^2+\frac{i\hbar t}{2m}\right)^{\frac{3}{2}}}\exp\left(-\frac{\boldsymbol{y}^2}{4(\Delta x^2+i\hbar t/2m)}\right) \tag{B.7}$$

solves it.

(b) Consider a general function of the form

$$F(\boldsymbol{x}, t) = A(t) \exp\left(-B(t)\boldsymbol{x}^2\right) \tag{B.8}$$

and plug it into the Schrödinger equation

$$i\hbar \frac{\partial F}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 F \, . \tag{B.9}$$

Show that equation (B.9) is satisfied if and only if the functions $A(t)$ and $B(t)$ satisfy

$$\begin{aligned} \frac{\partial B(t)}{\partial t} + \frac{2i\hbar}{m} B(t)^2 = 0, \\ \frac{\partial A(t)}{\partial t} + \frac{3i\hbar}{m} A(t)B(t) = 0. \end{aligned} \tag{B.10}$$

Choose $A(t)$ and $B(t)$ so that $F = \Phi$ with Φ defined in part **a** and verify that they indeed satisfy these equations.

(4) (Normalisable wave-functions)

(a) Are the wave-functions $\Psi_i(x, 0)$ given below normalisable in one dimension? I.e. are the integrals

$$\int_{-\infty}^{+\infty} dx |\Psi_i(x)|^2 \tag{B.11}$$

finite? (Recall that if the above integral is finite, it is possible to rescale Ψ_i to make it normalised, i.e. make this integral 1.) The aim in this question is not to compute the integrals, instead it is to think carefully about whether $|\psi|^2$ is integrable in principle and give supporting, analytical arguments.

(i) $\Psi_1(x) = \frac{1}{a^2+x^2}$,
(ii) $\Psi_2(x) = \frac{1}{a+|x|}$,
(iii) $\Psi_3(x) = \frac{1}{|x|}$, and
(iv) $\Psi_4(x) = \frac{e^{-x^2}}{|x|}$.

Above, a is some real, positive constant.

(b) Are the wave-functions $\Psi_i(x, 0)$ given below normalisable in three dimensions? I.e. are the integrals

$$\int_{\mathbb{R}^3} d^3x |\Psi_i(\boldsymbol{x})|^2 \tag{B.12}$$

finite?

(i) $\Psi_5(x) = \frac{1}{a^2+\boldsymbol{x}^2}$,

(ii) $\Psi_6(x) = \frac{1}{(a^2+\boldsymbol{x}^2)^2}$,

(ii) $\Psi_7(x) = \frac{1}{|\boldsymbol{x}|}$, and

(ii) $\Psi_8(x) = \frac{e^{-\boldsymbol{x}^2}}{|\boldsymbol{x}|}$.

Above, a is some real, positive constant. Note that you do not need to compute these integrals, just determine whether they are finite or not.

[Hint: recall that in spherical polar coordinates (r, θ, ϕ) then $\boldsymbol{x}^2 = r^2$ and

$$\int_{\mathbb{R}^3} d^3x = \int_{r=0}^{\infty} \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} r^2 \sin\theta \, dr \, d\theta \, d\phi \quad .]$$

INTRODUCTORY QUANTUM THEORY PROBLEMS III

In this problem set you will compute position and momentum expectation values for the free particle wave-function in one dimension and work with commutators for operators.

(5) (Expectation values) In this problem we will work with the one-dimensional wave-function[1]:

$$\Psi(x) = \frac{1}{(2\pi)^{1/4}(\Delta x)^{1/2}} \exp\left(-\frac{(x-x_0)^2}{4\Delta x^2} + ipx/\hbar\right). \tag{B.13}$$

Note that this wave-function is normalised.

Note: You will not need to compute non-trivial integrals in this problem.

(a) Show that the expectation value[2] of $\hat{x}$ is x_0,

$$\langle \hat{x} \rangle = x_0. \tag{B.14}$$

First argue that the following integral vanishes,

$$\int_{-\infty}^{+\infty} dx\,(x-x_0)|\Psi(x)|^2 = 0. \tag{B.15}$$

[Hint: Show that $|\Psi(x)|^2$ is even (symmetric) around $x = x_0$. Then combine the normalisation for $\Psi(x)$ with (B.15) and the definition of the expectation value to arrive at the result.]

(b) Show that the expectation value of $\hat{p}$ is p,

$$\langle \hat{p} \rangle = p. \tag{B.16}$$

To do this, first show that

$$(\hat{p}\Psi)(x) = \frac{i\hbar}{2\Delta x^2}(x-x_0)\Psi(x) + p\Psi(x). \tag{B.17}$$

[1] This is the wave-packet for a localised particle in one dimension, rather than the three dimensional version we have looked at in the lectures.

[2] Recall that for an operator $\hat{\mathcal{O}}$, its expectation value in a normalised state $\Psi(x)$ is given by (in one dimension) by $\langle \hat{\mathcal{O}} \rangle = \int_{-\infty}^{+\infty} dx (\Psi(x))^* (\hat{\mathcal{O}}\Psi)(x)$ and that the position operator $\hat{x}$ acts as $\hat{x}\Psi(x) = x\Psi(x)$ while the momentum operator $\hat{p}$ acts as $\hat{p}\Psi(x) = -i\hbar\frac{\partial\Psi(x)}{\partial x}$

(6) (Commutators)

(a) Prove that the following operator equations hold:

$$[\hat{p}^2, \hat{x}] = -2i\hbar\hat{p}, \tag{B.18}$$

$$[V(\hat{x}), \hat{x}] = 0. \tag{B.19}$$

where where $V(\hat{x})$ denotes the potential operator appearing in the Hamiltonian operator $\hat{H} \equiv \frac{\hat{p}^2}{2m} + V(\hat{x})$, $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$ is the commutator of $\hat{A}$ and $\hat{B}$. [Hint: Allow the operators to act on a generic wave-function $\Psi(x)$.]

(b) Use the definition of a commutator stated in part (a) to check it satisfies the following properties:

$$[\alpha\hat{A} + \beta\hat{B}, \hat{C}] = \alpha[\hat{A}, \hat{C}] + \beta[\hat{B}, \hat{C}], \tag{B.20}$$

$$[\hat{A}, \beta\hat{B} + \gamma\hat{C}] = \beta[\hat{A}, \hat{B}] + \gamma[\hat{A}, \hat{C}], \tag{B.21}$$

$$[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}], \tag{B.22}$$

where $\alpha, \beta, \gamma \in \mathbb{C}$ and $\hat{A}, \hat{B}, \hat{C}$ are any linear operators. Use some of these properties and the results of part (a) to show that

$$[\hat{H}, \hat{x}] = -\frac{i\hbar}{m}\hat{p}. \tag{B.23}$$

This is the equation that we used in the lectures.

INTRODUCTORY QUANTUM THEORY PROBLEMS IV

In this problem set we will think some more about solutions to the Schrödinger equation for a particle in an infinite potential well, we will look at the correspondence principle to see the connection with the classical problem, uncertainty and the standard deviation, and consider what happens as V_0 grows.

---

(7) (The correspondence principle for the particle in the box in one dimension). Consider a particle of mass m moving in one dimension, with coordinate x and confined to the region where $0 < x < L$. The particle moves freely in the region $0 < x < L$ and bounces elastically off the walls at $x = 0$ and $x = L$ when it hits them. We impose that its wave-function $\Psi(x,t)$ vanishes for $x < 0$ and $x > L$.

(a) State the normalisation condition for $\Psi(x,t)$, given that it is confined to the box.

(b) What is the probability of finding the particle in the interval $[a,b]$ within the box (i.e. $0 < a < b < L$)?

The n'th energy eigenstate for the particle in the box is given by the wave-function

$$\Psi_n(x,t) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right) e^{-iE_n t/\hbar} \tag{B.24}$$

where $E_n = \hbar^2 n^2 \pi^2/(2mL^2)$.

(c) Show that the wave-function $\Psi_n(x,t)$ is normalised.

(d) Suppose that the particle is in the n'th energy eigenstate, $\Psi_n(x,t)$, what is the probability of finding the particle in the interval $[a,b]$ within the box?

(e) Show that the probability found in part (d) tends to the classical result (namely $P^{Classical}_{[a,b]} = (b-a)/L$ as, classically, the particle is equally likely to be found in any part of the box at any time) as $n \to \infty$.

(8) (Uncertainty and the particle in an infinite potential well, or box)

(a) Define the uncertainty $\Delta\mathcal{O}^2$ of an operator $\hat{\mathcal{O}}$ as the expected squared deviation of $\hat{\mathcal{O}}$ from the expected value $\langle\hat{\mathcal{O}}\rangle$, i.e.

$$\Delta\mathcal{O}^2 \equiv \langle(\hat{\mathcal{O}} - \langle\hat{\mathcal{O}}\rangle)^2\rangle. \tag{B.25}$$

Show that in any state $\Psi(x)$ the following identity holds,

$$\Delta\mathcal{O}^2 \equiv \langle\hat{\mathcal{O}}^2\rangle - \langle\hat{\mathcal{O}}\rangle^2. \tag{B.26}$$

Hint: show that for any two linear operators $\hat{A}$ and $\hat{B}$ and $\alpha, \beta \in \mathbb{C}$ the following holds

$$\langle\alpha\hat{A} + \beta\hat{B}\rangle = \alpha\langle\hat{A}\rangle + \beta\langle\hat{B}\rangle. \tag{B.27}$$

Then use this and the fact that $\langle\hat{\mathcal{O}}\rangle$ is just a number.

(b) Recall that in the class we found the wave-functions for the energy eigenstates of a particle in a box to be given by

$$\Psi_n(x) = \sqrt{\frac{2}{L}} \sin \tfrac{n\pi x}{L}. \tag{B.28}$$

Use (B.26) and the definition of the expectation value to compute the uncertainty Δp^2 of the momentum operator $\hat{p}$ in the state Ψ_n. You should find that the uncertainty is non-zero, i.e. the state Ψ_n does not have a definite value of $\hat{p}$. How else can you check that Ψ_n does not have a definite value of $\hat{p}$?

Hint: to compute the integrals, use the identities

$$\sin\alpha\cos\alpha = \frac{\sin 2\alpha}{2} \tag{B.29}$$

$$\sin^2\alpha = \frac{1-\cos 2\alpha}{2}. \tag{B.30}$$

(9) (The Particle in a three-dimensional box[3])

Consider a wave-function $\Psi(\boldsymbol{x}) = \Psi(x_1, x_2, x_3)$, with the time-independent Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2\Psi(\boldsymbol{x}) = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}\right)\Psi(x_1, x_2, x_3) = E\Psi(x_1, x_2, x_3) \tag{B.31}$$

and the boundary conditions

$$\Psi(0, x_2, x_3) = \Psi(L, x_2, x_3) = 0, \tag{B.32}$$

$$\Psi(x_1, 0, x_2) = \Psi(x_1, L, x_2) = 0, \tag{B.33}$$

$$\Psi(x_1, x_2, 0) = \Psi(x_1, x_2, L) = 0. \tag{B.34}$$

This represents a particle that is confined to a $L \times L \times L$ box.

[3]This is the three-dimensional version of the particle in the infinite potential well studied in section 3.2 of the lecture notes.

(a) We will use separation of variables,

$$\Psi(x_1, x_2, x_3) = \Psi_1(x_1)\Psi_2(x_2)\Psi_3(x_3), \tag{B.35}$$

where Ψ_1, Ψ_2, Ψ_3 are new unknown functions.[4] Show that the time-independent Schrödinger equation implies

$$\frac{\Psi_1''(x_1)}{\Psi_1(x_1)} + \frac{\Psi_2''(x_2)}{\Psi_2(x_2)} + \frac{\Psi_3''(x_3)}{\Psi_3(x_3)} = -\frac{2mE}{\hbar^2}. \tag{B.36}$$

Argue that this is only possible if for $i = 1, 2, 3$

$$\frac{\Psi_i''(x_i)}{\Psi_i(x_i)} \tag{B.37}$$

is a constant. Denote the constant by $-k_i^2$,

$$\frac{\Psi_i''(x_i)}{\Psi_i(x_i)} = -k_i^2. \tag{B.38}$$

(b) The equation we have developed is solved by

$$\Psi_i(x_i) = A_i \sin k_i x + B_i \cos k_i x \tag{B.39}$$

for $i = 1, 2, 3$. Show that the boundary conditions (B.32-B.34) imply $B_i = 0$

$$k_i = \frac{\pi n_i}{L}, \tag{B.40}$$

where n_1, n_2, n_3 are (possibly distinct) positive integers.

(c) Using (B.36), (B.38), and (B.40), show that the allowed values of the energy are

$$E = \frac{\pi^2\hbar^2}{2mL^2}\left(n_1^2 + n_2^2 + n_3^2\right). \tag{B.41}$$

[4]The subscripts denote the three different coordinate directions. Do not confuse this notation with $\Psi_n(x)$ that we used in one-dimension to denote the energy eigenfunctions.

INTRODUCTORY QUANTUM THEORY
PROBLEMS V

In this problem set we will think some more about stationary states and solutions to the Schrödinger equation, beginning with the particle in the finite potential well, we will introduce degrenerate states and consider the particle on a circle.

(10) (Particle in a potential well)

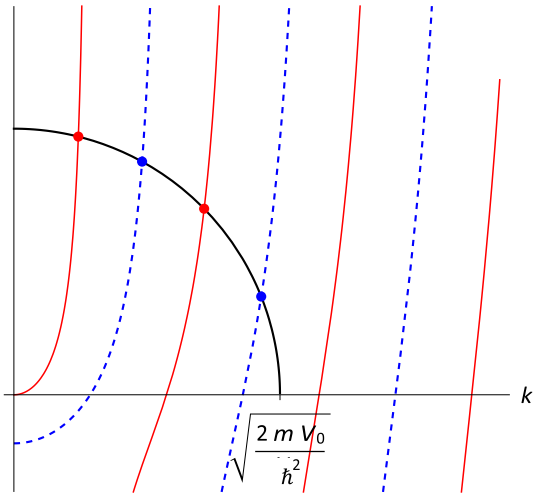


Figure B.3: A graphical solution of (3.119) and (3.124). The left-hand side of these equations is plotted in solid black (the quarter-circle). The right-hand side of (3.119) is plotted in solid red, and the right-hand side of (3.124) in dashed blue. The red and blue circles mark the even and odd solutions, respectively. The number of solutions depends on the parameters of the problem, and in this case is four.

The energy levels with even eigenfunctions[5] are found from the solutions $k \in (0, \sqrt{\frac{2mV_0}{\hbar}})$ to the equation[6]

$$\sqrt{\frac{2mV_0}{\hbar^2} - k^2} = k \tan\left(\frac{kL}{2}\right). \tag{B.42}$$

[5]See the section in the lecture notes on the particle in a finite potential well, in particular the equation here is (3.119).

[6]Given a solution k, the energy is $E = \frac{\hbar^2 k^2}{2m} - V_0$.

The energy levels with odd eigenfunctions[7] are instead determined by the solutions $k \in (0, \sqrt{\frac{2mV_0}{\hbar}})$ to

$$\sqrt{\frac{2mV_0}{\hbar^2} - k^2} = k \tan\left(\frac{kL + \pi}{2}\right). \tag{B.43}$$

As discussed in the lectures, the solutions can be visualised using figure B.3, and that as V_0 is increased, new solutions (and thus energy levels) appear.

(a) Suppose we start at $V_0 = 0$ and start increasing it. Find the values of V_0 at which new energy levels with even eigenfunctions appear.

(b) Suppose we start at $V_0 = 0$ and start increasing it. Find the values of V_0 at which new energy levels with odd eigenfunctions appear.

(11) (Degenerate states) If two or more linearly independent solutions to the time-independent Schrödinger equation have the same energy E, these states are called *degenerate* states.

(a) Explain why the stationary states $\Phi(x,t) = \phi(t)\varphi(x)$ for the free particle ($V = 0$) are degenerate states for each $E > 0$.

(b) For a general potential in one dimension that admits bound state solutions, prove that there are no degenerate bound states. [Method: Prove this by contradiction, i.e. assume there exist two linearly independent bound states φ_1 and φ_2 with the same energy E and then use the Schrödinger equation to show that $\varphi_2 \frac{d\varphi_1}{dx} - \varphi_1 \frac{d\varphi_2}{dx}$ is constant in space and from data at $x = \pm\infty$ that it is equal to zero. Finally show that your results imply that φ_2 must be a multiple of φ_1 and so not linearly independent.]

(12) (The free particle on a circle) Consider a particle of mass m restricted to move on a circle with circumference $L > 0$. Let x denote the coordinate on the circle with $x \in \mathbb{R}$ subject to the identification $x + L \sim x$. Let the circle be given by the coordinates $x \in [0, L]$ with $x = 0$ and $x = L$ identified, and hence $\varphi(x + L) = \varphi(x)$. Find the stationary states and explain why they are degenerate. Due to the constraint that the particle moves on a circle, these are bound states (they do not reach infinity), and yet they are degenerate - why does the theorem in the previous question fail to be valid in this case?

[7] As derived in equation (3.124) in the lecture notes.

INTRODUCTORY QUANTUM THEORY PROBLEMS VI

In this problem set we will think some more about stationary states and solutions to the Schrödinger equation, in two cases: the step-function potential, and a reflectionless potential.

(13) (The step-function potential) The step-function potential is given by

$$V(x) = \begin{cases} 0 & x < 0 \\ V_0 & x \geq 0 \end{cases} \tag{B.44}$$

where V_0 is a positive, real constant. Solve the time-independent Schrödinger equation for the step-function potential for a wave incident from the left (i.e. travelling from $x = -\infty$ towards $x = 0$) for the cases:

(i) $V_0 < E$ and

(ii) $0 \leq E < V_0$

In both cases compute the transmission and reflection probabilities. Use

$$\text{(Reflection probability)} \qquad R \equiv \frac{|J_R|}{|J_I|} \tag{B.45}$$

$$\text{(Transmission probability)} \qquad T \equiv \frac{|J_T|}{|J_I|}. \tag{B.46}$$

where J_I, J_R and J_T are the incident, reflected and transmitted probability currents, for the appropriate wave-function and the probability current in one dimension is

$$J = \frac{i\hbar}{2m}\left(\frac{\partial \Psi^*}{\partial x}\Psi - \Psi^*\frac{\partial \Psi}{\partial x}\right) \tag{B.47}$$

(14) (A reflectionless potential) Consider the potential

$$V(x) = -\frac{\hbar^2 a^2}{m}\text{sech}^2(ax) \tag{B.48}$$

where a is a positive, real constant.

(a) Sketch the potential.

(b) Check that the potential has a bound state given by

$$\varphi_0(x) = A\,\mathrm{sech}(ax) \tag{B.49}$$

where A is a real constant. Find its energy, normalise $\varphi_0(x)$ and sketch its graph. Comment: φ_0 is the ground state.

(c) Show that the function

$$\varphi_k(x) = A\left(\frac{ik - a\tanh(ax)}{ik + a}\right)e^{ikx} \tag{B.50}$$

where $k = \sqrt{2mE}/\hbar$, solves the time-independent Schrödinger equation for any positive energy E. Since $\tanh(\alpha) \to -1$ as $\alpha \to -\infty$ then we have the approximation:

$$\varphi_k(x) \approx e^{ikx} \quad \text{as } x \to -\infty \tag{B.51}$$

This represents an incident wave from the left with no reflected wave. What is the form of $\varphi_k(x)$ as $x \to \infty$? What are the reflection and transmission probabilities for this potential?

INTRODUCTORY QUANTUM THEORY PROBLEMS VII

In this problem set we will think about the space $L^2(\mathbb{R})$ as a Hilbert space

---

(15) (The triangle inequality for $L^2(\mathbb{R})$) Recall that the Hilbert space $L^2(\mathbb{R})$ is the set of equivalence classes[8] of square-integrable functions $f : \mathbb{R} \to \mathbb{C}$. The inner product on this space is defined as

$$\langle f, g\rangle = \int_{-\infty}^{+\infty} dx\, f^*(x)g(x) \tag{B.52}$$

where $*$ means complex conjugation. The norm on the space is defined from the inner product as $\|f\| \equiv \sqrt{(f, f)}$.

(a) Use the arithmetic-mean geometric-mean inequality (i.e. that $a^2+b^2 \geq 2ab$) to prove that $\langle f, g\rangle$ is finite for $f, g \in L^2(\mathbb{R})$.

(b) Prove the Cauchy-Schwarz inequality in this setting: $\|f\|\|g\| \geq |\langle f, g\rangle|$.

(c) Prove that the triangle inequality: $\|f + g\| \leq \|f\| + \|g\|$ is satisfied.

(16) (The Hilbert space $L^2(\mathbb{R})$)

(a) Let

$$f(x) = \begin{cases} 0, & x < 0, \\ 1, & 0 \leq x < 1, \\ i, & 1 \leq x < 2, \\ 0, & 2 \leq x. \end{cases} \tag{B.53}$$

$$g(x) = \begin{cases} 0, & x < \frac{1}{2}, \\ 1, & \frac{1}{2} \leq x < \frac{3}{2}, \\ 0, & x \geq \frac{3}{2}. \end{cases} \tag{B.54}$$

Compute $\|f\|, \|g\|, \langle f, g\rangle, \langle g, f\rangle$. Check the Cauchy-Schwarz inequality $|\langle f, g\rangle| \leq \|f\|\|g\|$.

[8] Two functions f_1, f_2 are considered equivalent if $f_1(x) = f_2(x)$ for almost all x. Almost all means that the set N of x for which $f_1(x) \neq f_2(x)$ has measure 0. Measure of N is, intuitively, the total length of all intervals in N. For example, any countable N has measure 0.

(b) Prove that $L^2(\mathbb{R})$ is infinite-dimensional. For this, show how to construct arbitrarily large linearly-independent subsets $\{f_1, \cdots, f_n\} \subset L^2(\mathbb{R})$. *Hint:* Try using functions $f_k(x) = g(x-k)$, where $g(x)$ is as in (B.54).

(17) (Finite-dimensional Hilbert spaces)

Let $\mathcal{H}$ be a Hilbert space of dimension n, where $n \in \mathbb{N}$ and $n > 0$. Suppose that

$$|e_1\rangle, |e_2\rangle, \cdots |e_n\rangle \tag{B.55}$$

is an orthonormal set $\langle e_i|e_j\rangle = \delta_{ij}$.

(a) Show that $\{|e_1\rangle, |e_2\rangle, \cdots |e_n\rangle\}$ is a linearly-independent set. [*Hint:* Recall the definition of a linearly-independent set of vectors and take some inner products.]

(b) Any linearly-independent subset of $\dim \mathcal{H}$ elements is a basis (hence by part (a) $|e_i\rangle$ form an orthonormal basis of $\mathcal{H}$). Using this, prove that for $\mathcal{H} = \mathbb{C}^2$

$$|e_1\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{pmatrix}, \quad |e_2\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} \end{pmatrix} \tag{B.56}$$

is a basis.

(c) How would you represent $\langle e_1|$ and $\langle e_2|$ by row vectors in $(\mathbb{C}^2)^*$?

(d) Recall that we used $|a\rangle\langle b|$ to represent the operator that sends $|v\rangle$ to $\langle b|v\rangle|a\rangle$. Focus on $\mathcal{H} = \mathbb{C}^n$, and suppose that

$$|a\rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \quad \langle b| = (b_1, b_2, \cdots, b_n). \tag{B.57}$$

Explain why the operator $|a\rangle\langle b|$ as defined above can be represented by the matrix product

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} (b_1, b_2, \cdots, b_n) = \begin{pmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_nb_1 & a_nb_2 & \cdots & a_nb_n \end{pmatrix}. \tag{B.58}$$

[*Hint:* Check that it acts as required on column vectors. Use associativity of matrix multiplication.]

(e) In the $\mathbb{C}^2$ example of equation (B.56) check that, as discussed during the lecture,

$$|e_1\rangle\langle e_1| + |e_2\rangle\langle e_2| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{B.59}$$

(18) (Change of basis in a Hilbert space)

Suppose we have two distinct orthonormal bases in a Hilbert space $\mathcal{H}$,

$$\{|a_1\rangle, |a_2\rangle, \cdots\}, \qquad \text{and} \qquad \{|b_1\rangle, |b_2\rangle, \cdots\}. \tag{B.60}$$

In particular, $\langle a_i|a_j\rangle = \delta_{ij}$ and $\langle b_i|b_j\rangle = \delta_{ij}$. Let $|v\rangle$ be in $\mathcal{H}$. By assumption, it can be expanded in either basis,

$$|v\rangle = \sum_{k=1}^{\infty} \alpha_k |a_k\rangle, \qquad |v\rangle = \sum_{k=1}^{\infty} \beta_k |b_k\rangle. \tag{B.61}$$

Express the coefficients β_k in terms of α_k.

INTRODUCTORY QUANTUM THEORY PROBLEMS VIII

In this problem set we will practise working with the Dirac delta-function, the Fourier transform, adjoint operators, and do some computations on a finite dimensional Hilbert space.

(19) (The Dirac delta-function and the Fourier transform)

Recall the following integral

$$\int_{-\infty}^{+\infty} dx\, e^{ix(k-k')} = 2\pi\delta(k-k'), \tag{B.62}$$

as well as the defining property of the delta-function

$$\int_{-\infty}^{+\infty} dx\, \delta(x-y)f(x) = f(y). \tag{B.63}$$

(a) Show that the integral

$$\int_{-\infty}^{+\infty} dx\, x^n e^{ikx}, \tag{B.64}$$

where n is a positive integer is given by

$$2\pi i^{-n}\delta^{(n)}(k), \tag{B.65}$$

where

$$\delta^{(n)}(k) = \frac{\partial^n}{\partial k^n}\delta(k) \tag{B.66}$$

is the n-th derivative of $\delta(k)$. [*Hint:* Differentiate (B.62) n times with respect to k.]

(b) Show that

$$\int_{-\infty}^{+\infty} dx\, \delta^{(n)}(x-y)f(x) = (-1)^n \frac{\partial^n f(y)}{\partial y^n}. \tag{B.67}$$

[*Hint:* Differentiate (B.63) n times with respect to y.]

(c) Consider a function $F(x)$ and recall the Fourier transform formulas used during the lecture (reviewed and derived in appendix A of the lecture notes),

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dk\, e^{-ikx} \widetilde{F}(k) \tag{B.68}$$

$$\widetilde{F}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dx\, e^{ikx} F(x). \tag{B.69}$$

Use the results of parts (a) and (b) to verify them for

$$F(x) = x^n, \tag{B.70}$$

$$\widetilde{F}(k) = \sqrt{2\pi} i^{-n} \delta^{(n)}(k). \tag{B.71}$$

(20) (More on the Dirac delta-function)

Recall that the defining property of the Dirac delta function is

$$\int_{-\infty}^{+\infty} dx\, \delta(x-y) f(x) x = f(y). \tag{B.72}$$

Use this definition and the change of variables (i.e. pretending that this is an ordinary integral) to prove the following properties of the Dirac delta-function,

$$\delta(x) = \delta(-x), \tag{B.73}$$

$$\delta(ax) = \frac{1}{|a|} \delta(x). \tag{B.74}$$

where $a \in \mathbb{R}, a \neq 0$.

(21) (More on the adjoint operators)

An operator $\hat{\mathcal{P}}$ is adjoint to $\hat{\mathcal{O}}$, i.e. $\hat{\mathcal{P}} = \hat{\mathcal{O}}^\dagger$ if and only if

$$\langle\, \hat{\mathcal{P}}|u\rangle\, ,\, |v\rangle\, \rangle = \langle\, |u\rangle\, ,\, \hat{\mathcal{O}}|v\rangle\, \rangle \tag{B.75}$$

for all vectors $|u\rangle, |v\rangle$ in the Hilbert space. Using this definition and the properties of the Hermitian inner product $\langle\, ,\, \rangle$, prove the following identities

$$(\alpha\hat{\mathcal{O}}_1 + \beta\hat{\mathcal{O}}_2)^\dagger = \alpha^* \hat{\mathcal{O}}_1^\dagger + \beta^* \hat{\mathcal{O}}_2^\dagger, \tag{B.76}$$

$$(\hat{\mathcal{O}}^\dagger)^\dagger = \hat{\mathcal{O}}, \tag{B.77}$$

$$(\hat{\mathcal{O}}_1 \hat{\mathcal{O}}_2)^\dagger = \hat{\mathcal{O}}_2^\dagger \hat{\mathcal{O}}_1^\dagger. \tag{B.78}$$

(22) (Measurement in a finite-dimensional Hilbert space)

Let $\mathcal{H}$ be a two-dimensional Hilbert space, and let $|0\rangle, |1\rangle \in \mathcal{H}$ form an orthonormal basis,[9]

$$\langle m|n\rangle = \delta_{m,n}. \tag{B.79}$$

Consider an operator $\hat{\mathcal{O}}$, which acts on this basis as follows,

$$\hat{\mathcal{O}}|0\rangle = |0\rangle + i|1\rangle, \tag{B.80}$$
$$\hat{\mathcal{O}}|1\rangle = -i|0\rangle + |1\rangle. \tag{B.81}$$

(a) Write the two-by-two matrix $\boldsymbol{O}$ that represents the action of $\hat{\mathcal{O}}$ in the basis $|0\rangle, |1\rangle$. This matrix is defined as follows: let $|v\rangle$ be any vector in $\mathcal{H}$. Write

$$|v\rangle = v_0|0\rangle + v_1|1\rangle, \tag{B.82}$$
$$\hat{\mathcal{O}}|v\rangle = v_0'|0\rangle + v_1'|1\rangle, \tag{B.83}$$

then $\boldsymbol{O}$ is such a matrix that

$$\begin{pmatrix} v_0' \\ v_1' \end{pmatrix} = \boldsymbol{O} \begin{pmatrix} v_0 \\ v_1 \end{pmatrix}. \tag{B.84}$$

Is $\boldsymbol{O}$ Hermitian?

(b) Check explicitly that

$$\boldsymbol{O}_{mn} = \langle m|\hat{\mathcal{O}}|n\rangle, \tag{B.85}$$

where as usual m labels the rows and n labels the columns (starting from 0). Prove that this had to be the case by starting from the definition of $\boldsymbol{O}$ in part (a) and using identities such as $v_k = \langle k|v\rangle$.

(c) Find the eigenvalues and the eigenvectors of the matrix $\boldsymbol{O}$. Denote the normalised eigenvectors of $\hat{\mathcal{O}}$ by $|\lambda_1\rangle$ and $|\lambda_2\rangle$, where λ_1, λ_2 are the eigenvalues of $\hat{\mathcal{O}}$.[10] Use the eigenvectors of $\boldsymbol{O}$ to express the normalised eigenvectors of $\hat{\mathcal{O}}$ in terms of the vectors $|0\rangle$ and $|1\rangle$. Check that $\langle\lambda_1|\lambda_2\rangle = 0$.

(d) Let $|\Psi\rangle = |0\rangle + |1\rangle$. First, normalise $|\Psi'\rangle = C|\Psi\rangle$ with C chosen so that

$$\langle\Psi'|\Psi'\rangle = 1. \tag{B.86}$$

[9] Note that $|0\rangle$ is just a name for a vector, it is not the zero-vector of $\mathcal{H}$, $|0\rangle \neq 0$.

[10] Recall that normalised means $\langle\lambda_i|\lambda_i\rangle = 1$.

Express $|\Psi'\rangle$ as a linear combination of $|\lambda_1\rangle$ and $|\lambda_2\rangle$,

$$|\Psi'\rangle = c_1|\lambda_1\rangle + c_2|\lambda_2\rangle. \tag{B.87}$$

What are the possible outcomes of measuring $\hat{\mathcal{O}}$ in the state $|\Psi'\rangle$? What are their probabilities? Why was it important to normalise $|\Psi'\rangle$?

(23) (Consecutive Measurements) An operator $\hat{A}$ representing an observable has two normalised eigenstates ψ_1 and ψ_2, with eigenvalues a_1 and a_2 respectively. Operator $\hat{B}$, representing another observable, has two normalised eigenstates ϕ_1 and ϕ_2 with respective eigenvalues b_1 and b_2. The eigenstates are related by

$$\psi_1 = \frac{1}{5}(3\phi_1 + 4\phi_2)$$
$$\psi_2 = \frac{1}{5}(4\phi_1 - 3\phi_2).$$

(a) A measurement of $\hat{A}$ is made and the result is a_1. What is the state of the system immediately after this measurement?

(b) If $\hat{B}$ is now measured, what are the possible results and what are the probabilities of obtaining each result?

(c) Immediately $\hat{A}$ is measured again. What is the probability of getting a_1 again? If we knew the outcome of the measurement of $\hat{B}$ in part (b) how would your answer to the first part of part (c) change?

INTRODUCTORY QUANTUM THEORY PROBLEMS IX

In this problem set we will look at when the uncertainty relation is saturated for two operators, and think about what this means for $\hat{x}$ and $\hat{p}$. In the second question we will introduce an observable called spin, and show that the Hilbert space for spin is $\mathbb{C}^2$, before do some computations on a finite dimensional Hilbert space.

---

(24) (Saturation in the Heisenberg uncertainty principle)

Let $\hat{A}$ and $\hat{B}$ be Hermitian operators on a Hilbert space $\mathcal{H}$. In this problem you will identify the conditions on the state $|\Psi\rangle$ under which the Heisenberg uncertainty principle in this state,

$$\Delta\hat{A}\Delta\hat{B} \geq \frac{1}{2}\left|\langle[\hat{A},\hat{B}]\rangle\right|, \tag{B.88}$$

turns into an equality,

$$\Delta\hat{A}\Delta\hat{B} = \frac{1}{2}\left|\langle[\hat{A},\hat{B}]\rangle\right|. \tag{B.89}$$

That is, you will identify the states which have the "smallest possible" amount of uncertainty.

(a) Let $\hat{A}_0 = \hat{A} - \langle\hat{A}\rangle$ and $\hat{B}_0 = \hat{B} - \langle\hat{B}\rangle$. By following the proof of (B.88) from lectures (or from the notes), show that we get (B.89) in a state $|\Psi\rangle$ only if one (or both) of the following two statements is true:

(i) The equation

$$\hat{A}_0|\Psi\rangle = -i\alpha\hat{B}_0|\Psi\rangle \tag{B.90}$$

holds for some $\alpha \in \mathbb{R}$.

(ii) $|\Psi\rangle$ is an eigenstate of $\hat{A}$ or $\hat{B}$ (possibly both).

Assuming (i) holds, express α in terms of $\Delta\hat{A}$ and $\Delta\hat{B}$.

You can use the fact that the Cauchy-Schwartz inequality is saturated for non-zero $|u\rangle$ and $|v\rangle$, i.e.

$$\left|\,\langle u|v\rangle\,\right| = \left\|\,|u\rangle\,\right\|\left\|\,|v\rangle\,\right\|, \qquad |u\rangle \neq 0, |v\rangle \neq 0 \tag{B.91}$$

if and only if $|u\rangle = c|v\rangle$ for some $c \in \mathbb{C}$.

(b) Set $\hat{A} = \hat{x}$ and $\hat{B} = \hat{p}$. Assuming that $\langle \hat{x} \rangle = \langle \hat{p} \rangle = 0$ in the state $|\Psi\rangle$, express the condition (B.90) in terms of the wave function $\Psi(x)$.

Solve this equation to find the wave function $\Psi(x)$.

Why does this imply that $\alpha > 0$?

(c) Repeat part (b) for the case $\Delta\hat{x} = x_0$ and $\Delta\hat{p} = p_0$ in the same state $|\Psi\rangle$. Derive the expression for the wave function $\Psi(x)$ that corresponds to the state $|\Psi\rangle$.

(25) (Electron spin and Larmor precession) Consider the following operators on $\mathcal{H} = \mathbb{C}^2$,

$$\hat{S}_1 = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{S}_2 = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{S}_3 = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{B.92}$$

These operators are used to describe the three components of the electron's spin. N.B. these are the Pauli matrices of SU(2) up to a factor of $\frac{\hbar}{2}$ i.e. $\hat{S}_i = \frac{\hbar}{2}\sigma_i$.

(a) Are the $\hat{S}_i$ Hermitian? Show that the spectrum of each of the $\hat{S}_i$ is $\{\frac{\hbar}{2}, -\frac{\hbar}{2}\}$.

(b) Show that the vector

$$\langle \hat{\boldsymbol{S}} \rangle = \begin{pmatrix} \langle \hat{S}_1 \rangle \\ \langle \hat{S}_2 \rangle \\ \langle \hat{S}_3 \rangle \end{pmatrix}$$

built out of the expectation values $\langle \hat{S}_i \rangle$ in any normalised state $|\Psi\rangle$ always has length $\frac{\hbar}{2}$.

[Hint: Choose $|\Psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in \mathbb{C}^2$ to be an arbitrary normalised state. Write $\langle \hat{S}_i \rangle$ in terms of the components of $|\Psi\rangle$ and their complex conjugates. Expand out the length-squared of vector $\langle \hat{\boldsymbol{S}} \rangle$ and simplify the answer]

(c) Suppose that the Hamiltonian is given by

$$\hat{H} = -\gamma B\, \hat{S}_3, \tag{B.93}$$

where $\gamma, B \in \mathbb{R}$ are some constants. This describes the interaction of the electron's spin with a constant magnetic field $\boldsymbol{B}$ (and $B = |\boldsymbol{B}|$) along the third coordinate axis, γ is called the gyromagnetic ratio (it is the ratio of a particle's magnetic moment to its angular momentum). Compute the evolution operator

$$\hat{\mathcal{U}}(t) = e^{-i\hat{H}t/\hbar}. \tag{B.94}$$

Verify explicitly the properties $\hat{\mathcal{U}}(t)^\dagger = \hat{\mathcal{U}}(-t)$ and $\hat{\mathcal{U}}(t_1)\hat{\mathcal{U}}(t_2) = \hat{\mathcal{U}}(t_1 + t_2)$.

(d) In this part, consider a general Hilbert space $\mathcal{H}$. Let $\hat{\mathcal{O}}$ be a linear operator, and

$$\langle\hat{\mathcal{O}}\rangle(t) \equiv \langle\Psi(t)|\hat{\mathcal{O}}|\Psi(t)\rangle \tag{B.95}$$

be the expectation value of $\hat{\mathcal{O}}$ in a time-dependent state $|\Psi(t)\rangle = \hat{\mathcal{U}}(t)|\Psi(0)\rangle$. Show that

$$\langle\hat{\mathcal{O}}\rangle(t) = \langle\hat{\mathcal{U}}(t)^\dagger\,\hat{\mathcal{O}}\,\hat{\mathcal{U}}(t)\rangle, \tag{B.96}$$

where the right-hand side is the expectation value in the state $|\Psi(0)\rangle$.

(e) Returning to the case $\mathcal{H} = \mathbb{C}^2$ and $\hat{H}$ given by (B.93), via an explicit calculation, express $\hat{\mathcal{U}}(t)^\dagger\,\hat{S}_1\,\hat{\mathcal{U}}(t)$ in terms of the $\hat{S}_i$. Do the same for $\hat{\mathcal{U}}(t)^\dagger\,\hat{S}_2\,\hat{\mathcal{U}}(t)$ and $\hat{\mathcal{U}}(t)^\dagger\,\hat{S}_3\,\hat{\mathcal{U}}(t)$.

(f) Interpreting $\langle\hat{S}_i\rangle(t)$ as the three components of a vector $\langle\hat{\boldsymbol{S}}\rangle$, use the above results to show that this vector rotates ("precesses") around the 3rd axis at a constant rate. Find the frequency of this rotation.

(26) (Uncertainty and the spin operators)

Consider again the spin operators on $\mathcal{H} = \mathbb{C}^2$, given by,

$$\hat{S}_1 = \frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \quad \hat{S}_2 = \frac{\hbar}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}, \quad \hat{S}_3 = \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}. \tag{B.97}$$

An electron is in the state

$$\chi = \frac{1}{5}\begin{pmatrix}3i\\ 4\end{pmatrix}.$$

(a) What are the expectation values for the operators $\hat{S}_1$, $\hat{S}_2$ and $\hat{S}_3$?

(b) What are the uncertainties ΔS_1, ΔS_2 and ΔS_3?
[*Hint*: Recall that $\langle(\hat{A} - \langle\hat{A}\rangle)^2\rangle = \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2$.]

(c) Show that the generalised uncertainty relations between any two pairs of the spin operators is satisfied.
[*Hint*: The generalised uncertainty relation for $\hat{A}$ and $\hat{B}$ is $\Delta A\Delta B \geq \frac{1}{2}|\langle[\hat{A},\hat{B}]\rangle|$.]

INTRODUCTORY QUANTUM THEORY PROBLEMS X

In this problem set we will work with the harmonic oscillator; we will revisit both the algebraic approach and the analytical approach, before looking at the harmonic oscillator in the Heisenberg picture. We recall here the definition of the annihilation operator

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{x} + \frac{i\hat{p}}{m\omega} \right)$$

which will be needed throughout this problem set.

---

(27) (The harmonic oscillator I)

(a) Express $\hat{p}$ and $\hat{x}$ in terms of $\hat{a}$ and $\hat{a}^\dagger$.

(b) Compute the expectation values of $\hat{x}$ and $\hat{p}$ in the state $|0\rangle \in \mathcal{H}$ defined by the condition

$$\hat{a}|0\rangle = 0, \quad \langle 0|0\rangle = 1. \tag{B.98}$$

[*Hint*: use the fact that $\langle 0|\hat{a}^\dagger = (\hat{a}|0\rangle)^\dagger$.]

(c) Using the commutator $[\hat{a}, \hat{a}^\dagger] = 1$ compute the uncertainties $\Delta\hat{x}$ and $\Delta\hat{p}$ in the same state $|0\rangle$. By comparing to the problem (24) on wave-functions which saturate the lower bound in the uncertainty relation for $\hat{x}$ and $\hat{p}$, derive the expression for the wave function $\Psi(x)$ that corresponds to the state $|0\rangle$.

(28) (The harmonic oscillator II)

Consider a state $|0\rangle$ which satisfies $\hat{a}\,|0\rangle = 0$.

Express this as a differential equation for the wave function $\Psi(x)$ corresponding to the state $|0\rangle$. Solve this equation to find $\Psi(x)$. Compare your answer here with the results of problem (27).

(29) (The harmonic oscillator III)

The state $|0\rangle$ is the unique (up to a phase) normalised state which satisfies $\hat{a}\,|0\rangle = 0$ and we found its wave function to be

$$\Psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} e^{-\frac{1}{2}m\omega x^2/\hbar}. \tag{B.99}$$

The states $|n\rangle$ for $n \geq 0$ are defined by

$$|n\rangle = \frac{1}{\sqrt{n!}}(\hat{a}^\dagger)^n|0\rangle. \tag{B.100}$$

(a) Let $\Psi_n(x)$ be the wave function describing the state $|n\rangle$. Show that

$$\Psi_n(x) = \frac{1}{\sqrt{n!2^n}}\left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}}(u+\partial_u)^n\, e^{-u^2/2}, \tag{B.101}$$

where $u = \sqrt{\frac{m\omega}{\hbar}}x$.

(b) Using the above, show that

$$\Psi_n(x) = \frac{1}{\sqrt{n!2^n}}\left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right) e^{-\frac{1}{2}m\omega x^2/\hbar} \tag{B.102}$$

where

$$H_n(u) = (-1)^n e^{u^2}\partial_u^n e^{-u^2} \tag{B.103}$$

are some functions.

[Hint: First show that $e^{-u^2/2}\partial_u e^{u^2/2}$ is an operator that acts on any function of u in the same way as the operator $u + \partial_u$. Hence deduce that $e^{-u^2/2}(\partial_u - u)e^{u^2/2} = \partial_u$. Use this to expand $(u+\partial_u)^n$.]

(c) The functions $H_n(u)$ are called the Hermite polynomials. Compute $H_0(u)$, $H_1(u)$, $H_2(u)$, and $H_3(u)$ explicitly. Prove that $H_n(u)$ is a polynomial for all n.

(30) (The harmonic oscillator IV: the Heisenberg picture)

Recall the Hamiltonian of the Harmonic oscillator

$$\hat{H} = \hbar\omega(\hat{a}^\dagger\hat{a} + \tfrac{1}{2}). \tag{B.104}$$

In Heisenberg picture we can write this as

$$\hat{H} = \hbar\omega(\hat{a}^\dagger(t)\hat{a}(t) + \tfrac{1}{2}). \tag{B.105}$$

We don't write $\hat{H}(t)$ because the Hamiltonian is time-independent even in the Heisenberg picture (energy is a conserved quantity). The Heisenberg equation of motion is

$$\frac{\partial}{\partial t}\hat{\mathcal{O}}(t) = \frac{i}{\hbar}[\hat{H}, \hat{\mathcal{O}}(t)]. \tag{B.106}$$

(a) Using $[\hat{a}(t), \hat{a}^\dagger(t)] = 1$, show that the Heisenberg equation of motion for $\hat{a}(t)$ is

$$\frac{\partial}{\partial t}\hat{a}(t) = -i\omega\hat{a}(t). \tag{B.107}$$

[Hint: First check the commutator identity $[A, BC] = [A, B]C + B[C, A]$ holds and then apply it in the problem.]

(b) Show that $\hat{a}(t) = e^{-i\omega t}\hat{a}(0)$ solves the Heisenberg equation of motion for $\hat{a}(t)$.