

INTRODUCTORY QUANTUM THEORY (CM332C)

Neil Lambert
Department of Mathematics
King's College London
Strand
London WC2R 2LS, U. K.
Email: `neil.lambert@kcl.ac.uk`

Contents

1	historical development and early quantum mechanics	3
1.1	the photoelectric effect	3
1.2	the Rutherford/Bohr atom	4
1.3	the Compton effect	5
1.4	diffraction	6
1.5	de Broglie waves	6
1.6	wave-particle duality and the correspondence principle	7
2	wave mechanics	7
2.1	the Schrödinger equation	7
2.2	the interpretation of the wavefunction and basic properties	10
2.3	Ehrenfest's theorem	12
3	some simple systems	14
3.1	a particle in a box	14
3.2	a potential well	18
3.3	scattering and tunneling through a potential barrier	21
3.4	the Dirac δ -function and planewaves	24

4	formal quantum mechanics and some Hilbert space theory	26
4.1	Hilbert spaces and Dirac bracket notation	27
4.2	The example of L^2	33
4.3	Observables, eigenvalues and expectation values	37
4.4	unitarity and time evolution	45
5	noncommuting observables and uncertainty	47
5.1	Measuring observables that do not commute	48
5.2	uncertainty of an observable	51
5.3	The Heisenberg uncertainty principle	52
6	additional topics	54
6.1	the harmonic oscillator	54
6.2	symmetries in quantum mechanics	59
6.3	the Heisenberg picture	61
6.4	a definition of quantum mechanics	62

Acknowledgement

These notes are indebted to the notes of M. Gaberdiel and others. The introduction is largely taken from the book by Messiah.

1 historical development and early quantum mechanics

By the end of the nineteenth century theoretical physicists thought that soon they could pack up their bags and go home. They had developed a powerful mathematical theory, classical mechanics, which seemed to describe just about all that they observed, with the exception of a few sticking points. In particular the classical world was ruled by Newtonian physics where matter (atoms) interacted with a radiation field (light) as described by Maxwell's equations. However as it turned out these sticking points were not smoothed over at all but rather were glimpses of the microscopic (and relativistic) world which was soon to be experientially discovered. As has been stated countless times, by the end of the first decade of the twentieth century quantum mechanics and relativity had appeared and would soon cause classical mechanics, with its absolute notions of space, time and determinacy to be viewed as an approximation.

Historically the first notion of quantized energy came in 1900 with Planck's hypothesis that the energy contained in radiation could only be exchanged in discrete lumps, known as "quanta", with matter. This in turn implies that the energy, E , of radiation is proportional to its frequency, ν ,

$$E = h\nu \tag{1.1}$$

with h a constant, Planck's constant. This allowed him to derive his famous formula for black body spectra which was in excellent agreement with experiment. From our perspective this is essentially a thermodynamic issue and the derivation is therefore out of the main theme of this course.

The derivation of such a formula from first principles was one of the sticking points mentioned above. While Planck's derivation is undoubtedly the first appearance of quantum ideas it was not at all clear at the time that this was a fundamental change, or that there was some underlying classical process which caused the discreteness, or that it was even correct. However there were further experiments which did deeply challenge the continuity of the world and whose resolution relies heavily on Planck's notion of quanta.

1.1 the photoelectric effect

When you shine a beam of light onto certain materials it will induce a flow of electrons. It was found experimentally that while the number of electrons produced depended on the intensity of the light, their speed, and hence their energy, only depended on the frequency of the light. In addition there was a critical frequency below which no electron current would be produced.

This was very mysterious at the time since the classical theory of light stated that light was a wave and the energy carried by a wave is proportional to the square of its amplitude, which in this case was the intensity. So one would expect the electrons to have an energy that depended on the intensity. Furthermore, no matter how small an

amount of energy the wave contains, if you wait long enough it should impart enough energy to the electrons to produce a current. So there shouldn't be a critical frequency.

In 1905 Einstein proposed a solution to this problem by picking up on Planck's hypothesis. He took the addition step of saying that light itself was made of particles (recall that Planck had merely said that light only transmits energy to matter in quantized amounts), each of which carries an energy $E = h\nu$. If we then postulate that the electrons are bound to the material with some constant potential $-W$, the energy of the electrons will be

$$E = \frac{1}{2}mv^2 = h\nu - W \quad (1.2)$$

This curve fitted the experimental data very well with the same choice of h that Planck had needed. Namely there is a cut-off in frequency below which no electron current is produced (since $E \geq 0$), followed by a simple linear behaviour, that depends on frequency by not intensity.

1.2 the Rutherford/Bohr atom

In 1911 Rutherford performed experiments to probe the structure of the atom. The surprising result was the the atom appear to be mainly empty space, with a positively charged nucleus at the centre about which the negatively charged electrons orbit. This raised an immediate concern. If the electrons are orbiting then they must be accelerating. Accelerating charges produce radiation and therefore lose energy. As a result the electrons should quickly spiral into the nucleus and the atom will be unstable, decaying in a burst of radiation. Our existence definitely contradicts this prediction.

A further experimental observation, dating as far back as 1884, was that the spectra of atoms was not continuous but discrete. Balmer observed that the frequency of emitted light by Hydrogen obeyed the formula

$$\nu = R\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right) \quad (1.3)$$

where n_1 and n_2 are integers and R is a constant.

In 1913 Bohr proposed a model for the hydrogen atom. He declared that the electrons were only allowed to sit in discrete orbits labeled by an integer n . In each orbit the angular momentum is quantized

$$mv_n r_n = n\hbar \quad (1.4)$$

where $\hbar = h/2\pi$ and m is the mass of the electron. The integer n is known as a "quantum number" and illustrates the fundamental discreteness of quantum processes.

Furthermore there must be a balance of forces in the orbit, the electrical attraction of the electron to the nucleus is equal to the centrifugal force

$$\frac{e^2}{r_n^2} = \frac{mv_n^2}{r_n} \quad (1.5)$$

From these equations we can determine the radius

$$\begin{aligned} m^2 \frac{e^2}{mr_n} r_n^2 &= n^2 \hbar^2 \\ r_n &= \frac{n^2 \hbar^2}{me^2}. \end{aligned} \tag{1.6}$$

The total energy of the electron is

$$\begin{aligned} E &= \frac{1}{2}mv_n^2 - \frac{e^2}{r_n} \\ &= \frac{1}{2}m \frac{e^2}{mr_n} - \frac{e^2}{r_n} \\ &= -\frac{me^4}{2\hbar^2} \frac{1}{n^2}. \end{aligned} \tag{1.7}$$

We can now see that if a photon is emitted from an electron which then jumps from the n_1 th to the n_2 th orbit ($n_1 > n_2$) it therefore has energy

$$E_\gamma = E_{n_1} - E_{n_2} = \frac{me^4}{2\hbar^2} \left(\frac{1}{n_2^2} - \frac{1}{n_1^2} \right) \tag{1.8}$$

And hence the frequency is

$$\nu = \frac{E_\gamma}{2\pi\hbar} = R \left(\frac{1}{n_2^2} - \frac{1}{n_1^2} \right) \tag{1.9}$$

with $R = me^4/4\pi\hbar^3$. Thus the Bohr model reproduces the Balmer formula and predicts the value of the constant R .

1.3 the Compton effect

Another crucial experiment was performed by Compton in 1924. He found that when light scattered off electrons there was shift in the wavelength of the light given by

$$\Delta\lambda = \frac{h}{mc}(1 - \cos\theta) \tag{1.10}$$

where m is the mass of the electron and θ the angle through which the light is deflected.

The problem is that in the classical theory of light interacting with matter the electron will continuously absorb the light and then radiate some of it off. However in the rest frame of the electron the radiated light will be spherically symmetric. In addition, some energy will be imparted to the electron and push it continuously along the direction of the incident beam of light. There will be a resulting shift in the wavelength

of the emitted light however due to the Doppler effect since the electron is now moving, but this implies that $\Delta\lambda$ should depend on λ .

On the other hand if we assume that light is made of particles and use the relativistic expressions for energy and momentum, along with Planck's formula, then one readily obtains the Compton scattering formula from conservation of energy and momentum (*see the problem sets*). In effect the difference is that Compton scattering is what one would expect from two particles, such as a marble bouncing off a tennis ball. On the other hand if light is a wave and energy is continuously transferred to the electron then the scattering is more like what you'd get by blowing on a tennis ball, namely the ball will move away along the direction of the wind.

1.4 diffraction

The experiments described above might leave one with the impression that after all light is just made of particles called photons. However this can not be the case. To illustrate this we consider another famous class of experiments where light is diffracted. In particular consider the double slit experiment where a coherent source of light is directed at a wall which has two slits. The light then spreads out from the two slits and the resulting pattern is observed on another wall.

The observed pattern on the wall shows constructive and destructive interference coming from the two slits. This is a familiar phenomenon of waves and you can reproduce it in your bath tub. One can lower the intensity of the light so that only one photon (assuming that such quanta exist) is emitted at a time however the pattern persists. How can this be produced if light is a particle? A further curiosity is that if you place a detector by the slits to see which route the light takes then the pattern disappears. Light therefore most definitely has wave-like properties.

1.5 de Broglie waves

In 1923 de Broglie postulated that if light has both wave and particle properties then matter does too. He proposed a generalization of Planck formula

$$p = \frac{h}{\lambda}, \quad (1.11)$$

where p is the momentum of the particle and λ is its wavelength. Note that if we use the relativistic formula for the energy of light, $E = pc$, along with $\nu = c/\lambda$, then we recover Planck's formula $E = h\nu$. This condition can also be viewed as the requirement that the wavelength of the electron in the n th orbit fits exactly n times around the orbit. It also follows that the wavelength of light emitted from the atom is much longer than the wavelength of the electron. This implies that the wave-like nature of the electron (or indeed the atom itself) can not be probed by the photons that arise in typical (low energy) atomic processes (*see the problem sets*).

If particles such as the electron also behave as waves and should therefore cause diffraction patterns. Remarkably the diffraction of electrons off crystals was observed by Davisson and Germer in 1927.

1.6 wave-particle duality and the correspondence principle

The ambiguous notion of light and matter as waves and/or particles lead to the notion of wave-particle duality. Namely that sometimes light and matter behave as though they are waves and sometimes as though they are particles. It seems odd but we must just get used to the idea. From our intuition gained in the classical world this seems contradictory by that is due to our limited experience. This is our fault.

Since the quantum world is not what we observe in our macroscopic experiences it must be that, in the limit of a large number of quantum mechanical processes, the results of classical mechanics are recovered. For example in the Compton effect we contrasted a marble scattering off a tennis ball with the effect of blowing on a tennis ball. The former being an analogue of the quantum scattering of a photon off an electron and the latter an analogue of the classical scattering of light off an electron. However wind is made up of many air molecules which do behave as particles and hence bounce off the tennis ball in much the same way as a marble does. In the limit of many such collisions of air molecules with the tennis ball one recovers the continuous, classical behaviour. The requirement on quantum theory that in the limit of large quantum numbers the classical answers are reproduced is known as the correspondence principle.

2 wave mechanics

Now that we have introduced the basic notions of quantum mechanics we can move on to develop a system of wave mechanics. The guiding point is to accept that particles, as well as light, should be described by a kind of wave. This was developed primarily by Heisenberg and Schrödinger in the late 1920's.

2.1 the Schrödinger equation

Let us begin by restricting our attention to one spatial dimension. To proceed we suppose that a free particle is described by a plane wave

$$\Psi(x, t) = e^{ikx - i\omega t} . \quad (2.1)$$

Note that the x variable is periodic with a period $2\pi/k$ thus the wavelength is $\lambda = 2\pi/k$. On the other hand the frequency is the number of wavelengths that pass a given point per unit time, $\nu = 2\pi\omega$. Using the relations

$$p = \frac{2\pi\hbar}{\lambda} = \hbar k , \quad E = 2\pi\hbar\nu = \hbar\omega , \quad (2.2)$$

we see that

$$i\hbar \frac{\partial \Psi}{\partial t} = E\Psi, \quad -i\hbar \frac{\partial \Psi}{\partial x} = p\Psi \quad (2.3)$$

Thus the energy and momentum of the wave are eigenvalues of the operators

$$\hat{E} \equiv i\hbar \frac{\partial}{\partial t} \quad \text{and} \quad \hat{p} \equiv -i\hbar \frac{\partial}{\partial x} \quad (2.4)$$

Here a hat over a symbol implies that we are considering it as an operator, *i.e.* something that acts on our “wavefunction” Ψ and produces another wavefunction.

Definition An operator \hat{O} is simply a map from the space of functions to itself.

We will always consider here linear operators, *i.e.* ones which obey $\hat{O}(a\Psi_1 + b\Psi_2) = a\hat{O}\Psi_1 + b\hat{O}\Psi_2$ for any two functions Ψ_1 and Ψ_2 and numbers a and b .

Definition If for some operator \hat{O} there is a function Ψ and a number λ such that $\hat{O}\Psi = \lambda\Psi$ then Ψ is called an eigenfunction of \hat{O} and λ its eigenvalue.

Our next step is to construct an equation for the wavefunction. Since we want the solution to have wave-like properties, such as destructive and constructive interference, we should look for a linear equation. This is known as the superposition principle.

We further want to impose the equation for the energy, namely $E = p^2/2m$. Thus we postulate that the correct equation is

$$\hat{E}\Psi = \frac{\hat{p}^2}{2m}\Psi \quad \text{or} \quad i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2}. \quad (2.5)$$

This is the free Schrödinger equation in one spatial dimension. Substituting in our plane wave (2.1) we indeed find that $E^2 = p^2/2m$. It is clear that there are other equations that might satisfy our requirements (linearity and the correct energy/momentum relation) but this is the simplest.

Definition: A solution to the Schrödinger equation at fixed time t is called the state of the system at time t .

There are two generalizations that we can make. If we instead have a particle that is not free but moving in a potential then we simply use the energy relation, $E = p^2/2m + V(x)$ to obtain

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V(x)\Psi. \quad (2.6)$$

Our second generalization is to return to three-dimensional space. In three dimensions a plane wave takes the form

$$\Psi(\vec{x}, t) = e^{i\vec{k}\cdot\vec{x} - i\omega t}. \quad (2.7)$$

where now the “wave number” \vec{k} is a three component vector. The energy operator remains the same but now the momentum operator is also three component vector

$$\hat{p} \equiv -i\hbar\vec{\nabla} = -i\hbar \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} . \quad (2.8)$$

The full Schrödinger equation in three-dimensional space is therefore

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi + V(\vec{x}) \Psi . \quad (2.9)$$

where

$$\nabla^2 = \vec{\nabla} \cdot \vec{\nabla} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} . \quad (2.10)$$

For most of this course we will restrict attention to the the one-dimensional Schrödinger equation since this captures all of the essential physics and provides solvable problems.

Often the potential V is independent of time. In which case we can use separation of variables to write $\Psi(\vec{x}, t) = \psi(\vec{x})f(t)$. Substituting into the Schrödinger equation and dividing by ψf gives

$$i\hbar \frac{1}{f} \frac{\partial f}{\partial t} = -\frac{\hbar^2}{2m} \frac{1}{\psi} \nabla^2 \psi + V(\vec{x}) \quad (2.11)$$

Now the left and right hand sides depend only on t and \vec{x} respectively and hence they must in fact be constant since they are equal. Thus we find two equations

$$i\hbar \frac{\partial f}{\partial t} = Ef , \quad -\frac{\hbar^2}{2m} \nabla^2 \psi + V(\vec{x})\psi = E\psi , \quad (2.12)$$

where E is a constant. We can immediately solve the first equation to find

$$\Psi(\vec{x}, t) = e^{-iEt/\hbar} \psi(\vec{x}) . \quad (2.13)$$

The second equation for $\psi(\vec{x})$ is known as the time independent Schrödinger equation. For a given V and set of boundary conditions the general theory of differential equations says that there will be a linearly independent set of solutions ψ_n to the time independent Schrödinger equation with eigenvalues E_n . Linearity of the Schrödinger equation implies that the general solution has the form

$$\Psi(\vec{x}, t) = \sum_n A_n e^{-iE_n t/\hbar} \psi_n(\vec{x}) \quad (2.14)$$

note that solutions may well be labeled by more than one quantum number n .

2.2 the interpretation of the wavefunction and basic properties

We need to provide an interpretation for $\Psi(\vec{x}, t)$ and establish some basic properties.

In analogy with light, where the square of the amplitude of a wave gives the intensity of the light, we make the identification that

$$\rho(\vec{x}, t) = |\Psi(\vec{x}, t)|^2 \quad (2.15)$$

gives the probability density of finding the particle at the point \vec{x} at time t . With this interpretation it is clear that we must consider wavefunctions which are normalized

$$\int d^3x |\Psi(\vec{x}, t)|^2 = 1 \quad (2.16)$$

so that the probability that the particle is somewhere is one.

Definition: A state is called normalizable if it satisfies 2.16

Next we'd like to introduce a notion of a probability current, $\vec{j}(\vec{x}, t)$. This is defined so that

$$\frac{d\rho}{dt} + \vec{\nabla} \cdot \vec{j} = 0. \quad (2.17)$$

To construct it we notice that

$$\begin{aligned} \frac{d\rho}{dt} &= \Psi^* \frac{\partial \Psi}{\partial t} + \frac{\partial \Psi^*}{\partial t} \Psi \\ &= \frac{i\hbar}{2m} \Psi^* \nabla^2 \Psi - \frac{i}{\hbar} V \Psi^* \Psi - \frac{i\hbar}{2m} \Psi \nabla^2 \Psi^* + \frac{i}{\hbar} V \Psi \Psi^* \\ &= \vec{\nabla} \cdot \frac{i\hbar}{2m} (\Psi \vec{\nabla} \Psi^* - \Psi^* \vec{\nabla} \Psi) \end{aligned} \quad (2.18)$$

Thus we find

$$\vec{j}(\vec{x}, t) = \frac{i\hbar}{2m} (\Psi \vec{\nabla} \Psi^* - \Psi^* \vec{\nabla} \Psi). \quad (2.19)$$

The probability current measures the amount of probability that leaks out of a given region through its surface. That is, if the probability of finding the particle in a region R is

$$P(t) = \int_R d^3x \rho(\vec{x}, t) \quad (2.20)$$

then

$$\begin{aligned} \frac{dP}{dt} &= \int_R d^3x \frac{d\rho}{dt} \\ &= - \int_R d^3x \vec{\nabla} \cdot \vec{j} \\ &= - \int_{\partial R} d^2\vec{x} \cdot \vec{j} \end{aligned} \quad (2.21)$$

With this probabilistic interpretation we can define the expectation value of the location of the particle as

$$\langle \vec{x} \rangle = \int d^3x \vec{x} \rho(\vec{x}, t) = \int d^3x \Psi^* \vec{x} \Psi \quad (2.22)$$

Similarly the expectation values of the energy and momentum are

$$\langle E \rangle = \int d^3x \Psi^* \hat{E} \Psi, \quad \langle \vec{p} \rangle = \int d^3x \Psi^* \hat{\vec{p}} \Psi, \quad (2.23)$$

These are interpreted as the average value of the quantity as measured after many experiments. Note the position of the operator on the right hand side. Since it acts on wavefunctions, and not the product of two wavefunctions, we must insert it so that it only acts on one and not the other. With this in mind we see that the position operator \hat{x} acts on wavefunctions as

$$\hat{x} \Psi(\vec{x}, t) = x \Psi(\vec{x}, t) \quad (2.24)$$

Note that this probabilistic definition implies that there is an uncertainty in the position of the particle and also its energy and momentum, analogous to the notion of standard deviation. To see this we can introduce the uncertainty in the components of the position, say x , as

$$(\Delta \hat{x})^2 = \langle (\hat{x})^2 \rangle - (\langle \hat{x} \rangle)^2 \quad (2.25)$$

and similarly for $\Delta \hat{y}$, $\Delta \hat{z}$, $\Delta \hat{E}$ and $\Delta \vec{\hat{p}}$. Generically $\Delta \hat{x}$ is non-vanishing. For example if $\Psi(\vec{x}, t)$ is an even or odd function of x , then $\langle \hat{x} \rangle$ vanishes however $\langle \hat{x}^2 \rangle$ is the integral of an everywhere positive function and so cannot vanish.

On the other hand if Ψ is in an eigenstate of \hat{E} with eigenvalue E then

$$\begin{aligned} \langle \hat{E} \rangle &= \int d^3x \Psi^* \hat{E} \Psi = E \int d^3x \Psi^* \Psi = E \\ \langle \hat{E}^2 \rangle &= \int d^3x \Psi^* \hat{E}^2 \Psi = E \int d^3x \Psi^* \hat{E} \Psi = E^2 \int d^3x \Psi^* \Psi = E^2 \end{aligned} \quad (2.26)$$

so that $\langle \Delta \hat{E} \rangle = \sqrt{E^2 - E^2} = 0$. Thus these states have a definite value of the energy. But a general solution, which is a superposition of different energy eigenstates will not satisfy such a simple relation. Clearly a similar result holds for $\langle \vec{\hat{p}} \rangle$.

N.B.: Ψ itself is not directly measurable. In particular rotating Ψ by a constant phase has no physical effect.

From here we prove the following theorems:

Theorem: $\langle \vec{\hat{p}} \rangle$ is real (for a normalizable state).

Proof

$$\begin{aligned}
\langle \vec{p} \rangle &= -i\hbar \int d^3x \Psi^* \vec{\nabla} \Psi \\
&= i\hbar \int d^3x (\vec{\nabla} \Psi^*) \Psi - i\hbar \int d^3x \vec{\nabla} (\Psi^* \Psi) \\
&= i\hbar \int d^3x \Psi \vec{\nabla} \Psi^* \\
&= \langle \vec{p} \rangle^*
\end{aligned} \tag{2.27}$$

where we dropped a total derivative which vanishes if Ψ is normalizable.

Another important theorem is

Theorem: If $V(\vec{x}, t) \geq V_0$ for all \vec{x} then $\langle \hat{E} \rangle$ is real and $\langle \hat{E} \rangle \geq V_0$ for any normalizable state.

Proof: We have that

$$\begin{aligned}
\langle \hat{E} \rangle &= i\hbar \int d^3x \Psi^* \frac{\partial \Psi}{\partial t} \\
&= -\frac{\hbar^2}{2m} \int d^3x \Psi^* \nabla^2 \Psi + \int d^3x \Psi^* V \Psi \\
&= -\frac{\hbar^2}{2m} \int d^3x \vec{\nabla} \cdot (\Psi^* \vec{\nabla} \Psi) + \frac{\hbar^2}{2m} \int d^3x \vec{\nabla} \Psi^* \cdot \vec{\nabla} \Psi + \int d^3x V \Psi^* \Psi \\
&= \frac{\hbar^2}{2m} \int d^3x \vec{\nabla} \Psi^* \cdot \vec{\nabla} \Psi + \int d^3x V \Psi^* \Psi \\
&\geq \int d^3x V_0 \Psi^* \Psi = V_0
\end{aligned} \tag{2.28}$$

Here in the second line we used the Schrödinger equation, in the third line we integrated by parts and in the fourth line we dropped the total derivative term since normalizability of the wavefunction implies that $\Psi(\vec{x}, t) \rightarrow 0$ as $x \rightarrow \infty$. The reality of $\langle \hat{E} \rangle$ follows from the fourth line.

Note that the essential point of this theorem can also be stated as $\langle \hat{p}^2 \rangle \geq 0$.

N.B.: In infinite volume a normalisable state must vanish suitably quickly at infinity for the integral to be convergent. In general we will assume that wavefunctions always vanish sufficiently quickly so that various surface terms evaluated at infinity also vanish.

2.3 Ehrenfest's theorem

Where does classical mechanics arise from? The answer is generally given in part by the following theorem:

Erhenfest's Theorem: The expectation value $\langle \vec{\hat{x}} \rangle$ satisfies the classical equations

$$m \frac{d}{dt} \langle \vec{\hat{x}} \rangle = - \langle \vec{\hat{p}} \rangle, \quad m \frac{d^2}{dt^2} \langle \vec{\hat{x}} \rangle = - \langle \vec{\nabla} V \rangle \quad (2.29)$$

Proof: It is sufficient to consider a single component of \vec{x} , say x , the other components follow in the same way.

$$\begin{aligned} m \frac{d}{dt} \langle \hat{x} \rangle &= m \int d^3x x \left(\frac{d\Psi^*}{dt} \Psi + \Psi^* \frac{d\Psi}{dt} \right) \\ &= \int d^3x x \left(-\frac{i\hbar}{2} \nabla^2 \Psi^* \Psi + \frac{im}{\hbar} V \Psi^* \Psi + \frac{i\hbar}{2} \Psi^* \nabla^2 \Psi - \frac{im}{\hbar} V \Psi^* \Psi \right) \\ &= \int d^3x x \left(-\frac{i\hbar}{2} \nabla^2 \Psi^* \Psi + \frac{i\hbar}{2} \Psi^* \nabla^2 \Psi \right) \\ &= -\frac{i\hbar}{2} \int d^3x \vec{\nabla} \cdot (x (\vec{\nabla} \Psi^* \Psi - \Psi^* \vec{\nabla} \Psi)) \\ &\quad + \frac{i\hbar}{2} \int d^3x x (\vec{\nabla} \Psi^* \cdot \vec{\nabla} \Psi - \vec{\nabla} \Psi^* \cdot \vec{\nabla} \Psi) \\ &\quad + \frac{i\hbar}{2} \int d^3x \left(\frac{\partial \Psi^*}{\partial x} \Psi - \Psi^* \frac{\partial \Psi}{\partial x} \right) \end{aligned} \quad (2.30)$$

In the last line the second term vanishes and the first term is a total derivative which can be dropped. The last term is simply $(\langle \hat{p}_x \rangle^* - \langle \hat{p}_x \rangle)/2 = \langle \hat{p}_x \rangle$.

To prove the second equation we differentiate and use the Schrödinger equation again

$$\begin{aligned} m \frac{d^2}{dt^2} \langle \hat{x} \rangle &= -i\hbar \frac{\partial}{\partial t} \int d^3x \Psi^* \frac{\partial \Psi}{\partial x} \\ &= -i\hbar \int d^3x \left(\Psi^* \frac{\partial^2 \Psi}{\partial x \partial t} + \frac{\partial \Psi^*}{\partial t} \frac{\partial \Psi}{\partial x} \right) \\ &= - \int d^3x \Psi^* \frac{\partial}{\partial x} \left(-\frac{\hbar^2}{2m} \nabla^2 \Psi + V \Psi \right) \\ &\quad + \int d^3x \left(-\frac{\hbar^2}{2m} \nabla^2 \Psi^* + V \Psi^* \right) \frac{\partial \Psi}{\partial x} \\ &= - \int d^3x \Psi^* \frac{\partial V}{\partial x} \Psi \\ &\quad + \frac{\hbar^2}{2m} \int d^3x \left(\Psi^* \nabla^2 \frac{\partial \Psi}{\partial x} - \nabla^2 \Psi^* \frac{\partial \Psi}{\partial x} \right) \\ &= - \int d^3x \Psi^* \frac{\partial V}{\partial x} \Psi \end{aligned}$$

$$\begin{aligned}
& + \frac{\hbar^2}{2m} \int d^3x \vec{\nabla} \cdot \left(\Psi^* \vec{\nabla} \frac{\partial \Psi}{\partial x} - \vec{\nabla} \Psi^* \frac{\partial \Psi}{\partial x} \right) \\
& = - \left\langle \frac{\partial V}{\partial x} \right\rangle
\end{aligned}$$

where we dropped a total derivative in the last line.

Thus the expectation values of the physical quantities in quantum mechanics satisfy the classical equations of motion. The classical limit is therefore reached when the uncertainties in the various quantities are small. According to the correspondence principle we expect that for large quantum numbers the probability distributions of physical quantities become sharply peaked around the classical answers. We will see this in all examples however the general theory of this is rather involved. Indeed there are examples of macroscopic systems which display distinctly quantum behaviour, such as superfluidity and superconductivity.

3 some simple systems

Before going on to understand the formalism in greater detail it will be helpful to first consider some simple physical applications.

3.1 a particle in a box

Perhaps the simplest example of a quantum mechanical system is to consider a particle confined to a box $0 \leq x \leq l$, but otherwise free. Since the wavefunction is interpreted as the probability of finding the particle at a given point, it must vanish outside the box. Therefore it must vanish on the side of the box. So we need to solve the free one-dimensional Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} \quad (3.31)$$

subject to $\Psi(0, t) = \Psi(l, t) = 0$.

We can go to a basis of energy eigenstates and use the time-independent Schrödinger equation. To further illustrate the point let us go through the separation of variables steps again in detail using $\Psi(x, t) = \psi(x)f(t)$. Substituting into the Schrödinger equation and dividing by ψf gives

$$i\hbar \frac{1}{f} \frac{\partial f}{\partial t} = -\frac{\hbar^2}{2m} \frac{1}{\psi} \frac{\partial^2 \psi}{\partial x^2} \quad (3.32)$$

Now the left and right hand sides depend only on t and x respectively and hence they must in fact be constant since they are equal. Thus we find two equations

$$i\hbar \frac{\partial f}{\partial t} = \frac{\alpha \hbar^2}{2m} f, \quad \frac{\partial^2 \psi}{\partial x^2} = -\alpha \psi, \quad (3.33)$$

where α is a constant. What we called energy before is now $E = \alpha\hbar^2/2m$.

Now if $\alpha < 0$ then the solution to the second equation is simply

$$\psi = A \cosh(\sqrt{-\alpha}x) + B \sinh(-\sqrt{-\alpha}x). \quad (3.34)$$

Imposing $\psi(0) = 0$ implies that $A = 0$. Similarly imposing $\psi(l) = 0$ implies that $B \sinh(-\sqrt{-\alpha}l) = 0$. However there is no non-trivial solution to this. Hence there are no relevant solutions when $\alpha < 0$. If $\alpha = 0$ then the solution is clearly just a linear function which cannot vanish at the end points if it doesn't vanish everywhere.

For $\alpha > 0$ we have

$$\psi = A \cos(\sqrt{\alpha}x) + B \sin(-\sqrt{\alpha}x). \quad (3.35)$$

Imposing $\psi(0) = 0$ implies that $A = 0$. Similarly imposing $\psi(l) = 0$ implies that $B \sin(-\sqrt{\alpha}l) = 0$ or

$$\sqrt{\alpha}l = \pi n \leftrightarrow \alpha = \frac{\pi^2 n^2}{l^2}, \quad (3.36)$$

for some integer n . Here n is an example of a “quantum number” and takes any positive value, $n=1,2,3,\dots$ (if $n = 0$ then $\psi = 0$, in addition n and $-n$ give the same ψ up to a minus sign).

The solution for f is

$$f(t) = f(0)e^{-\frac{i\hbar\alpha}{2m}t} = f(0)e^{-\frac{i\hbar\pi^2 n^2}{2ml^2}t} \quad (3.37)$$

so that the full solution is

$$\Psi_n(x, t) = N_n e^{-\frac{i\hbar\pi^2 n^2}{2ml^2}t} \sin\left(\frac{\pi n}{l}x\right). \quad (3.38)$$

We need to determine N_n by normalizing the wavefunction

$$\begin{aligned} 1 &= \int_0^l dx |\Psi_n|^2 \\ &= N_n^2 \int_0^l dx \sin^2\left(\frac{\pi n}{l}x\right) \\ &= N_n^2 \frac{l}{2} \end{aligned} \quad (3.39)$$

hence $N_n = \sqrt{2/l}$.

So we have constructed an infinite set of wavefunctions labeled by n

$$\Psi_n(x, t) = \sqrt{\frac{2}{l}} e^{-\frac{i\hbar\pi^2 n^2}{2ml^2}t} \sin\left(\frac{\pi n}{l}x\right). \quad (3.40)$$

In fact on general grounds these form a basis for the set of all solutions, the general solution taking the form

$$\Psi(x, t) = \sum_n A_n \sqrt{\frac{2}{l}} e^{-\frac{i\hbar\pi^2 n^2}{2ml^2}t} \sin\left(\frac{\pi n}{l}x\right). \quad (3.41)$$

Normalization now implies that $\sum_n |A_n|^2 = 1$.

However the solutions $\Psi(x, t)$ are special since they are eigenfunctions of the energy operator $\hat{E} = i\hbar\partial/\partial t$

$$i\hbar \frac{\partial \Psi_n}{\partial t} = \frac{\hbar^2 \pi^2 n^2}{2ml^2} \Psi_n \quad (3.42)$$

As we saw this means that they have a definite value of energy given by its eigenvalue of \hat{E} ;

$$E_n = \frac{\hbar^2 \pi^2 n^2}{2ml^2} . \quad (3.43)$$

These are therefore the allowed energies for a particle in a box. The more general solutions must represent a system with no definite value of the energy.

We can also evaluate $\langle \hat{x} \rangle$ and $\langle \hat{x}^2 \rangle$. These give

$$\begin{aligned} \langle \hat{x} \rangle &= \int_0^l dx x |\Psi_n|^2 \\ &= \frac{2}{l} \int_0^l dx x \sin^2\left(\frac{\pi n}{l} x\right) \\ &= \frac{2}{l} \left(\frac{l}{\pi n}\right)^2 \int_0^{\pi n} dy y \sin^2(y) \\ &= \frac{l}{\pi^2 n^2} \int_0^{n\pi} dy y (1 - \cos(2y)) \\ &= \frac{l}{\pi^2 n^2} \left(\frac{1}{2}y^2 - \frac{1}{2}y \sin(2y) - \frac{1}{4} \cos(2y)\right)_0^{n\pi} \\ &= \frac{l}{2} \end{aligned} \quad (3.44)$$

and also

$$\begin{aligned} \langle x^2 \rangle &= \int_0^l dx x^2 |\Psi_n|^2 \\ &= \frac{2}{l} \int_0^l dx x^2 \sin^2\left(\frac{\pi n}{l} x\right) \\ &= \frac{2}{l} \left(\frac{l}{\pi n}\right)^3 \int_0^{\pi n} dy y^2 \sin^2(y) \\ &= \frac{l^2}{n^3 \pi^3} \int_0^{n\pi} dy y^2 (1 - \cos(2y)) \\ &= \frac{l^2}{n^3 \pi^3} \left(\frac{1}{3}y^3 - \frac{1}{2}y^2 \sin(2y) - \frac{1}{2}y \cos(2y) + \frac{1}{4} \sin(2y)\right)_0^{n\pi} \\ &= \frac{l^2}{n^3 \pi^3} \left(\frac{n^3 \pi^3}{3} - \frac{n\pi}{2}\right) \end{aligned} \quad (3.45)$$

Thus we find that the uncertainty in the position is

$$(\Delta\hat{x})^2 = \langle \hat{x}^2 \rangle^2 - \langle \hat{x} \rangle^2 = \frac{l^2}{3} - \frac{l^2}{2n^2\pi^2} - \left(\frac{l}{2}\right)^2 = \frac{l^2}{12} - \frac{l^2}{2n^2\pi^2} \quad (3.46)$$

The n -independent terms are “classical”, *i.e.* they are what you’d expect from saying that the particle was in the box but that you didn’t know where in the box it is, *viz.*,

$$\begin{aligned} \langle x \rangle_{cl} &= \frac{1}{l} \int_0^l dx x = \frac{l}{2} \\ \langle x^2 \rangle_{cl} &= \frac{1}{l} \int_0^l dx x^2 = \frac{l^2}{3} \end{aligned} \quad (3.47)$$

Note that the $1/l$ factor in front is there as a normalisation, *i.e.* so that the probability that the particle is somewhere is one

$$1 = \frac{1}{l} \int_0^l dx \quad (3.48)$$

We see that as $n \rightarrow \infty$ the quantum mechanical predictions approach those of the classical theory.

Note also that Ψ_n modes are not eigenfunctions of the momentum operator $\hat{p} = -i\hbar\partial/\partial x$. Therefore they do not have a definite momentum, *i.e.* $\Delta p \neq 0$. To see this let us calculate $\langle \hat{p} \rangle$ and $\langle \hat{p}^2 \rangle$.

$$\begin{aligned} \langle \hat{p} \rangle &= -i\hbar \int_0^l dx \Psi_n^* \frac{\partial \Psi_n}{\partial x} \\ &= -i\hbar \frac{2}{l} \frac{\pi n}{l} \int_0^l dx \sin\left(\frac{\pi n}{l}x\right) \cos\left(\frac{\pi n}{l}x\right) \\ &= -\frac{i\hbar\pi n}{l^2} \int_0^l dx \sin\left(2\frac{\pi n}{l}x\right) \\ &= 0 \end{aligned} \quad (3.49)$$

This is to be expected from the symmetry of the problem. On the other hand

$$\begin{aligned} \langle \hat{p}^2 \rangle &= -\hbar^2 \int_0^l dx \Psi_n^* \frac{\partial^2 \Psi_n}{\partial x^2} \\ &= \hbar^2 \frac{2}{l} \left(\frac{\pi n}{l}\right)^2 \int_0^l dx \sin^2\left(\frac{\pi n}{l}x\right) \\ &= \frac{\hbar^2 n^2 \pi^2}{l^2} \end{aligned} \quad (3.50)$$

so that

$$(\Delta\hat{p})^2 = \frac{\hbar^2 n^2 \pi^2}{l^2} \quad (3.51)$$

Thus we see that Erhenfest's theorem is indeed satisfied:

$$\langle \hat{E} \rangle = E_n = \frac{\hbar^2 \pi^2 n^2}{2ml^2} = \frac{\langle \hat{p}^2 \rangle}{2m} \quad (3.52)$$

Note that due to uncertainty

$$\langle \hat{E} \rangle \neq \frac{\langle \hat{p} \rangle^2}{2m} \quad (3.53)$$

3.2 a potential well

Next we consider a particle in a potential well. Here it is allowed to live over the entire real line but there is a potential well:

$$V(x) = \begin{cases} -V_0 & \text{if } |x| \leq l, \\ 0 & \text{if } |x| > l, \end{cases} \quad (3.54)$$

where l and V_0 are constants with $V_0 > 0$. We break the problem up into three regions $x < -l$, $|x| \leq l$ and $x > l$ which we label by I, II and III respectively.

We will consider energy eigenstates

$$\Psi(x, t) = e^{-iEt/\hbar} \psi(x) \quad (3.55)$$

so that

$$-\frac{\hbar^2}{2m} \frac{d^2 \psi}{dx^2} + V\psi = E\psi. \quad (3.56)$$

The wavefunction in regions I and III are solutions of the free Schrödinger equation and thus

$$E\psi_{I/III} = -\frac{\hbar^2}{2m} \frac{d^2 \psi_{I/III}}{dx^2} \quad (3.57)$$

As we have seen the solutions to this equation will be exponentials or sine/cosine waves depending on the sign of $\hbar^2/2mE$. Since we are interested in normalizable solution, *i.e.* those that decay as $|x| \rightarrow \infty$ we need to have exponential solutions. Thus $E < 0$ and

$$\psi_{I/III} = A_{I/III} e^{-\kappa x} + B_{I/III} e^{\kappa x}, \quad \kappa = \sqrt{\frac{-2mE}{\hbar^2}} \quad (3.58)$$

The conditions $\psi \rightarrow 0$ as $x \rightarrow -\infty$ and $\psi \rightarrow 0$ as $x \rightarrow \infty$ then imply that $A_I = B_{III} = 0$.

For region II we must include the effect of V_0 . This gives

$$(E + V_0)\psi_{II} = -\frac{\hbar^2}{2m} \frac{d^2 \psi_{II}}{dx^2} \quad (3.59)$$

and hence the solution is

$$\psi_{II} = A_{II} e^{-ikx} + B_{II} e^{ikx}, \quad k = \sqrt{\frac{2m(E + V_0)}{\hbar^2}} \quad (3.60)$$

Note that k is real since $E \geq -V_0$.

Next we need to determine the four integration constants $B_I, A_{III}, A_{II}^\pm, B_{II}^\pm$ and determine the energy E . To do this we need to determine the matching condition on the wavefunction at the boundaries $x = \pm l$. Since $|\psi|^2$ is the probability density we must ensure that it is continuous. This gives one matching condition at $x = \pm l$. Next we integrate the time independent Schrödinger equation around a neighbourhood of the boundaries $x = \pm l$

$$-\frac{\hbar^2}{2m} \int_{\pm l-\epsilon}^{\pm l+\epsilon} \frac{d^2\psi}{dx^2} = \int_{\pm l-\epsilon}^{\pm l+\epsilon} (E - V(x))\psi(x) \quad (3.61)$$

The left hand side integrates to

$$-\frac{\hbar^2}{2m} \left(\frac{d\psi}{dx}(\pm l + \epsilon) - \frac{d\psi}{dx}(\pm l - \epsilon) \right) \quad (3.62)$$

Next we note that $V(x)$ is piecewise continuous and hence $(E - V(x))\psi(x)$ is too. Furthermore the integral of a piecewise continuous function is continuous. Thus if we take the limit $\epsilon \rightarrow 0$ it follows that the right hand side of 3.61 vanishes. Hence it follows that the derivative of the wavefunction is continuous. In particular the matching conditions are, in this case,

$$\begin{aligned} \psi_I(-l) &= \psi_{II}(-l) \\ B_I e^{-\kappa l} &= A_{II} e^{ikl} + B_{II} e^{-ikl} \end{aligned} \quad (3.63)$$

$$\begin{aligned} \psi'_I(-l) &= \psi'_{II}(-l) \\ \kappa B_I e^{-\kappa l} &= -ik A_{II} e^{ikl} + ik B_{II} e^{-ikl} \end{aligned} \quad (3.64)$$

$$\begin{aligned} \psi_{III}(l) &= \psi_{II}(l) \\ A_{III} e^{-\kappa l} &= A_{II} e^{-ikl} + B_{II} e^{ikl} \end{aligned} \quad (3.65)$$

$$\begin{aligned} \psi'_{III}(l) &= \psi'_{II}(l) \\ -\kappa A_{III} e^{-\kappa l} &= -ik A_{II} e^{-ikl} + ik B_{II} e^{ikl} \end{aligned} \quad (3.66)$$

These matching conditions, along with the over-all normalization condition, gives five equations for the five unknowns.

To help analyze these solutions we note that $\partial^2/\partial x^2$ and $V(x)$ are even, *i.e.* invariant under $x \leftrightarrow -x$. It follows that a basis of solutions can be split into even and odd wavefunctions. To see this in detail we observe that, due to the invariance of d^2/dx^2 and $V(x)$, the time independent Schrödinger equations for $\psi(x)$ and $\psi(-x)$ are

$$\begin{aligned} E\psi(x) &= -\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) \\ E\psi(-x) &= -\frac{\hbar^2}{2m} \frac{d^2\psi(-x)}{dx^2} + V(x)\psi(-x) \end{aligned} \quad (3.67)$$

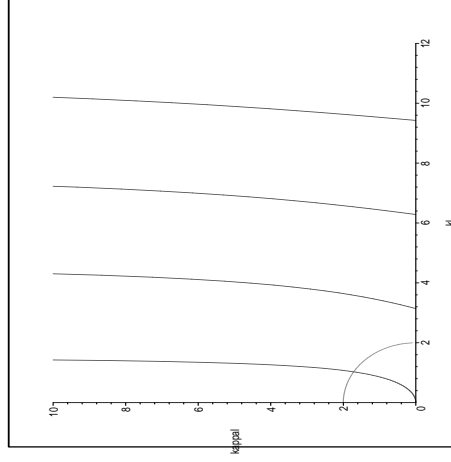


Figure 1:

Therefore taking the sum or the difference of these two equations we learn that $\psi_{\text{even/odd}}(x) = (\psi(x) \pm \psi(-x))/2$ are solutions to the Schrödinger equation.

Let us consider the even wavefunctions first. Here $B_I = A_{III} = C$ and $A_{II} = B_{II} = D$ so that the conditions are

$$\begin{aligned} C e^{-\kappa l} &= D e^{ikl} + D e^{-ikl} \\ \kappa C e^{-\kappa l} &= -ik D e^{ikl} + ik D e^{-ikl} \end{aligned} \quad (3.68)$$

Thus

$$C = D(e^{(ik+\kappa)l} + e^{(-ik+\kappa)l}) = -\frac{ikD}{\kappa}(e^{(ik+\kappa)l} - e^{(-ik+\kappa)l}) \quad (3.69)$$

From which we learn that

$$-i \frac{e^{ikl} - e^{-ikl}}{e^{ikl} + e^{-ikl}} = \tan(kl) = \frac{\kappa}{k} \quad (3.70)$$

It also follows from the definition of k and κ that

$$\kappa^2 l^2 + k^2 l^2 = \frac{2mV_0 l^2}{\hbar^2} \quad (3.71)$$

A plot of $\kappa l = kl \tan(kl)$ and the circle 3.71 (see figure 1) shows that the number of solutions to these two equations is finite but grows with the size of $2mV_0 l^2/\hbar^2$. In particular if $2mV_0 l^2/\hbar^2 < \pi^2/4$ (recall that both κl and kl are positive) then there is a unique solution.

To fix the overall normalization we note that the even wavefunctions have the form

$$\psi_I = 2D e^{\kappa l} \cos(kl) e^{\kappa x} \quad x < -l$$

$$\begin{aligned}
\psi_{II} &= 2D \cos(kx) \\
\psi_{III} &= 2De^{\kappa l} \cos(kl) e^{-\kappa x} \quad x > l
\end{aligned}
\tag{3.72}$$

so that

$$\begin{aligned}
1 &= \int dx |\psi|^2 \\
&= 2 \int_0^l dx |\psi_{II}|^2 + 2 \int_l^\infty |\psi_{III}|^2 \\
&= 8|D|^2 \left(\int_0^l dx \cos^2(kx) + e^{2\kappa l} \cos^2(kl) \int_l^\infty e^{-2\kappa x} \right) \\
&= 8|D|^2 \left(\frac{l}{2} + \frac{1}{4k} \sin(2kl) + \frac{e^{2\kappa l} \cos^2(kl)}{2\kappa} \right)
\end{aligned}
\tag{3.73}$$

Hence

$$D = \sqrt{\frac{1}{8}} \left(\frac{l}{2} + \frac{1}{4k} \sin(2kl) + \frac{e^{2\kappa l} \cos^2(kl)}{2\kappa} \right)^{-\frac{1}{2}}
\tag{3.74}$$

Of course to complete the solution we must also consider the odd-wavefunctions (*see the problem sets*).

These solutions are known as bound states since the wavefunction is localized around the potential minimum. The probability of finding a particle a distance $|x| > 1/\kappa \sim l$ is exponentially small (but non-zero). We see that there are only a finite number of these states which come with discrete negative energies. Note that for $kl = n\pi/2$ with n an odd integer we apparently find that the wavefunction vanishes outside of $|x| \leq l$. However according to 3.70 this implies that κ and hence E are infinite.

3.3 scattering and tunneling through a potential barrier

Next we can consider a similar problem to the potential well but now with

$$V(x) = \begin{cases} V_0 & \text{if } |x| \leq l, \\ 0 & \text{if } |x| > l, \end{cases}
\tag{3.75}$$

where $V_0 > 0$. This is a potential barrier. The analysis of the previous section can be used but with $V_0 \rightarrow -V_0$. It is clear from the condition $\kappa^2 l^2 + k^2 l^2 = -2mV_0 l^2 / \hbar^2$ that there are no solutions which decay at infinity, *i.e.* no bound states with $E < 0$.

Therefore we need to look for other solutions, known as scattering states. Here we give up the notion that the total probability of finding the particle somewhere is one. Instead we imagine firing a continuous beam of particles at the barrier. Far to the left (in region I) we therefore imagine a plane wave solution ($k > 0$)

$$\psi_I = e^{ikx} + R e^{-ikx}
\tag{3.76}$$

Recall (*see the problems sets*) that the probability current for a simple plane wave $\psi_k = e^{\pm ikx}$ is $j_k = \pm \hbar k/m = \pm p_k/m = \pm v$ where $\pm v$ is the “classical velocity” (note that we haven’t, and won’t, define velocity in quantum mechanics). The sign distinguishes between a current of particles moving to the left or right. The first term represents a beam of particles coming in from $x \rightarrow -\infty$. We choose our normalization so that its coefficient is one. The second term represents a beam of particles exiting to the left and is interpreted as the reflected modes. Their density, relative to the incoming modes, is given by $|R|^2$. Similarly in region III we have

$$\psi_{III} = Te^{ikx} \quad (3.77)$$

corresponding transmitted modes going out to the right. Their density relative to the incoming mode is $|T|^2$. Here we have chosen boundary conditions so that there are no modes coming in from $x = \infty$, *i.e.* no e^{-ikx} term. It is clear that since $V = 0$ in these regions, the Schrödinger equation implies that

$$E = \frac{k^2}{2m} \quad (3.78)$$

note that we have changed the definition of k from that used in the potential well.

Finally in region II we have

$$\psi_{II} = Ae^{iKx} + Be^{-iKx} \quad (3.79)$$

and the Schrödinger equation implies that

$$E - V_0 = \frac{K^2}{2m} \quad (3.80)$$

It is important to note that K need not be real since the proof that $E > V_0$, *i.e.* the positivity of \hat{p}^2 , relied on the normalizability of the wavefunction.

Note that we have the same number of variables as in the potential well; four coefficients R, T, A, B and the energy E . But without the overall normalization condition there is one less equation than before. However our task here is slightly different. We do not what to calculate the energy E , instead we consider it as labeling the boundary conditions of our incoming wave. In particular we can choose to take $E < V_0$ so that the particles do not have enough energy (classically) to make it over the barrier. We are primarily interested in the coefficients R and T that determine the amount of reflected and transmitted particles.

To proceed we note again that $\psi(x)$ and its derivative are continuous at the boundaries

$$\begin{aligned} \psi_I(-l) &= \psi_{II}(-l) \\ e^{-ikl} + Re^{ikl} &= Ae^{-iKl} + Be^{iKl} \end{aligned} \quad (3.81)$$

$$\begin{aligned} \psi'_I(-l) &= \psi'_{II}(-l) \\ ike^{-ikl} - ikRe^{ikl} &= iKAe^{-iKl} - iKBe^{iKl} \end{aligned} \quad (3.82)$$

$$\psi_I(III) = \psi_{II}(l) \quad Te^{ikl} = Ae^{iKl} + Be^{-iKl} \quad (3.83)$$

$$\psi'_I(III) = \psi'_{II}(l) \quad ikTe^{ikl} = iKAe^{iKl} - iKBe^{-iKl} \quad (3.84)$$

After some calculations (*see the problem sets*) one can find

$$\begin{aligned} T &= \frac{4Kke^{-2ikl}}{(K+k)^2e^{-2iKl} - (K-k)^2e^{2iKl}} \\ R &= (K^2 - k^2)e^{-2ikl} \frac{e^{2iKl} - e^{-2iKl}}{(K+k)^2e^{-2iKl} - (K-k)^2e^{2iKl}} \end{aligned} \quad (3.85)$$

A bit more calculation shows that (*see the problem sets*), if K is real or pure imaginary,

$$|T|^2 + |R|^2 = 1 \quad (3.86)$$

This can also be seen but considering the probability that a particle is found in the region $|x| < l$,

$$P = \int_{-l}^l dx \rho = \int_{-l}^l dx |\Psi|^2 \quad (3.87)$$

Since we have energy eigenstates, $\Psi = e^{-iEt/\hbar}\psi(x)$ it follows that $|\Psi|^2 = |\psi|^2$ is time independent. Therefore $dP/dt = 0$. On the other hand we can use the probability current to evaluate

$$\begin{aligned} 0 &= \frac{dP}{dt} \\ &= \int_{-l}^l dx \frac{\partial \rho}{\partial t} \\ &= - \int_{-l}^l dx \frac{\partial j}{\partial x} \\ &= j(-l) - j(l) \end{aligned} \quad (3.88)$$

Now since ψ and its derivative are continuous it follows from the formula 2.19 that j is continuous. Therefore we can evaluate it for $|x| > l$ and take the limit to obtain $j(\pm l)$. In region I, where $\psi_I = e^{ikx} + Re^{-ikx}$, we saw in the problem sets that $j_I = \frac{\hbar k}{m}(1 - |R|^2)$. Similarly in region III, $\psi_{III} = Te^{ikx}$ and hence $j_{III} = \frac{\hbar k}{m}|T|^2$. From 3.88 it follows that $1 - |R|^2 - |T|^2 = 0$. Thus the total flux of reflected and transmitted particles is equal to the flux of incoming particles.

It is clear from the expression for R/T that there will be no reflected particles, *i.e.* $R = 0$ if $K = n\pi/2l$ for any integer n . Since $K^2 > 0$ in these cases we have that $E > V_0$. Thus, with the exception of these special values of E , some particles are always reflected. On the other hand there are always transmitted particles.

Note that we haven't needed to assume that $V_0 > 0$ so these scattering states and formulae also apply to the potential well.

Note also that the formulae for T and R are valid even when K is pure imaginary, *i.e.* $E < V_0$. In this case the particle does not have enough energy to pass over the barrier in the classical theory. Here we see that the wavefunction ψ_{II} becomes exponentially suppressed in region II. Thus there is still some small probability to find the particle on the other side of the barrier. Indeed consider the limit $V_0 \rightarrow \infty$ so that $K = iP$ with $P \rightarrow \infty$, here we see that $T \rightarrow 4ike^{-2ikl-2Pl}/P$ becomes exponentially suppressed whereas $R \rightarrow -e^{-2ikl}$.

This process is one of the most interesting quantum phenomena. Even though the particle does not have enough energy to make it over the barrier in the classical theory, $E < V_0$, there is a non-zero probability that it makes it over in the quantum theory. This is called tunneling since the wavefunction becomes exponentially suppressed while passing through the barrier and 'tunnels' across to the other side.

In effect we also saw this in the potential well. The probability to find the particle outside the well was small but non-zero. However the particle had negative energy, whereas a free particle away from the well has positive energy.

One consequence of tunneling is that in a quantum theory, you cannot stay forever in a 'false vacuum', that is a minimum of the potential which is not the global minimum, the 'true' vacuum. The wavefunction will always spread out, leading to a small but non-zero probability of being in the true vacuum. This is a big problem in theories of quantum gravity since it would appear that our universe is not in a global minimum and hence should decay away.

3.4 the Dirac δ -function and planewaves

Several times we have used the wavefunctions $\Psi = e^{ikx-i\omega t}$ even though they are not normalizable in infinite volume. In particular

$$\rho(x, t) = |\Psi|^2 = 1 \quad (3.89)$$

However these functions are so-called δ -function normalizable.

Here the Dirac δ -function $\delta(x)$ has the property that, in one dimension,

$$\int dx \delta(x) f(x) = f(0) \quad (3.90)$$

for any suitable smooth function. Of course in the strict sense $\delta(x)$ is not at all a function. If it were it would need to vanish everywhere except at $x = 0$, where is it so large that its total area is one. In essence whenever you see a δ -function you must interpret it as applying only under an integral.

From the general theory of Fourier analysis we have that

$$f(x) = \int \frac{dk}{2\pi} e^{ikx} \tilde{f}(k), \quad \tilde{f}(k) = \int dx e^{-ikx} f(x) \quad (3.91)$$

So that from the second equation

$$\tilde{\delta}(k) = 1 \quad (3.92)$$

and hence from the first equation, formally,

$$\delta(x) = \int \frac{dk}{2\pi} e^{ikx} . \quad (3.93)$$

To see that this “definition” works consider

$$\begin{aligned} \int dx \delta(x) f(x) &= \int dx \frac{dk}{2\pi} \frac{dp}{2\pi} \tilde{f}(k) e^{ikx} e^{ipx} \\ &= \int \frac{dk}{2\pi} \frac{dp}{2\pi} 2\pi \delta(k+p) \tilde{f}(k) \\ &= \int \frac{dk}{2\pi} \tilde{f}(k) \\ &= f(0) \end{aligned} \quad (3.94)$$

where in the first line we expanded out the functions δ and f in terms of their Fourier modes, in the second line we integrated over x and in the third line we integrated over p .

The extension to three-dimensional space is straightforward with

$$\delta(\vec{x}) = \int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot\vec{x}} \quad (3.95)$$

Thus if we consider two plane waves $\Psi_{\vec{k}} = e^{i\vec{k}\cdot\vec{x} - i\omega_k t}$ and $\Psi_{\vec{p}} = e^{i\vec{p}\cdot\vec{x} - i\omega_p t}$ then

$$\int dx \Psi_{\vec{p}}^* \Psi_{\vec{k}} = \int dx e^{i(\vec{k}-\vec{p})\cdot\vec{x}} e^{i(\omega_p - \omega_k)t} = (2\pi)^3 \delta(\vec{k} - \vec{p}) e^{i(\omega_p - \omega_k)t} \quad (3.96)$$

and this is what we mean by δ -function normalizable. Note that the δ -function imposes $\vec{k} = \vec{p}$ so that $\omega_k = \omega_p$ and hence the exponential can be dropped from the right hand side.

Returning to our plane wave solutions it is possible to construct normalizable solutions from them. For example in the case of the time-independent free Schrödinger equation

$$-\frac{\hbar^2}{2m} \nabla^2 \psi = E \psi \quad (3.97)$$

Solutions are of the form $\psi = e^{i\vec{k}\cdot\vec{x}}$ with

$$E_k = \frac{\hbar^2 k^2}{2m} , \quad (3.98)$$

and the momentum of the waves are

$$\vec{p} = \hbar \vec{k} . \quad (3.99)$$

The general solution is called a “wave packet” and takes the form

$$\Psi(\vec{x}, t) = \int \frac{d^3k}{(2\pi)^3} \phi(\vec{k}) e^{i\vec{k}\cdot\vec{x} - iE_k t/\hbar} \quad (3.100)$$

Here the energy eigenvalues are in fact continuous, rather than discrete as usual in problems with a potential. The superposition is now over the various continuous momenta and the weight is a continuous function of the momentum $\phi(\vec{k})$. From our discussion above it follows that (*see the problem sets*)

$$\int |\Psi(\vec{x}, t)|^2 = \int \frac{d^3k}{(2\pi)^3} |\phi(\vec{k})|^2 \quad (3.101)$$

and hence Ψ can be made normalizable.

4 formal quantum mechanics and some Hilbert space theory

Our next aim is to cast Quantum Mechanics into a more abstract and formal form. Until now we have described the various systems in terms of a complex valued wavefunction, $\Psi(\vec{x}, t)$, whose time evolution is described by the (time-dependent) Schrödinger equation. An important feature was that the Schrödinger equation is a linear equation, and therefore that any multiple of a solution is a solution, and sums of solutions define solutions. This gave us the constructive and destructive inference patterns associated with waves. Furthermore, in order for the wavefunction to have a probabilistic interpretation as describing a single particle, the wavefunction has to be normalized,

$$\int d^3x |\Psi(\vec{x}, t)|^2 = 1 \quad \text{for each } t. \quad (4.1)$$

The various physical quantities were represented by operators on the space of wavefunction. In particular the time evolution was determined by the energy operator through the the Schrödinger equation. We may rewrite this in a more abstract form by introducing an operator corresponding to the classical Hamiltonian:

$$\begin{aligned} i\hbar \frac{\partial \Psi}{\partial t} &= -\frac{\hbar^2}{2m} \nabla^2 \Psi + V\Psi \\ \text{iff} & \\ \hat{E}\Psi &= \hat{H}\Psi \end{aligned} \quad (4.2)$$

where

$$\hat{H}\Psi = \left(\frac{\hat{p}^2}{2m} + \hat{V} \right) \Psi. \quad (4.3)$$

Recall that

$$\hat{E}\Psi = i\hbar \frac{\partial \Psi}{\partial t} \quad \hat{p}\Psi = -i\hbar \hat{\nabla}\Psi \quad \text{and} \quad \hat{x}\Psi = x\Psi \quad (4.4)$$

In this section we want to see how the natural setting for these statements is that of a Hilbert space.

In what follows we will content ourselves with a discussion of quantum mechanics in one spatial dimension x .

4.1 Hilbert spaces and Dirac bracket notation

Let us start by introducing the formal definitions required in Hilbert space theory. Note that the study of Hilbert spaces is a very well defined and highly developed branch of pure mathematics. We will not do justice to this subject here and merely introduce the concepts and theorems that we need at the level of rigour that we need them.

First we recall the notion of a complex vector space.

Definition: A *complex vector space* is a set V with an operation of addition $+$ on V and multiplication by scalars in \mathbb{C} , such that

(i) addition is symmetric and associative, *i.e.*

$$\begin{aligned} \mathbf{a} + \mathbf{b} &= \mathbf{b} + \mathbf{a} & \forall \mathbf{a}, \mathbf{b} \in V \\ (\mathbf{a} + \mathbf{b}) + \mathbf{c} &= \mathbf{a} + (\mathbf{b} + \mathbf{c}) & \forall \mathbf{a}, \mathbf{b}, \mathbf{c} \in V \end{aligned} \quad (4.5)$$

(ii) there exists an element $\mathbf{0} \in V$ such that

$$\mathbf{a} + \mathbf{0} = \mathbf{a} \quad \forall \mathbf{a} \in V \quad (4.6)$$

(iii) For each $\mathbf{a} \in V$, there exists a vector $-\mathbf{a} \in V$ such that

$$\mathbf{a} + (-\mathbf{a}) = \mathbf{0}. \quad (4.7)$$

(iv) Furthermore we have

$$\begin{aligned} \lambda(\mu\mathbf{a}) &= (\lambda\mu)\mathbf{a} & \forall \mathbf{a} \in V, \quad \lambda, \mu \in \mathbb{C} \\ 1\mathbf{a} &= \mathbf{a} & \forall \mathbf{a} \in V \\ \lambda(\mathbf{a} + \mathbf{b}) &= \lambda\mathbf{a} + \lambda\mathbf{b} & \forall \mathbf{a}, \mathbf{b} \in V, \quad \lambda \in \mathbb{C} \\ (\lambda + \mu)\mathbf{a} &= \lambda\mathbf{a} + \mu\mathbf{a} & \forall \mathbf{a} \in V, \quad \lambda, \mu \in \mathbb{C}. \end{aligned}$$

You should convince yourselves that various other properties follow from this definition, for example, $-1\mathbf{a} = -\mathbf{a}$, $0\mathbf{a} = \mathbf{0}$, $-\mathbf{0} = \mathbf{0}$, $\lambda\mathbf{0} = \mathbf{0}$. The elements of V are usually called *vectors*.

Before we move on to Hilbert spaces we need one more definition:

Definition: An inner product on \mathcal{H} is a mapping $\langle | \rangle : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{C}$ such that for all vectors $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathcal{H}$ and scalars $\lambda \in \mathbb{C}$

(i) The inner product is linear in the right entry, *i.e.*

$$\begin{aligned}\langle \mathbf{a} | \mathbf{b} + \mathbf{c} \rangle &= \langle \mathbf{a} | \mathbf{b} \rangle + \langle \mathbf{a} | \mathbf{c} \rangle \\ \langle \mathbf{a} | \lambda \mathbf{b} \rangle &= \lambda \langle \mathbf{a} | \mathbf{b} \rangle.\end{aligned}$$

(ii) We have

$$\langle \mathbf{a} | \mathbf{b} \rangle = \langle \mathbf{b} | \mathbf{a} \rangle^*, \quad (4.8)$$

where $*$ denotes complex conjugation.

It can then be shown that

$$\begin{aligned}\text{(a)} \quad \langle \mathbf{a} | \lambda \mathbf{b} + \mu \mathbf{c} \rangle &= \lambda \langle \mathbf{a} | \mathbf{b} \rangle + \mu \langle \mathbf{a} | \mathbf{c} \rangle \\ \text{(b)} \quad \langle \lambda \mathbf{a} + \mu \mathbf{b} | \mathbf{c} \rangle &= \lambda^* \langle \mathbf{a} | \mathbf{c} \rangle + \mu^* \langle \mathbf{b} | \mathbf{c} \rangle.\end{aligned}$$

This last condition shows that the inner-product is complex anti-linear in the first entry.

Furthermore a *positive definite* inner-product satisfies:

$$\langle \mathbf{a} | \mathbf{a} \rangle \geq 0 \quad \text{with} \quad \langle \mathbf{a} | \mathbf{a} \rangle = 0 \quad \text{if and only if} \quad \mathbf{a} = \mathbf{0}. \quad (4.9)$$

Given a positive definite inner product, we can define the corresponding norm by

$$\|\mathbf{a}\| = \sqrt{\langle \mathbf{a} | \mathbf{a} \rangle}. \quad (4.10)$$

Here the square root denotes the *positive* square root; this is well defined since $\langle \mathbf{a} | \mathbf{a} \rangle \geq 0$. The norm has the property that

$$\|\lambda \mathbf{a}\| = |\lambda| \cdot \|\mathbf{a}\|. \quad (4.11)$$

A Hilbert space can now be defined as follows:

Definition: A *Hilbert space* is a complex vector space \mathcal{H} together with a positive definite inner product on \mathcal{H} .

N.B.: Often an additional requirement of a Hilbert space is that it is complete with respect to its norm 4.10. Recall that completeness is the statement that every Cauchy sequence \mathbf{a}_n , *i.e.* a sequence for which $\lim_{n,m \rightarrow \infty} \|\mathbf{a}_n - \mathbf{a}_m\| = 0$, converges. Since the proof of completeness is beyond the level of detail provided in this course we will simply assume that all the Hilbert spaces we encounter are complete (since it is true).

The archetypel example of a complex vector space is the space \mathbb{C}^n where $n \in \mathbb{N}$. In the simplest (non-trivial) case, where $n = 2$, we can write the vectors as

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad (4.12)$$

where $a_1, a_2 \in \mathbb{C}$. Addition is then defined by

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \end{pmatrix}, \quad (4.13)$$

while scalar multiplication is given by

$$\lambda \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \end{pmatrix}. \quad (4.14)$$

The inner product is defined by

$$\left\langle \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \middle| \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right\rangle = (a_1^* \ a_2^*) \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = a_1^* b_1 + a_2^* b_2. \quad (4.15)$$

To check that this defines indeed a Hilbert space:

First we check that we have a vector space. This is quite standard linear algebra, and we will not give the details here.

Next we check that $\langle | \rangle$ is indeed a positive definite inner product.

(i) Define

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}. \quad (4.16)$$

Then

$$\begin{aligned} \langle \mathbf{a} | \mathbf{b} + \mathbf{c} \rangle &= \left\langle \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \middle| \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \right\rangle \\ &= \left\langle \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \middle| \begin{pmatrix} b_1 + c_1 \\ b_2 + c_2 \end{pmatrix} \right\rangle \\ &= a_1^*(b_1 + c_1) + a_2^*(b_2 + c_2) \\ &= a_1^* b_1 + a_2^* b_2 + a_1^* c_1 + a_2^* c_2, \end{aligned}$$

while

$$\begin{aligned} \langle \mathbf{a} | \mathbf{b} \rangle + \langle \mathbf{a} | \mathbf{c} \rangle &= \left\langle \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \middle| \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right\rangle + \left\langle \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \middle| \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \right\rangle \\ &= a_1^* b_1 + a_2^* b_2 + a_1^* c_1 + a_2^* c_2. \end{aligned}$$

Thus the two expressions agree.

Likewise we can show

$$\begin{aligned} \langle \mathbf{a} | \lambda \mathbf{b} \rangle &= \left\langle \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \middle| \lambda \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right\rangle \\ &= \left\langle \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \middle| \begin{pmatrix} \lambda b_1 \\ \lambda b_2 \end{pmatrix} \right\rangle \\ &= a_1^* \lambda b_1 + a_2^* \lambda b_2 \\ &= \lambda (a_1^* b_1 + a_2^* b_2) \\ &= \lambda \langle \mathbf{a} | \mathbf{b} \rangle. \end{aligned} \quad (4.17)$$

(ii) Finally, we have

$$\begin{aligned}\langle \mathbf{a} | \mathbf{b} \rangle &= \left\langle \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \middle| \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right\rangle \\ &= a_1^* b_1 + a_2^* b_2\end{aligned}$$

while on the other hand

$$\begin{aligned}\langle \mathbf{b} | \mathbf{a} \rangle &= \left\langle \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \middle| \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \right\rangle \\ &= b_1^* a_1 + b_2^* a_2 \\ &= (a_1^* b_1 + a_2^* b_2)^* \\ &= \langle \mathbf{a} | \mathbf{b} \rangle^*.\end{aligned}$$

Finally we must consider whether or not the inner product is positive definite. Let \mathbf{a} be given as in (4.12). Then we have

$$\langle \mathbf{a} | \mathbf{a} \rangle = \left\langle \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \middle| \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \right\rangle = a_1^* a_1 + a_2^* a_2 = |a_1|^2 + |a_2|^2. \quad (4.18)$$

Now $|a| \geq 0$ with $|a| = 0$ if and only if $a = 0$. Thus $\langle \mathbf{a} | \mathbf{a} \rangle \geq 0$; moreover, $\langle \mathbf{a} | \mathbf{a} \rangle = 0$ if and only if both a_1 and a_2 are zero, *i.e.* if and only if $\mathbf{a} = \mathbf{0}$.

Analogous statements also hold for \mathbb{C}^n , the set of complex n -tuples with addition and scalar multiplication defined by

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}. \quad (4.19)$$

With these operations, \mathbb{C}^n is a vector space. If we define the inner product by

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \middle| \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{i=1}^n x_i^* y_i, \quad (4.20)$$

\mathbb{C}^n actually becomes a Hilbert space; this is left as an exercise.

In most uses Hilbert spaces are actually infinite dimensional. However things can get out of control and so we will assume that the the Hilbert spaces in quantum theory are “seperable”

Definition: A Hilbert space is said to be seperable if it has a countable basis *i.e.* every element can be uniquely determined by a sum of the form

$$\mathbf{x} = \sum_{n=1}^{\infty} x_n \mathbf{e}_n \quad (4.21)$$

with $x_n \in \mathbb{C}$ and $\mathbf{x} = \mathbf{y}$ iff $x_n = y_n$ for all n . Of course we include finite dimensional spaces in this class. Furthermore by using the Gramm-Schmidt process we can always find a basis which is orthonormal

$$\langle \mathbf{e}_n | \mathbf{e}_m \rangle = \delta_{mn} \quad (4.22)$$

For every Hilbert space we can construct another Hilbert space from it, known as its dual

Definition: The dual to a Hilbert space \mathcal{H} is denoted by \mathcal{H}^* and is defined to be the space of linear maps from \mathcal{H} to \mathbb{C} .

It should be clear that \mathcal{H}^* is a complex vector space. What we need to establish is the existence of an inner-product. To do this we will actually first prove a stronger result

Theorem: \mathcal{H}^* and \mathcal{H} are isomorphic as vector spaces.

Proof: We need to construct a linear map from \mathcal{H} to \mathcal{H}^* which is one-to-one and onto. For any vector $\mathbf{v} \in \mathcal{H}$ we can consider the map $f_{\mathbf{v}} : \mathcal{H} \rightarrow \mathbb{C}$ defined by

$$f_{\mathbf{v}}(\mathbf{x}) = \langle \mathbf{v} | \mathbf{x} \rangle \quad (4.23)$$

From the properties of the inner-product it is clear that $f_{\mathbf{v}}$ is a linear map from \mathcal{H} to \mathbb{C} and hence defines an element of \mathcal{H}^* .

Note that the map which takes $\mathbf{v} \rightarrow f_{\mathbf{v}}$ is an anti-complex linear map from \mathcal{H} to \mathcal{H}^* , viz,

$$f_{\mu\mathbf{v} + \nu\mathbf{u}}(\mathbf{x}) = \langle \mu\mathbf{v} + \nu\mathbf{u} | \mathbf{x} \rangle = \mu^* \langle \mathbf{v} | \mathbf{x} \rangle + \nu^* \langle \mathbf{u} | \mathbf{x} \rangle = \mu^* f_{\mathbf{v}}(\mathbf{x}) + \nu^* f_{\mathbf{u}}(\mathbf{x}) \quad (4.24)$$

Let us check that it is one-to-one. Suppose that $f_{\mathbf{v}} = f_{\mathbf{u}}$, *i.e.*

$$\langle \mathbf{v} | \mathbf{x} \rangle = \langle \mathbf{u} | \mathbf{x} \rangle \quad (4.25)$$

for all \mathbf{x} . Choosing $\mathbf{x} = \mathbf{v} - \mathbf{u}$ we see that $\langle \mathbf{v} - \mathbf{u} | \mathbf{v} - \mathbf{u} \rangle = 0$ and hence $\mathbf{v} - \mathbf{u} = \mathbf{0}$. Thus our map is indeed one-to-one.

Our last step is to show that it is also onto. To this end consider any element f of \mathcal{H}^* , that is f is a linear map from \mathcal{H} to \mathbb{C} . As is known from linear algebra, such a map is determined uniquely by its action on a basis: $f(\mathbf{e}_n) = f_n$. We can therefore construct the element

$$\mathbf{v} = \sum f_n^* \mathbf{e}_n \in \mathcal{H} . \quad (4.26)$$

Hence if $\mathbf{x} = \sum_n x_n \mathbf{e}_n$ is any vector in \mathcal{H} we have that

$$f_{\mathbf{v}}(\mathbf{x}) = \langle \mathbf{v} | \mathbf{x} \rangle = \sum_n \langle f_n^* \mathbf{e}_n | \mathbf{x} \rangle = \sum_n f_n \langle \mathbf{e}_n | \mathbf{x} \rangle \quad (4.27)$$

whereas

$$f(\mathbf{x}) = f\left(\sum_n x_n \mathbf{e}_n\right) = \sum_n x_n f(\mathbf{e}_n) = \sum_n x_n f_n \quad (4.28)$$

Choosing an orthonormal basis we have that

$$\langle \mathbf{e}_n | \mathbf{x} \rangle = \langle \mathbf{e}_n | \sum_m x_m \mathbf{e}_m \rangle = \sum_m x_m \langle \mathbf{e}_n | \mathbf{e}_m \rangle = x_n \quad (4.29)$$

and hence $f(\mathbf{x}) = f_{\mathbf{v}}(\mathbf{x})$. Thus our map is one-to-one and onto.

It now follows that we can define an inner product on \mathcal{H}^* by simply

$$\langle f_{\mathbf{v}} | f_{\mathbf{u}} \rangle = \langle \mathbf{u} | \mathbf{v} \rangle \quad (4.30)$$

Note the change in order between the left and right hand sides. To see that this is indeed an innerproduct we consider property (i)

$$\langle f_{\mathbf{v}} | f_{\mathbf{u}} + f_{\mathbf{w}} \rangle = \langle f_{\mathbf{v}} | f_{\mathbf{u}+\mathbf{w}} \rangle = \langle \mathbf{u} + \mathbf{w} | \mathbf{v} \rangle = \langle \mathbf{u} | \mathbf{v} \rangle + \langle \mathbf{w} | \mathbf{v} \rangle = \langle f_{\mathbf{v}} | f_{\mathbf{u}} \rangle + \langle f_{\mathbf{v}} | f_{\mathbf{w}} \rangle \quad (4.31)$$

and

$$\langle f_{\mathbf{v}} | \lambda f_{\mathbf{u}} \rangle = \langle f_{\mathbf{v}} | f_{\lambda^* \mathbf{u}} \rangle = \langle \lambda^* \mathbf{u} | \mathbf{v} \rangle = \lambda \langle \mathbf{u} | \mathbf{v} \rangle = \lambda \langle f_{\mathbf{v}} | f_{\mathbf{u}} \rangle \quad (4.32)$$

where we have used 4.24, along with the fact that $\langle \cdot | \cdot \rangle$ is an inner product on \mathcal{H} . We must also check condition (ii)

$$\langle f_{\mathbf{v}} | f_{\mathbf{u}} \rangle^* = \langle \mathbf{u} | \mathbf{v} \rangle^* = \langle \mathbf{v} | \mathbf{u} \rangle = \langle f_{\mathbf{u}} | f_{\mathbf{v}} \rangle \quad (4.33)$$

And finally we see that it is manifestly positive definite since

$$\langle f_{\mathbf{u}} | f_{\mathbf{u}} \rangle = \langle \mathbf{u} | \mathbf{u} \rangle \geq 0 \quad (4.34)$$

with equality iff $\mathbf{u} = \mathbf{0}$, corresponding to $f_{\mathbf{0}} = 0$.

We can now introduce the so-called Dirac notation. We denote vectors in \mathcal{H} by 'kets'

$$|\mathbf{a}\rangle \in \mathcal{H} \quad (4.35)$$

and we can, via our isomorphism, denote vectors in \mathcal{H}^* by 'bra's'

$$f_{\mathbf{b}} = \langle \mathbf{b} | \in \mathcal{H}^* \quad (4.36)$$

so that $f_{\mathbf{b}}(\mathbf{a}) = \langle \mathbf{b} | \mathbf{a} \rangle$ is a bra-c-ket! Note that since there is no ambiguity now we need not use a bold-faced letter to refer to a vector if we use Dirac notation, *i.e.* $|\mathbf{a}\rangle = |a\rangle$.

Suppose now that f is an arbitrary element of \mathcal{H}^* . Then we claim that

$$f = \sum_{n=1}^{\infty} f(|\mathbf{e}_n\rangle) \langle \mathbf{e}_n |. \quad (4.37)$$

Indeed applying this to any given vector $|\mathbf{x}\rangle$ we find

$$f(|\mathbf{x}\rangle) = \sum_n f(|\mathbf{e}_n\rangle) \langle \mathbf{e}_n | \mathbf{x} \rangle = \sum_n f_n \langle \mathbf{e}_n | \mathbf{x} \rangle \quad (4.38)$$

where we again have that $f(|\mathbf{e}_n\rangle) = f_n$, which agrees with 4.27 since $x_n = \langle \mathbf{e}_n | \mathbf{x} \rangle$.

Actually, for the example of \mathbb{C}^2 (or indeed \mathbb{C}^n), we can think of the conjugate vector to be the *Hermitian conjugate* of the original vector; thus if

$$|\mathbf{a}\rangle = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad |\mathbf{b}\rangle = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \quad (4.39)$$

we have

$$f_{\mathbf{a}}(\mathbf{b}) = \langle \mathbf{a} | \mathbf{b} \rangle = a_1^* b_1 + a_2^* b_2 \quad (4.40)$$

Hence if we identify

$$\langle \mathbf{a} | = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}^\dagger = (a_1^* \quad a_2^*) . \quad (4.41)$$

then the inner product is just the usual multiplication of matrices

$$\langle \mathbf{a} | \mathbf{b} \rangle = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}^\dagger \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \quad (4.42)$$

From this point of view, property (ii) is simply the rule of taking Hermitian conjugates of products of matrices:

$$\langle \mathbf{a} | \mathbf{b} \rangle^* = \left[\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}^\dagger \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right]^\dagger = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}^\dagger \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \langle \mathbf{b} | \mathbf{a} \rangle . \quad (4.43)$$

4.2 The example of L^2

Now that we have introduced the necessary notions, we want to show that the space of wavefunctions actually defines a Hilbert space. The Hilbert space in question is usually called the L^2 space. More precisely, we define the set $L^2(\mathbb{R})$ to consist of those functions $f : \mathbb{R} \rightarrow \mathbb{C}$ which are square integrable, *i.e.* which satisfy

$$\int_{-\infty}^{\infty} dx |f(x)|^2 < \infty . \quad (4.44)$$

Any function in this set must fall off to zero (sufficiently fast) as $x \rightarrow \pm\infty$. In particular if for large x , $f(x) \sim x^{-\alpha}$ then

$$\int dx |f(x)|^2 \sim \int dx x^{-2\alpha} \sim \frac{1}{1-2\alpha} x^{1-2\alpha} \quad (4.45)$$

which will finite if $\text{Re}(\alpha) > 1/2$. However it is important to note that the failure of a function to be in L^2 can also arise due to singular behaviour at a point. For example if for small x , $f(x) \sim x^{-\alpha}$ then the same argument shows that $f(x)$ is in L^2 if $\text{Re}(\alpha) < 1/2$.

The set $L^2(\mathbb{R})$ is a complex vector space provided we define

$$\begin{aligned} (f + g)(x) &= f(x) + g(x) \\ (\lambda f)(x) &= \lambda f(x) . \end{aligned}$$

Let us first check that these operations are well defined on $L^2(\mathbb{R})$. For the case of the scalar multiplication this is obvious since

$$\int_{-\infty}^{\infty} dx |(\lambda f)(x)|^2 = |\lambda|^2 \int_{-\infty}^{\infty} dx |f(x)|^2 \quad (4.46)$$

and thus the integral will be finite provided that $f \in L^2(\mathbb{R})$. The addition requires a bit more thought since we find that

$$\begin{aligned} \int_{-\infty}^{\infty} dx |(f+g)(x)|^2 &= \int_{-\infty}^{\infty} dx [f(x)^* + g(x)^*][f(x) + g(x)] \\ &= \int_{-\infty}^{\infty} dx |f(x)|^2 + \int_{-\infty}^{\infty} dx |g(x)|^2 \\ &\quad + \int_{-\infty}^{\infty} dx f(x)^* g(x) + \int_{-\infty}^{\infty} dx g(x)^* f(x). \end{aligned} \quad (4.47)$$

The second but last line is finite since both f and g are in $L^2(\mathbb{R})$, but we need to show that this is the case for both terms in the last line. In order to do so we use a trick that is quite common in analysis.

Before we describe this (standard) argument let us first introduce the inner product (we need to define it later on anyway)

$$\langle g|f \rangle = \int_{-\infty}^{\infty} dx g(x)^* f(x). \quad (4.48)$$

The statement that (4.47) is finite now follows from the assertion that the inner product is a finite number (and thus well-defined). Now in order to prove this statement let λ be any complex number, and consider the integral

$$\int_{-\infty}^{\infty} dx |f(x) + \lambda g(x)|^2 \geq 0. \quad (4.49)$$

This integral is non-negative since we are integrating a non-negative function. Now we expand

$$\begin{aligned} |f(x) + \lambda g(x)|^2 &= (f(x) + \lambda g(x))^*(f(x) + \lambda g(x)) \\ &= |f(x)|^2 + \lambda^* g(x)^* f(x) + \lambda g(x) f(x)^* + |\lambda|^2 |g(x)|^2 \end{aligned}$$

and thus (4.49) becomes

$$\begin{aligned} 0 &\leq \int_{-\infty}^{\infty} dx (|f(x)|^2 + \lambda^* g(x)^* f(x) + \lambda g(x) f(x)^* + |\lambda|^2 |g(x)|^2) \\ &= \langle f|f \rangle + \lambda^* \langle g|f \rangle + \lambda \langle f|g \rangle + |\lambda|^2 \langle g|g \rangle. \end{aligned} \quad (4.50)$$

If $g(x) \equiv 0$, the original statement is trivial since then also $f(x)g(x) \equiv 0$, and thus there is nothing to check. Otherwise it follows from the positivity of the integrand that

$$\langle g|g \rangle = \int_{-\infty}^{\infty} dx |g(x)|^2 > 0. \quad (4.51)$$

Next we observe that the right hand side of (4.50) takes the smallest value provided that

$$\lambda = -\frac{\langle g|f \rangle}{\langle g|g \rangle}. \quad (4.52)$$

In order to check this simply differentiate the right hand side of 4.50 with respect to $\lambda = \lambda_1 + i\lambda_2$ and set it equal to zero to find the extremum. In terms of λ_1 and λ_2 the right hand side is

$$\langle f|f \rangle + (\lambda_1 + i\lambda_2)\langle f|g \rangle + (\lambda_1 - i\lambda_2)\langle g|f \rangle + (\lambda_1^2 + \lambda_2^2)\langle g|g \rangle \quad (4.53)$$

Differentiation with respect to λ_1 and λ_2 gives

$$\begin{aligned} 2\lambda_1\langle g|g \rangle + \langle f|g \rangle + \langle g|f \rangle &= 0 \\ 2\lambda_2\langle g|g \rangle + i\langle f|g \rangle - i\langle g|f \rangle &= 0 \end{aligned} \quad (4.54)$$

respectively. The second equation is equivalent to

$$2i\lambda_2\langle g|g \rangle - \langle f|g \rangle + \langle g|f \rangle = 0 \quad (4.55)$$

Thus we indeed see that

$$\lambda = \lambda_1 + i\lambda_2 = -\frac{\langle g|f \rangle}{\langle g|g \rangle}, \quad \lambda^* = \lambda_1 - i\lambda_2 = -\frac{\langle f|g \rangle}{\langle g|g \rangle} \quad (4.56)$$

At any rate, we can choose this value for λ (since λ is arbitrary), and (4.50) therefore becomes

$$\langle f|f \rangle - \frac{\langle g|f \rangle^* \langle g|f \rangle}{\langle g|g \rangle} - \frac{\langle f|g \rangle \langle g|f \rangle}{\langle g|g \rangle} + \frac{|\langle g|f \rangle|^2}{\langle g|g \rangle} \geq 0. \quad (4.57)$$

Observing that $\langle f|g \rangle \langle g|f \rangle = |\langle g|f \rangle|^2$, it then follows (by multiplying through with $\langle g|g \rangle$),

$$|\langle g|f \rangle|^2 \leq \langle g|g \rangle \langle f|f \rangle. \quad (4.58)$$

Definition: The inequality 4.58 is called the *Cauchy-Schwarz inequality*.

N.B.: If you look back at the steps you will see that the Cauchy-Schwarz inequality can be derived for any Hilbert space (*see the problems sets*).

Returning to our proof we see that if f and g are in $L^2(\mathbb{R})$, then by definition $\langle f|f \rangle$ and $\langle g|g \rangle$ are finite, and thus the right hand side of (4.58) is finite, thus establishing our claim.

Thus we have now shown that $L^2(\mathbb{R})$ is a vector space, and that the inner product defined by (4.48) is well-defined. It thus only remains to show that the inner product has the necessary properties.

(i) The linearity of the inner product follows from

$$\begin{aligned}
\langle f|g+h\rangle &= \int_{-\infty}^{\infty} dx f(x)^*(g(x)+h(x)) \\
&= \int_{-\infty}^{\infty} dx f(x)^*g(x) + \int_{-\infty}^{\infty} dx f(x)^*h(x) \\
&= \langle f|g\rangle + \langle f|h\rangle,
\end{aligned} \tag{4.59}$$

and

$$\langle f|\lambda g\rangle = \int_{-\infty}^{\infty} dx f(x)^*\lambda g(x) = \lambda \int_{-\infty}^{\infty} dx f(x)^*g(x) = \lambda \langle f|g\rangle. \tag{4.60}$$

(ii) It follows directly from the definition that

$$\langle g|f\rangle^* = \int_{-\infty}^{\infty} dx g(x)f(x)^* = \langle f|g\rangle. \tag{4.61}$$

Lastly we wish to show that the inner product is positive definite. We observe that

$$\langle f|f\rangle = \int_{-\infty}^{\infty} dx f(x)^*f(x) = \int_{-\infty}^{\infty} dx |f(x)|^2 \geq 0 \tag{4.62}$$

since $|f(x)|^2 \geq 0$ for each x . Furthermore the integral is only zero when $f = 0$ (actually this requires some technical assumptions which can be imposed upon the functions in $L^2(\mathbb{R})$).

Thus we have shown that $L^2(\mathbb{R})$ is a Hilbert space. Its dimension is infinite; this simply means that it is not spanned by any finite set of vectors. (There are infinitely many linearly independent L^2 functions!) However it is separable, for example a basis can be provided by monomials, multiplied by a factor of $e^{-x^2/2a^2}$ to ensure that their norm is finite, *i.e.* $\mathbf{e}_n(x) = x^n e^{-x^2/2a^2}$ (*see the problem sets*).

You should familiarise yourself with the idea that this vector space is a vector space of functions. Very roughly speaking you can think of $x \in \mathbb{R}$ labelling the different components; so instead of having in \mathbb{C}^n v_i with $i = 1, \dots, n$, we now have $f(x)$ where $x \in \mathbb{R}$. However, you should not take this analogy too seriously (since it is only an analogy!)

As we have defined before, given a positive definite inner product we can always defined a norm by

$$\|f\| \equiv \sqrt{\langle f|f\rangle}. \tag{4.63}$$

The Cauchy-Schwarz inequality (4.58) above then simply means that

$$|\langle f|g\rangle| \leq \|f\| \cdot \|g\|. \tag{4.64}$$

The actual wave-functions $\Psi(x, t)$ are elements in the $L^2(\mathbb{R})$ space (for every t) that are normalised; in the current formalism this is simply the statement that $\|\Psi(\cdot, t)\| = 1$ for all t .

4.3 Observables, eigenvalues and expectation values

In the first part of these lectures we encountered a number of so-called ‘observables’, in particular the momentum operator

$$\hat{p} = -i\hbar \frac{\partial}{\partial x} \quad (4.65)$$

the energy operator

$$\hat{E} = i\hbar \frac{\partial}{\partial t} \quad (4.66)$$

and the Hamiltonian operator

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x). \quad (4.67)$$

These operators act on the space of wavefunctions. Thus they define a map

$$\hat{p} : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R}) \quad f(x) \mapsto \hat{p} f(x) \equiv -i\hbar \frac{\partial f(x)}{\partial x}, \quad (4.68)$$

and similarly for \hat{H} . They are actually linear operators. This is to say,

$$\hat{p}(\lambda f + \mu g) = \lambda \hat{p} f + \mu \hat{p} g, \quad (4.69)$$

where $\lambda, \mu \in \mathbb{C}$ and $f, g \in L^2(\mathbb{R})$.

Actually, strictly speaking, this is not true. Functions in $L^2(\mathbb{R})$ need not be differentiable and so \hat{p} isn’t defined over the whole space but rather some subset called its domain. In addition it is not necessarily the case that if a function in $L^2(\mathbb{R})$ is differentiable that its derivative is also in $L^2(\mathbb{R})$. Thus the space of functions obtained by acting on $L^2(\mathbb{R})$ with \hat{p} , called its range, is different than $L^2(\mathbb{R})$. We will ignore here any such subtleties regarding domain and range. You can learn more about it elsewhere in the department. The important point is that \hat{p} can be defined on a dense subset of $L^2(\mathbb{R})$ so that it maps into $L^2(\mathbb{R})$.

Linear operators are the analogues of matrices acting on \mathbb{C}^n and play an important role in quantum mechanics. Let us study a few of their basic features.

Recall that for each Hilbert space \mathcal{H} we constructed another Hilbert space, \mathcal{H}^* called its dual. The same ideas allow us to construct from an arbitrary linear operator \hat{A} another linear operator known as its Hermitian conjugate (or adjoint) \hat{A}^\dagger . We start with the following observation: $\langle \mathbf{x} | \hat{A} \mathbf{y} \rangle$ defines a map $f_{\mathbf{x}} : \mathcal{H} \rightarrow \mathbb{C}$ by

$$f_{\mathbf{x}}(\mathbf{y}) = \langle \mathbf{x} | \hat{A} \mathbf{y} \rangle \quad (4.70)$$

Clearly $f_{\mathbf{x}}$ is a linear mapping and hence is in \mathcal{H}^* . Thus there must be a vector \mathbf{z} such that

$$f_{\mathbf{x}}(\mathbf{y}) = \langle \mathbf{z} | \mathbf{y} \rangle \quad (4.71)$$

for all \mathbf{y} . In other words, for each vector \mathbf{x} there is a vector \mathbf{z} such that

$$\langle \mathbf{x} | \hat{A} \mathbf{y} \rangle = \langle \mathbf{z} | \mathbf{y} \rangle \quad (4.72)$$

for all \mathbf{y} . Thus we have a map $\hat{A}^\dagger : \mathcal{H} \rightarrow \mathcal{H}$ which takes \mathbf{x} to \mathbf{z} so that

$$\langle \mathbf{x} | \hat{A} \mathbf{y} \rangle = \langle \hat{A}^\dagger(\mathbf{x}) | \mathbf{y} \rangle \quad (4.73)$$

for all \mathbf{y} . Furthermore \hat{A}^\dagger is linear

$$\begin{aligned} \langle \hat{A}^\dagger(\mathbf{x}_1 + \lambda \mathbf{x}_2) | \mathbf{y} \rangle &= \langle \mathbf{x}_1 + \lambda \mathbf{x}_2 | \hat{A} \mathbf{y} \rangle \\ &= \langle \mathbf{x}_1 | \hat{A} \mathbf{y} \rangle + \lambda^* \langle \mathbf{x}_2 | \hat{A} \mathbf{y} \rangle \\ &= \langle \hat{A}^\dagger(\mathbf{x}_1) | \mathbf{y} \rangle + \lambda^* \langle \hat{A}^\dagger(\mathbf{x}_2) | \mathbf{y} \rangle \\ &= \langle \hat{A}^\dagger(\mathbf{x}_1) + \lambda \hat{A}^\dagger(\mathbf{x}_2) | \mathbf{y} \rangle \end{aligned} \quad (4.74)$$

Therefore we find

$$\langle \hat{A}^\dagger(\mathbf{x}_1 + \lambda \mathbf{x}_2) - \hat{A}^\dagger(\mathbf{x}_1) - \lambda \hat{A}^\dagger(\mathbf{x}_2) | \mathbf{y} \rangle = 0 \quad (4.75)$$

for all vectors \mathbf{x}_1 , \mathbf{x}_2 and \mathbf{y} . Thus if we take $\mathbf{y} = \hat{A}^\dagger(\mathbf{x}_1 + \lambda \mathbf{x}_2) - \hat{A}^\dagger(\mathbf{x}_1) - \lambda \hat{A}^\dagger(\mathbf{x}_2)$ we see that $\langle \mathbf{y} | \mathbf{y} \rangle = 0$ which means that $\mathbf{y} = \mathbf{0}$ and hence

$$\hat{A}^\dagger(\mathbf{x}_1 + \lambda \mathbf{x}_2) = \hat{A}^\dagger(\mathbf{x}_1) + \lambda \hat{A}^\dagger(\mathbf{x}_2) \quad (4.76)$$

Thus we have the following definition:

Definition: The Hermitian conjugate or adjoint of a linear operator $\hat{A} : \mathcal{H} \rightarrow \mathcal{H}$ is the linear operator $\hat{A}^\dagger : \mathcal{H} \rightarrow \mathcal{H}$ such that

$$\langle \hat{A}^\dagger \mathbf{x} | \mathbf{y} \rangle = \langle \mathbf{x} | \hat{A} \mathbf{y} \rangle \quad (4.77)$$

for all vectors \mathbf{x} and \mathbf{y}

Next we may ask what is $(\hat{A}^\dagger)^\dagger$? From the definition and properties of the inner product we see that

$$\langle (\hat{A}^\dagger)^\dagger \mathbf{x} | \mathbf{y} \rangle = \langle \mathbf{x} | \hat{A}^\dagger \mathbf{y} \rangle = \langle \hat{A}^\dagger \mathbf{y} | \mathbf{x} \rangle^* = \langle \mathbf{y} | \hat{A} \mathbf{x} \rangle^* = \langle \hat{A} \mathbf{x} | \mathbf{y} \rangle \quad (4.78)$$

This shows that $\langle ((\hat{A}^\dagger)^\dagger - \hat{A}) \mathbf{x} | \mathbf{y} \rangle = 0$ for all vectors \mathbf{x} and \mathbf{y} . Thus we see that, choosing $\mathbf{y} = ((\hat{A}^\dagger)^\dagger - \hat{A}) \mathbf{x}$, $\langle \mathbf{y} | \mathbf{y} \rangle = 0$ which means that $\mathbf{y} = \mathbf{0}$ and hence

$$(\hat{A}^\dagger)^\dagger = \hat{A} \quad (4.79)$$

Finally we can consider the product of two operators $(\hat{A}\hat{B})^\dagger$. Here we see that

$$\langle (\hat{A}\hat{B})^\dagger \mathbf{x} | \mathbf{y} \rangle = \langle \mathbf{x} | \hat{A}\hat{B} \mathbf{y} \rangle = \langle \hat{A}^\dagger \mathbf{x} | \hat{B} \mathbf{y} \rangle = \langle \hat{B}^\dagger \hat{A}^\dagger \mathbf{x} | \mathbf{y} \rangle \quad (4.80)$$

for all \mathbf{x} and \mathbf{y} . Thus choosing $\mathbf{y} = ((\hat{A}\hat{B})^\dagger - \hat{B}^\dagger\hat{A}^\dagger)\mathbf{x}$ we see that $\mathbf{y} = \mathbf{0}$ and hence

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger \quad (4.81)$$

The operators that correspond to (measurable) observables are actually quite special operators. As we have argued before, the expectation value of measuring the operator \hat{A} is given by

$$\langle \hat{A}(t) \rangle = \int_{-\infty}^{\infty} dx \Psi(x, t)^* \hat{A} \Psi(x, t). \quad (4.82)$$

Given our definition of the inner product for the $L^2(\mathbb{R})$ space, we can now write this more formally as

$$\langle f | \hat{A} f \rangle. \quad (4.83)$$

This has the interpretation as the average result of measuring \hat{A} . In particular, this must therefore be a real number. In order to ensure this reality property, we require that the observables correspond to *Hermitian* (or *self-adjoint*) operators. For a general Hilbert space we have the following definition

Definition: A Hermitian or self-adjoint operator satisfies $\hat{A}^\dagger = \hat{A}$, that is

$$\langle \mathbf{x} | \hat{A} \mathbf{y} \rangle = \langle \hat{A} \mathbf{x} | \mathbf{y} \rangle. \quad (4.84)$$

for all vectors \mathbf{x} and \mathbf{y} .

Thus if \hat{A} is Hermitian, *i.e.* $\hat{A} = \hat{A}^\dagger$, then we have

$$\langle \mathbf{x} | \hat{A} \mathbf{x} \rangle = \langle \hat{A}^\dagger \mathbf{x} | \mathbf{x} \rangle = \langle \hat{A} \mathbf{x} | \mathbf{x} \rangle = \langle \mathbf{x} | \hat{A} \mathbf{x} \rangle^*. \quad (4.85)$$

i.e. the expectation values of an Hermitian operator are real. Applying this to the case of $L^2(\mathbb{R})$ we see that

$$\langle f | \hat{A} f \rangle. \quad (4.86)$$

is real for Hermitian operators \hat{A} . It is not difficult to check that (4.84) is actually satisfied for the momentum operator \hat{p} . The same is also true for the position, energy and Hamilton operators. These are all called observables.

Definition: An observable is a linear Hermitian operator on the Hilbert space of wavefunctions.

In the finite-dimensional case where the Hilbert space is \mathbb{C}^n , a Hermitian operator is simply a *Hermitian matrix*, *i.e.* a matrix satisfying $\hat{A}^\dagger = \hat{A}$. Here \dagger means that we take the complex conjugate of the transpose. Let us illustrate this for the case of \mathbb{C}^2 a 2×2 matrix. Suppose \hat{A} is the matrix

$$\hat{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad (4.87)$$

then for any vectors $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ in \mathbb{C}^2

$$\begin{aligned}
\langle \mathbf{x} | \hat{A} \mathbf{y} \rangle &= (x_1^* \ x_2^*) \begin{pmatrix} ay_1 + by_2 \\ cy_1 + dy_2 \end{pmatrix} \\
&= x_1^*(ay_1 + by_2) + x_2^*(cy_1 + dy_2) \\
&= (a^*x_1 + c^*x_2)^*y_1 + (b^*x_1 + d^*x_2)^*y_2 \\
&= (a^*x_1 + c^*x_2 \ b^*x_1 + d^*x_2) \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\
&= \langle \hat{A}^\dagger \mathbf{x} | \mathbf{y} \rangle
\end{aligned} \tag{4.88}$$

where

$$\hat{A}^\dagger = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix}. \tag{4.89}$$

Thus this definition of the Hermitian conjugate agrees with the more familiar one in matrix theory. Furthermore it is clear in this definition that

$$(\hat{A}^\dagger)^\dagger = \hat{A} \tag{4.90}$$

as expected.

In the finite-dimensional case (*i.e.* for $\mathcal{H} = \mathbb{C}^n$) any Hermitian matrix \hat{A} has a basis consisting entirely of eigenvectors of \hat{A} . Recall that an eigenvector of a matrix \hat{A} is a vector \mathbf{x} satisfying

$$\hat{A} \mathbf{x} = \lambda \mathbf{x}, \tag{4.91}$$

where $\lambda \in \mathbb{C}$. We then call λ the eigenvalue of \mathbf{x} . This is (presumably) something you have learned in Linear Algebra before; we shall not prove this statement here, we will content ourselves to illustrate it with an example. (This will also remind you of how to find the eigenvectors and eigenvalues.) Let us thus consider the Hermitian matrix

$$\hat{C} = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}. \tag{4.92}$$

To find the eigenvalues of \hat{C} we solve

$$\det \begin{pmatrix} 0 - \lambda & i \\ -i & 0 - \lambda \end{pmatrix} = 0, \tag{4.93}$$

i.e. $\lambda^2 - 1 = 0$. Thus has the two solutions $\lambda = \pm 1$. In order to find the eigenvector corresponding to $\lambda = 1$ we then look for the kernel of the matrix $\hat{C} - \hat{1}$, *i.e.* we are looking for a vector such that

$$\begin{pmatrix} -1 & i \\ -i & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0, \tag{4.94}$$

where $\mathbf{x} \neq \mathbf{0}$. It is easy to see that the solution is of the form

$$\mathbf{e}_1 = a \begin{pmatrix} 1 \\ -i \end{pmatrix}, \quad (4.95)$$

where a is an arbitrary (complex) constant. We can choose \mathbf{e}_1 to have norm one, and up to an arbitrary phase we then have

$$\mathbf{e}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}. \quad (4.96)$$

Similarly we find

$$\mathbf{e}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}. \quad (4.97)$$

Since we have found two eigenvectors of \hat{C} , we can expand every vector in \mathbb{C}^2 in terms of \mathbf{e}_1 and \mathbf{e}_2 . Indeed, it is not difficult to check that

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{\sqrt{2}}(x_1 - ix_2) \mathbf{e}_1 + \frac{1}{\sqrt{2}}(x_1 + ix_2) \mathbf{e}_2. \quad (4.98)$$

The two eigenvectors are orthogonal: this is to say,

$$\begin{aligned} \langle \mathbf{e}_1 | \mathbf{e}_2 \rangle &= \frac{1}{2} \begin{pmatrix} 1 \\ -i \end{pmatrix}^\dagger \cdot \begin{pmatrix} 1 \\ i \end{pmatrix} \\ &= \frac{1}{2} (1 \quad i) \cdot \begin{pmatrix} 1 \\ i \end{pmatrix} \\ &= \frac{1}{2} (1 - 1) = 0. \end{aligned} \quad (4.99)$$

This is actually true more generally:

Proposition Suppose \hat{A} is a Hermitian operator in a Hilbert space \mathcal{H} . Then all eigenvalues of \hat{A} are real, and eigenvectors corresponding to different eigenvalues are orthogonal.

Proof: Let \mathbf{e}_1 and \mathbf{e}_2 be two normalised eigenvectors of \hat{A} , *i.e.*

$$\hat{A}\mathbf{e}_i = \lambda_i \mathbf{e}_i \quad \langle \mathbf{e}_i | \mathbf{e}_i \rangle = 1. \quad (4.100)$$

Then we calculate

$$\begin{aligned} \lambda_i \langle \mathbf{e}_j | \mathbf{e}_i \rangle &= \langle \mathbf{e}_j | \hat{A} \mathbf{e}_i \rangle \\ &= \langle \hat{A} \mathbf{e}_j | \mathbf{e}_i \rangle \\ &= \langle \lambda_j \mathbf{e}_j | \mathbf{e}_i \rangle \\ &= \lambda_j^* \langle \mathbf{e}_j | \mathbf{e}_i \rangle, \end{aligned} \quad (4.101)$$

where we have used the Hermiticity of \hat{A} in the second line, and the properties of the inner product in the final line. First consider now the case that $i = j$. Then the above equation becomes

$$\lambda_i = \lambda_i \langle \mathbf{e}_i | \mathbf{e}_i \rangle = \lambda_i^* \langle \mathbf{e}_j | \mathbf{e}_i \rangle = \lambda_i^*, \quad (4.102)$$

thus showing that all eigenvalues are real. We can therefore replace λ_j^* by λ_j in (4.101), and thus we find

$$(\lambda_i - \lambda_j) \langle \mathbf{e}_j | \mathbf{e}_i \rangle = 0. \quad (4.103)$$

Hence eigenvectors corresponding to different eigenvalues are orthogonal.

N.B.: If there are several eigenvectors with the same eigenvalue then they need not be orthonormal. However together they span a subvector space of \mathcal{H} , all which are eigenvectors with the same eigenvalue. One can then apply the Gram-Schmidt process to this subvector space and find an orthonormal basis.

Some of the above analysis was phrased in terms of the finite dimensional vector space \mathbb{C}^n . In particular, this was the case for the assertion that we can find a basis for \mathbb{C}^n that consists only of eigenvectors of a Hermitian operator. The corresponding statement is still true in spirit in a general Hilbert space, but this is somewhat more subtle and requires more care.

We first note that if $\hat{A} : \mathcal{H} \rightarrow \mathcal{H}$ can be written as, for any operator and orthonormal basis $|\mathbf{e}_n\rangle$,

$$\hat{A} = \sum_{m,n} A_{mn} |\mathbf{e}_n\rangle \langle \mathbf{e}_m| \quad (4.104)$$

where

$$\hat{A} |\mathbf{e}_m\rangle = \sum_n A_{mn} |\mathbf{e}_n\rangle \quad (4.105)$$

To see this we recall that an operator is defined by its action on a basis. Thus if

$$|\mathbf{x}\rangle = \sum_n x_n |\mathbf{e}_n\rangle \quad (4.106)$$

i.e. $\langle \mathbf{e}_m | \mathbf{x} \rangle = x_m$ then

$$\sum_{m,n} A_{mn} |\mathbf{e}_n\rangle \langle \mathbf{e}_m | \mathbf{x} \rangle = \sum_{m,n} x_m A_{mn} |\mathbf{e}_n\rangle = \sum_m x_m \hat{A} |\mathbf{e}_m\rangle = \hat{A} \left(\sum_m x_m |\mathbf{e}_m\rangle \right) = \hat{A} |\mathbf{x}\rangle \quad (4.107)$$

In particular if \hat{A} is Hermitian and has an basis of eigenvectors $|\mathbf{e}_n\rangle$ with

$$\hat{A} |\mathbf{e}_n\rangle = A_n |\mathbf{e}_n\rangle, \quad \langle \mathbf{e}_m | \mathbf{e}_n \rangle = \delta_{mn} \quad (4.108)$$

Therefore $A_{mn} = A_m \delta_{mn}$ and hence

$$\hat{A} = \sum_n A_n |\mathbf{e}_n\rangle \langle \mathbf{e}_n| \quad (4.109)$$

Definition The set of eigenvalues of \hat{A} is called its spectrum.

More generally \hat{A} may have a continuous part of its spectrum in which case there is the important theorem (somewhat simplified):

Spectral Theorem (Lite): If \hat{A} is Hermitian then it can be written as

$$\hat{A} = \sum_n \lambda_n |\mathbf{e}_n\rangle \langle \mathbf{e}_n| + \int d\lambda \lambda |\lambda\rangle \langle \lambda| \quad (4.110)$$

where the continuous eigenstates $|\lambda\rangle$ are δ -function normalisable,

$$\langle \lambda | \lambda' \rangle = \delta(\lambda - \lambda') \quad (4.111)$$

Note that the sum can be finite and the continuous part of the spectrum is not necessarily the whole of \mathbb{R} . This result lies at the heart of quantum mechanics. We will not be able to give a complete proof of it here. One subtlety is the fact that the continuous eigenstates are not really elements of a Hilbert space due to their infinite norm.

One example is the momentum operator \hat{p} which is Hermitian (*see the problem sheets*). We saw that \hat{p} had plane-waves for its eigenfunctions and that these were δ -function normalisable.

Another important example is the Hamiltonian, which is also Hermitian (*see the problem sheets*). We saw that for a particle in a box the Hamiltonian had a set of eigenvectors which were sine-waves and a discrete set of eigenvalues given by the energies. From Fourier theory we know that this sine-waves in fact form a basis for all square-integrable functions on the interval which vanish on the boundary. Thus in this case there were no continuous eigenvalues in the spectrum, which just consists of discrete positive energy eigenvalues

$$E_n = \frac{n^2 \pi^2 \hbar^2}{2ml^2} \quad (4.112)$$

for integer n . On the other hand we also constructed eigenvectors of the Hamiltonian for the case of a particle in a potential well. There we found only a finite number of normalisable modes which had discrete and negative energy eigenvalues. However we also found non-normalisable scattering eigenstates (these are essentially δ -function normalisable) with any positive energy eigenvalue $E > 0$. Thus in this case the Hamiltonian had both a discrete and continuous part of its spectrum.

At any rate, in what follows we shall always assume that the operators corresponding to observables are such that any element of the Hilbert space \mathcal{H} (*i.e.* the space of wavefunctions) can be expanded in terms of eigenfunctions of an observable \hat{A} , *i.e.*

$$|\psi\rangle \in \mathcal{H} : \quad |\psi\rangle = \sum_n c_n |\mathbf{e}_n\rangle, \quad (4.113)$$

where $c_n \in \mathbb{C}$ and $|\mathbf{e}_n\rangle$ are the normalised eigenvectors of the observable \hat{A} with eigenvalues

$$\hat{A} |\mathbf{e}_n\rangle = \lambda_n |\mathbf{e}_n\rangle. \quad (4.114)$$

Typically if the spectrum contains continuous eigenvalues then these can be made discrete by placing the system in a box, that is restricting x into a finite but large range.

Suppose $|\psi\rangle \in \mathcal{H}$ has been written as in (4.113). We can then determine the expectation value of the observable \hat{A} in the state described by $|\psi\rangle$. This is simply

$$\begin{aligned}\langle \hat{A} \rangle &= \langle \psi | \hat{A} \psi \rangle \\ &= \sum_{n,m} c_n^* c_m \langle \mathbf{e}_n | \hat{A} \mathbf{e}_m \rangle \\ &= \sum_{n,m} A_m c_n^* c_m \langle \mathbf{e}_n | \mathbf{e}_m \rangle.\end{aligned}\tag{4.115}$$

If we assume in addition that the different eigenvalues are distinct, *i.e.* $A_m \neq A_n$ if $m \neq n$, then it follows from the above proposition that $\langle \mathbf{e}_n | \mathbf{e}_m \rangle = \delta_{m,n}$, and thus

$$\langle \hat{A} \rangle = \sum_m A_m |c_m|^2.\tag{4.116}$$

Furthermore, we have

$$c_n = \langle \mathbf{e}_n | \psi \rangle.\tag{4.117}$$

As before for the case when \hat{A} was the Hamilton operator we now interpret (4.116) as the statement that the possible results of a measurement of \hat{A} are the eigenvalues A_m , and that the probability with which A_m is measured is precisely $|c_m|^2$. In order for this interpretation to make sense we have to have that

$$\sum_m |c_m|^2 = 1.\tag{4.118}$$

This follows simply from the fact that ψ is normalised since

$$\begin{aligned}1 = \langle \psi | \psi \rangle &= \sum_{m,n} c_m^* c_n \langle \mathbf{e}_m | \mathbf{e}_n \rangle \\ &= \sum_{m,n} c_m^* c_n \delta_{m,n} \\ &= \sum_m |c_m|^2.\end{aligned}\tag{4.119}$$

If some of the eigenvalues appear with a non-trivial multiplicity, *i.e.* if $A_m \neq A_n$ for $m \neq n$ is not correct, then the different eigenvectors are not necessarily orthonormal, but we can always *choose* a basis of orthonormal eigenvectors. This is familiar from \mathbb{C}^n but actually holds in the general case. If the $|\mathbf{e}_n\rangle$ form an orthonormal basis of eigenvectors of the operator \hat{A} , then the probability of measuring a is equal to

$$P(a) = \sum_{m:A_m=a} |c_m|^2.\tag{4.120}$$

After a measurement has been performed and the value $A = a$ has been measured, the wavefunction collapses to an eigenfunction with eigenvalue $A = a$, *i.e.*

$$|\psi\rangle = \sum_{m:A_m=a} d_m |\mathbf{e}_m\rangle,\tag{4.121}$$

where

$$\sum_{m:A_m=a} |d_m|^2 = 1. \quad (4.122)$$

The most common example of this is the solution of the full time-dependent Schrödinger equation in terms of the time independent Schrödinger equation. We saw earlier that typically if we use separation of variables

$$\Psi(x, t) = e^{-iEt/\hbar} \psi(x) \quad (4.123)$$

then we find the equation

$$E\psi = -\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + V(x)\psi \quad (4.124)$$

The theory of differential equations tells us that we will typically find a discrete family of solutions ψ_n with energies E_n .

Thus we are really solving the eigenvalue equation

$$\hat{H}\psi = E\psi \quad (4.125)$$

The general solution to the Schrödinger equation was then a superposition

$$\Psi(x, t) = \sum_n c_n e^{-iE_n t/\hbar} \psi_n(x) \quad (4.126)$$

and this in turn is, for each t , a general element in $L^2(\mathbb{R})$. Our discussion above tells us that $|c_n|^2 = |\langle \Psi | \psi_n \rangle|^2$ is the probability that the wave function is measured to have energy E_n and the average or expected energy of the system after several measurements is

$$\langle \hat{H} \rangle = \langle \Psi | \hat{H} \Psi \rangle = \sum_n E_n |c_n|^2 \quad (4.127)$$

4.4 unitarity and time evolution

You may recall from the theory of matrices that a unitary operator \hat{U} is an operator that preserves the inner product

$$\langle \hat{U}\psi | \hat{U}\phi \rangle = \langle \psi | \phi \rangle \quad (4.128)$$

for all $\psi, \phi \in \mathcal{H}$. Using the definition of the Hermitian conjugate, we can rewrite this as

$$\langle \psi | \hat{U}^\dagger \hat{U} \phi \rangle = \langle \psi | \phi \rangle, \quad (4.129)$$

and therefore conclude that

$$\hat{U}^\dagger \hat{U} = \hat{1}. \quad (4.130)$$

In the finite dimensional context, *i.e.* if $\mathcal{H} = \mathbb{C}^n$, this then also implies that

$$\hat{U} \hat{U}^\dagger = \hat{1} \quad \textit{i.e. that} \quad \hat{U}^{-1} = \hat{U}^\dagger. \quad (4.131)$$

However, in general this latter statement does not hold in the infinite dimensional case (*see the problem sets*).

In the finite dimensional situation, a unitary operator is again simply a unitary matrix. For example, for $\mathcal{H} = \mathbb{C}^2$, the matrix

$$\hat{U} = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \quad (4.132)$$

is unitary. Another class of examples can be obtained from a Hermitian operator:

Proposition: Suppose that \hat{A} is Hermitian operator. Then

$$\hat{U} = e^{i\hat{A}} \quad (4.133)$$

is a unitary operator. Here \hat{U} is defined by the power series expansion, *i.e.*

$$\hat{U} = \sum_{n=0}^{\infty} \frac{i^n}{n!} \hat{A}^n. \quad (4.134)$$

Proof: We calculate

$$\begin{aligned} \langle \hat{U}\psi | \hat{U}\phi \rangle &= \sum_{n=0}^{\infty} \left\langle \frac{i^n}{n!} \hat{A}^n \psi \middle| \hat{U}\phi \right\rangle \\ &= \sum_{n=0}^{\infty} \left\langle \psi \middle| \frac{(-i)^n}{n!} \hat{A}^n \hat{U}\phi \right\rangle \\ &= \langle \psi | e^{-i\hat{A}} e^{i\hat{A}} \phi \rangle \\ &= \langle \psi | \phi \rangle. \end{aligned} \quad (4.135)$$

Here we have used the Hermiticity of \hat{A}^n in the second line.

As an application of this proposition we note that the operator

$$\hat{U}_t = e^{-i\frac{t}{\hbar}\hat{H}} \quad (4.136)$$

is unitary since the Hamilton operator \hat{H} is Hermitian. By construction

$$|\Psi(t)\rangle = \hat{U}_t |\Psi(0)\rangle \quad (4.137)$$

is a solution to the Schrödinger equation provided that \hat{H} is independent of time

$$i\hbar \frac{\partial |\Psi\rangle}{\partial t} = i\hbar \frac{\partial}{\partial t} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{-it}{\hbar}\right)^n \hat{H}^n |\Psi(0)\rangle = \sum_{n=0}^{\infty} \frac{1}{(n-1)!} \left(\frac{-it}{\hbar}\right)^{n-1} \hat{H}^n |\Psi(0)\rangle = \hat{H} |\Psi\rangle \quad (4.138)$$

Clearly also $\Psi(t=0) = \Psi(0)$ since $\hat{U}_0 = \hat{1}$. Thus the above proposition implies that $\Psi(t)$ will be correctly normalised if and only if $\Psi(0)$ is, since

$$\langle \Psi(t) | \Psi(t) \rangle = \langle \hat{U}_t \Psi(0) | \hat{U}_t \Psi(0) \rangle = \langle \Psi(0) | \Psi(0) \rangle. \quad (4.139)$$

Note further that if \hat{A} is a Hermitian operator and \hat{U} is a unitary operator, then $\hat{U}\hat{A}\hat{U}^\dagger$ is again a Hermitian operator. Indeed we calculate

$$\begin{aligned}\langle f|\hat{U}\hat{A}\hat{U}^\dagger g\rangle &= \langle \hat{U}^\dagger f|\hat{A}\hat{U}^\dagger g\rangle \\ &= \langle \hat{A}\hat{U}^\dagger f|\hat{U}^\dagger g\rangle \\ &= \langle \hat{U}\hat{A}\hat{U}^\dagger f|g\rangle.\end{aligned}\tag{4.140}$$

Furthermore, if \hat{A} has eigenvalues A_n with eigenvectors ψ_n , *i.e.*

$$\hat{A}|\psi_n\rangle = A_n|\psi_n\rangle,\tag{4.141}$$

then the eigenvalues of $\hat{U}\hat{A}\hat{U}^\dagger$ are also A_n , where now the eigenvectors are given as

$$|\hat{\psi}_n\rangle = \hat{U}|\psi_n\rangle.\tag{4.142}$$

Indeed, we calculate

$$\begin{aligned}\hat{U}\hat{A}\hat{U}^\dagger|\hat{\psi}_n\rangle &= \hat{U}\hat{A}\hat{U}^\dagger\hat{U}|\psi_n\rangle \\ &= \hat{U}\hat{A}|\psi_n\rangle \\ &= A_n\hat{U}|\psi_n\rangle \\ &= A_n|\hat{\psi}_n\rangle,\end{aligned}\tag{4.143}$$

where we have used (4.130)

5 noncommuting observables and uncertainty

Many quantum phenomena are ultimately a consequence of the fact that observables correspond to Hermitian operators that may not commute. We have already encountered two operators that actually do not commute, namely the position operator \hat{x} and the momentum operator \hat{p} . As we have explained before, \hat{p} acts on $L^2(\mathbb{R})$ as

$$\hat{p} : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R}) \quad f(x) \mapsto \hat{p}f(x) \equiv -i\hbar \frac{\partial f(x)}{\partial x}.\tag{5.144}$$

On the other hand, it follows from our interpretation of $|\Psi(x, t)|^2$ as the probability of finding the particle at x at time t that

$$\langle \hat{x} \rangle = \int_{-\infty}^{\infty} dx \Psi^*(x, t) \hat{x} \Psi(x, t) = \int_{-\infty}^{\infty} dx \Psi^*(x, t) x \Psi(x, t),\tag{5.145}$$

and therefore that the position operator \hat{x} simply acts as multiplication by x , *i.e.*

$$\hat{x} : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R}) \quad f(x) \mapsto \hat{x}f(x) \equiv x f(x).\tag{5.146}$$

Now let us determine the *commutator* of \hat{x} and \hat{p} , *i.e.* let us determine

$$[\hat{x}, \hat{p}] = \hat{x}\hat{p} - \hat{p}\hat{x}. \quad (5.147)$$

We evaluate this commutator by studying its action on an arbitrary element $|f\rangle \in L^2(\mathbb{R})$. This is to say, we calculate

$$\hat{x}\hat{p}f(x) = x \left(-i\hbar \frac{\partial f(x)}{\partial x} \right) = -i\hbar x \frac{\partial f(x)}{\partial x} \quad (5.148)$$

and

$$\hat{p}\hat{x}f(x) = -i\hbar \frac{\partial}{\partial x} (xf(x)) = -i\hbar f(x) - i\hbar x \frac{\partial f(x)}{\partial x}. \quad (5.149)$$

Subtracting these two expressions we therefore find that

$$[\hat{x}, \hat{p}]f = i\hbar f(x) \quad (5.150)$$

which we can write as the multiplication operator

$$[\hat{x}, \hat{p}] = i\hbar. \quad (5.151)$$

Thus, in particular, the commutator does not vanish. As we shall see shortly the fact that this commutator is non-zero is at the heart of the ‘Heisenberg uncertainty principle’. Before we begin to derive this, let us first try to get a better feel for some of the implications of non-commuting observables.

5.1 Measuring observables that do not commute

In order to illustrate some of the complications that arise whenever we are dealing with observables that do not commute, let us consider the following toy example. Suppose that our quantum system is two-dimensional with orthonormal vectors $|\chi_1\rangle$ and $|\chi_2\rangle$ which are eigenvectors of the Hamilton operator with eigenvalues

$$\hat{H}|\chi_1\rangle = E_1|\chi_1\rangle \quad \hat{H}|\chi_2\rangle = E_2|\chi_2\rangle. \quad (5.152)$$

Thus $|\chi_1\rangle$ and $|\chi_2\rangle$ are solutions to the time-independent Schrödinger equation. Now suppose that we are interested in the observable \hat{S} that acts as

$$\hat{S}|\chi_1\rangle = |\chi_2\rangle \quad \hat{S}|\chi_2\rangle = |\chi_1\rangle. \quad (5.153)$$

It is easy to see that \hat{S} is Hermitian (see the problem sets). If we represent the two operators by 2×2 matrices we therefore have

$$\hat{H} = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} \quad \hat{S} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (5.154)$$

where

$$|\chi_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\chi_2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (5.155)$$

Thus the commutator is

$$\begin{aligned} [\hat{H}, \hat{S}] &= \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & E_1 \\ E_2 & 0 \end{pmatrix} - \begin{pmatrix} 0 & E_2 \\ E_1 & 0 \end{pmatrix} \\ &= (E_1 - E_2) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \end{aligned} \quad (5.156)$$

which is non-zero provided that $E_1 - E_2 \neq 0$, *i.e.* that the two eigenvectors of the Hamilton operator have different eigenvalues.

The eigenvectors of the Hamilton operator are $|\chi_1\rangle$ and $|\chi_2\rangle$, but these are clearly not eigenvectors of \hat{S} . If we want to work out the probabilities of measuring \hat{S} , it is useful to determine these. In order to determine the eigenvalues we perform the standard calculation: first we determine the eigenvalues by finding the zeros of the characteristic equation,

$$\det(\hat{S} - \lambda \hat{1}) = \det \begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 - 1. \quad (5.157)$$

Thus the two eigenvalues are ± 1 . By inspection, the (normalised) eigenvector corresponding to $s_1 = +1$ is

$$|\psi_1\rangle = \frac{1}{\sqrt{2}} (|\chi_1\rangle + |\chi_2\rangle) \quad (5.158)$$

while the normalised eigenvector corresponding to $s_2 = -1$ is

$$|\psi_2\rangle = \frac{1}{\sqrt{2}} (|\chi_1\rangle - |\chi_2\rangle). \quad (5.159)$$

Suppose now that we are interested in the following problem: At time $t = 0$, \hat{S} is measured and the value $+1$ is found. The system is then left undisturbed for time t , and then \hat{S} is measured again. Find the probability that the value is again $+1$.

According to our rules, the system is in the eigenstate corresponding to $s_1 = +1$ after the measurement at $t = 0$. This is to say, the system is described by

$$|\Psi(t=0)\rangle = |\psi_1\rangle = \frac{1}{\sqrt{2}} (|\chi_1\rangle + |\chi_2\rangle). \quad (5.160)$$

(Without loss of generality we have chosen here a definite value of the overall phase.) As we have seen before, the solution to the time-dependent Schrödinger equation is then for $t > 0$

$$|\Psi(t)\rangle = \frac{1}{\sqrt{2}} \left(|\chi_1\rangle e^{-i\frac{E_1 t}{\hbar}} + |\chi_2\rangle e^{-i\frac{E_2 t}{\hbar}} \right). \quad (5.161)$$

In order to determine the probability of the two possible outcomes of measuring \hat{S} at $t > 0$, we now have to write this again in terms of ψ_1 and ψ_2 (with time-dependent coefficients). In order to determine these coefficient we use the by now familiar trick: make the ansatz

$$|\Psi(t)\rangle = a_1(t)|\psi_1\rangle + a_2(t)|\psi_2\rangle, \quad (5.162)$$

and then use the fact that $|\psi_1\rangle$ and $|\psi_2\rangle$ are orthonormal to deduce that

$$a_i(t) = \langle\psi_i|\Psi(t)\rangle \quad i = 1, 2. \quad (5.163)$$

Thus, in particular, we have

$$\begin{aligned} a_1(t) &= \frac{1}{2} \langle\chi_1 + \chi_2|\chi_1 e^{-i\frac{E_1 t}{\hbar}} + \chi_2 e^{-i\frac{E_2 t}{\hbar}}\rangle \\ &= \frac{1}{2} \left(e^{-i\frac{E_1 t}{\hbar}} + e^{-i\frac{E_2 t}{\hbar}} \right) \\ &= \frac{1}{2} e^{-i\frac{(E_1+E_2)t}{2\hbar}} \left(e^{-i\frac{(E_1-E_2)t}{2\hbar}} + e^{i\frac{(E_1-E_2)t}{2\hbar}} \right) \\ &= e^{-i\frac{(E_1+E_2)t}{2\hbar}} \cos\left(\frac{(E_1 - E_2)t}{2\hbar}\right) \end{aligned} \quad (5.164)$$

and

$$\begin{aligned} a_2(t) &= \frac{1}{2} \langle\chi_1 - \chi_2|\chi_1 e^{-i\frac{E_1 t}{\hbar}} + \chi_2 e^{-i\frac{E_2 t}{\hbar}}\rangle \\ &= \frac{1}{2} \left(e^{-i\frac{E_1 t}{\hbar}} - e^{-i\frac{E_2 t}{\hbar}} \right) \\ &= \frac{1}{2} e^{-i\frac{(E_1+E_2)t}{2\hbar}} \left(e^{-i\frac{(E_1-E_2)t}{2\hbar}} - e^{i\frac{(E_1-E_2)t}{2\hbar}} \right) \\ &= -i e^{-i\frac{(E_1+E_2)t}{2\hbar}} \sin\left(\frac{(E_1 - E_2)t}{2\hbar}\right). \end{aligned} \quad (5.165)$$

Given these formulas for $a_1(t)$ and $a_2(t)$ we can now read off the probability of measuring $S = +1$ after time t : it is simply

$$P(S = +1, t) = |a_1(t)|^2 = \cos^2\left(\frac{(E_1 - E_2)t}{2\hbar}\right), \quad (5.166)$$

while the probability of measuring $S = -1$ after time t is

$$P(S = -1, t) = |a_2(t)|^2 = \sin^2\left(\frac{(E_1 - E_2)t}{2\hbar}\right). \quad (5.167)$$

Obviously the two probabilities add to one. However, in contrast to what we have seen before, the probabilities depend on time in a fairly ‘quantum’ (*i.e.* non-classical) fashion.

5.2 uncertainty of an observable

We briefly introduced the notion of the uncertainty in the measurement of position, energy and momentum early on when we discussed the probabilistic interpretation of the wavefunction. Let us now return in more detail to uncertainty and see in particular that its roots lie in the non-commutivity of observables.

The expectation value of the observable \hat{A} in the state $|\psi\rangle$ is given as

$$\langle \hat{A} \rangle = \langle \psi | \hat{A} \psi \rangle. \quad (5.168)$$

Earlier we introduced the notion of the *uncertainty* of observing \hat{A} as

$$(\Delta \hat{A})^2 = \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2 \quad (5.169)$$

However we did not say much about it or even show that it is real and positive. But we are now in a position to do better. We start by noting that

$$\begin{aligned} \langle (\hat{A} - \langle \hat{A} \rangle)^2 \rangle &= \langle \hat{A}^2 - 2\hat{A}\langle \hat{A} \rangle + \langle \hat{A} \rangle^2 \rangle \\ &= \langle \hat{A}^2 \rangle - 2\langle \hat{A}\langle \hat{A} \rangle \rangle + \langle \langle \hat{A} \rangle^2 \rangle \\ &= \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2 \\ &= (\Delta \hat{A})^2. \end{aligned} \quad (5.170)$$

We can now also prove that the square of the uncertainty is a non-negative real number since

$$\begin{aligned} (\Delta \hat{A})^2 &= \langle (\hat{A} - \langle \hat{A} \rangle)^2 \rangle \\ &= \langle \psi | (\hat{A} - \langle \hat{A} \rangle) (\hat{A} - \langle \hat{A} \rangle) \psi \rangle \\ &= \langle (\hat{A} - \langle \hat{A} \rangle) \psi | (\hat{A} - \langle \hat{A} \rangle) \psi \rangle \\ &\geq 0, \end{aligned} \quad (5.171)$$

where we have used that $(\hat{A} - \langle \hat{A} \rangle)$ is a Hermitian operator since \hat{A} is Hermitian and therefore its expectation value, $\langle \hat{A} \rangle$, is real. Since the square of the uncertainty is a non-negative real number, we can define the uncertainty itself to be non-negative (and real). This is what we shall always do in the following.

Recall that we calculated the uncertainty before for the case of the particle in the box. For this example, the n th eigenfunction of the Hamilton operator was given by

$$\psi_n(x) = \sqrt{\frac{2}{l}} \sin\left(\frac{n\pi x}{l}\right) \quad \text{with} \quad E_n = \frac{n^2 \pi^2 \hbar^2}{2ml^2}, \quad (5.172)$$

where $0 \leq x \leq l$ and $\psi_n(x) = 0$ for $|x| > l$. In the state described by $\psi_n(x)$ we calculated the expectation value of x to be

$$\langle \hat{x} \rangle = \frac{l}{2} \quad (5.173)$$

as well as

$$(\Delta \hat{x})^2 = \langle (\hat{x} - \langle \hat{x} \rangle)^2 \rangle = \frac{l^2}{12} \left(1 - \frac{6}{n^2 \pi^2} \right). \quad (5.174)$$

and the corresponding expressions for momentum

$$\langle \hat{p} \rangle = 0, \quad (\Delta \hat{p})^2 = \frac{n^2 \pi^2 \hbar^2}{l^2}. \quad (5.175)$$

Here we see that

$$\begin{aligned} (\Delta \hat{p})^2 (\Delta \hat{x})^2 &= \frac{\hbar^2 n^2 \pi^2}{12} - \frac{\hbar^2}{2} \\ &\geq \frac{\hbar^2 \pi^2}{12} - \frac{\hbar^2}{2} \\ &\sim 0.32 \hbar^2 \end{aligned} \quad (5.176)$$

Here we have noted that the right hand side is minimized for $n = 1$. This is an example of the Heisenberg uncertainty principle

$$(\Delta \hat{p})^2 (\Delta \hat{x})^2 \geq \hbar^2 / 4 \quad (5.177)$$

which is valid for any quantum mechanical system.

5.3 The Heisenberg uncertainty principle

Here we want to prove the famous Heisenberg uncertainty theorem:

Theorem Suppose that \hat{A} and \hat{B} are two observables of a system in an arbitrary state. Then

$$\Delta \hat{A} \Delta \hat{B} \geq \frac{1}{2} \left| \langle i[\hat{A}, \hat{B}] \rangle \right|. \quad (5.178)$$

Here the operator $[\hat{A}, \hat{B}] = \hat{A} \hat{B} - \hat{B} \hat{A}$ is the commutator of \hat{A} and \hat{B} .

proof: It is easy to see that $i[\hat{A}, \hat{B}]$ is an Hermitian operator provided that \hat{A} and \hat{B} are. The expectation value of a Hermitian operator is always real, and therefore the expression in the modulus is actually a real number. Thus (5.178) is equivalent to the statement

$$(\Delta \hat{A})^2 (\Delta \hat{B})^2 \geq \frac{1}{4} \langle i[\hat{A}, \hat{B}] \rangle^2. \quad (5.179)$$

This is the statement that we shall attempt to prove in the following.

Let us define $\hat{C} = \hat{A} - \langle \hat{A} \rangle$ and $\hat{D} = \hat{B} - \langle \hat{B} \rangle$. Since \hat{A} and \hat{B} are Hermitian operators, their expectation values are real and therefore \hat{C} and \hat{D} are also Hermitian operators. We observe that

$$[\hat{C}, \hat{D}] = [\hat{A} - \langle \hat{A} \rangle, \hat{B} - \langle \hat{B} \rangle] = [\hat{A}, \hat{B}]. \quad (5.180)$$

In terms of \hat{C} and \hat{D} , (5.179) is therefore simply the statement that

$$\langle \hat{C}^2 \rangle \langle \hat{D}^2 \rangle \geq \frac{1}{4} \langle i[\hat{C}, \hat{D}] \rangle^2, \quad (5.181)$$

where \hat{C} and \hat{D} are Hermitian. Let us denote by ψ the state of the system, and let s be an arbitrary real number. Then

$$0 \leq \langle (\hat{C} + is\hat{D})\psi | (\hat{C} + is\hat{D})\psi \rangle = \langle \psi | (\hat{C} + is\hat{D})^\dagger (\hat{C} + is\hat{D})\psi \rangle = \langle (\hat{C} + is\hat{D})^\dagger (\hat{C} + is\hat{D}) \rangle. \quad (5.182)$$

Since both C and D are Hermitian, we therefore obtain

$$0 \leq \langle (\hat{C} - is\hat{D})(\hat{C} + is\hat{D}) \rangle = \langle \hat{C}^2 \rangle + s \langle i[\hat{C}, \hat{D}] \rangle + s^2 \langle \hat{D}^2 \rangle. \quad (5.183)$$

Next we observe that \hat{C}^2 , \hat{D}^2 and $i[\hat{C}, \hat{D}]$ are all Hermitian and so have real expectation values. Thus 5.183 is a real quadratic expression in s with at most one zero. Therefore its discriminant must be non-positive and hence

$$\langle i[\hat{C}, \hat{D}] \rangle^2 - 4 \langle \hat{C}^2 \rangle \langle \hat{D}^2 \rangle \leq 0, \quad (5.184)$$

which is indeed (5.181).

On slight subtlety comes if $\langle \hat{D}^2 \rangle = 0$. To see that this is harmless we first note that

$$\langle \hat{D}^2 \rangle = \langle \psi | \hat{D}^2 \psi \rangle = \langle \hat{D} \psi | \hat{D} \psi \rangle \geq 0, \quad (5.185)$$

where we have used that \hat{D} is Hermitian. We can thus conclude that $\langle \hat{D}^2 \rangle = 0$ if and only if $\hat{D}|\psi\rangle = 0$. If this is the case, then (5.181) is trivially true since we then also have

$$\langle [\hat{C}, \hat{D}] \rangle = \langle \psi | [\hat{C}, \hat{D}] \psi \rangle = \langle \psi | \hat{C} \hat{D} \psi \rangle - \langle \psi | \hat{D} \hat{C} \psi \rangle = \langle \hat{C} \psi | \hat{D} \psi \rangle - \langle \hat{D} \psi | \hat{C} \psi \rangle = 0, \quad (5.186)$$

where we have used that \hat{C} and \hat{D} are Hermitian and $\hat{D}|\psi\rangle = 0$.

The familiar Heisenberg uncertainty relation between position and momentum is now a simple consequence of this theorem. Recall that the commutator of \hat{x} and \hat{p} is given by

$$[\hat{x}, \hat{p}] = i\hbar. \quad (5.187)$$

Thus we obtain

$$\Delta \hat{x} \Delta \hat{p} \geq \frac{\hbar}{2}, \quad (5.188)$$

which is indeed the Heisenberg uncertainty relation.

6 additional topics

6.1 the harmonic oscillator

After having introduced a more formal notion of quantum mechanics let us consider another physical problem, the harmonic oscillator. Here we have particle subject to a simple quadratic potential

$$V(x) = \frac{1}{2}\omega x^2, \quad (6.1)$$

hence the Schrödinger equation is

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + \frac{1}{2}\omega x^2 \Psi \quad (6.2)$$

We could proceed by finding the general solution to this differential equation. However this is a difficult task and a much better analysis can be done using the ideas of Hilbert spaces.

Again our Hilbert space is $\mathcal{H} = L^2(\mathbb{R})$. We can write the Schrödinger equation as

$$\hat{E}|\Psi\rangle = \left(\frac{1}{2m}\hat{p}^2 + \frac{\omega}{2}\hat{x}^2 \right) |\Psi\rangle \quad (6.3)$$

Next we introduce new operators \hat{a} and \hat{a}^\dagger

$$\hat{a} = \left(\frac{1}{2\hbar\sqrt{m\omega}} \right)^{\frac{1}{2}} \hat{p} - i \left(\frac{\sqrt{m\omega}}{2\hbar} \right)^{\frac{1}{2}} \hat{x}, \quad \hat{a}^\dagger = \left(\frac{1}{2\hbar\sqrt{m\omega}} \right)^{\frac{1}{2}} \hat{p} + i \left(\frac{\sqrt{m\omega}}{2\hbar} \right)^{\frac{1}{2}} \hat{x} \quad (6.4)$$

These satisfy

$$\begin{aligned} \hat{a}\hat{a}^\dagger &= \frac{1}{2\hbar\sqrt{m\omega}}\hat{p}^2 + \frac{\sqrt{m\omega}}{2\hbar}\hat{x}^2 - \frac{i}{2\hbar}[\hat{x}, \hat{p}] = \frac{1}{2\hbar\sqrt{m\omega}}\hat{p}^2 + \frac{\sqrt{m\omega}}{2\hbar}\hat{x}^2 + \frac{1}{2} \\ \hat{a}^\dagger\hat{a} &= \frac{1}{2\hbar\sqrt{m\omega}}\hat{p}^2 + \frac{\sqrt{m\omega}}{2\hbar}\hat{x}^2 + \frac{i}{2\hbar}[\hat{x}, \hat{p}] = \frac{1}{2\hbar\sqrt{m\omega}}\hat{p}^2 + \frac{\sqrt{m\omega}}{2\hbar}\hat{x}^2 - \frac{1}{2} \end{aligned} \quad (6.5)$$

and hence

$$[\hat{a}, \hat{a}^\dagger] = 1 \quad (6.6)$$

Thus the Hamiltonian can be written as

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + \frac{\omega}{2}\hat{x}^2 = \sqrt{\frac{\omega\hbar^2}{m}} \left(\hat{a}^\dagger\hat{a} + \frac{1}{2} \right) \quad (6.7)$$

We know that the normalisable modes have energies which are bounded below by zero. In particular there is a lowest energy state that we call the ground state $|0\rangle$. Let us introduce another operator \hat{N} called the number operator

$$\hat{N} = \hat{a}^\dagger\hat{a} \quad (6.8)$$

i.e. $\hat{H} = \sqrt{\frac{\omega\hbar^2}{m}}(\hat{N} + \frac{1}{2})$. Note that \hat{N} is Hermitian with positive eigenvalues

$$\langle \Psi | \hat{N} \Psi \rangle = \langle \Psi | \hat{a}^\dagger \hat{a} \Psi \rangle = \langle \hat{a} \Psi | \hat{a} \Psi \rangle \geq 0 \quad (6.9)$$

Thus it follows that the lowest energy eigenstate $|0\rangle$ must satisfy

$$\hat{N}|0\rangle = 0 \quad (6.10)$$

which means that

$$\hat{a}|0\rangle = 0 \quad (6.11)$$

hence $\hat{H}|0\rangle = \frac{1}{2}\sqrt{\frac{\omega\hbar^2}{m}}|0\rangle$, *i.e.* the ground state energy is

$$E_0 = \frac{1}{2}\sqrt{\frac{\omega\hbar^2}{m}} \quad (6.12)$$

We can create new states by acting with \hat{a}^\dagger

$$|n\rangle = N_n(\hat{a}^\dagger)^n|0\rangle \quad (6.13)$$

where N_n is a normalisation constant chosen so that $\langle n|n\rangle = 1$. In particular it is (*see the problem sets*)

$$N_n = \sqrt{\frac{1}{n!}} \quad (6.14)$$

Next we prove the following lemma: $\hat{N}|n\rangle = n|n\rangle$. To do this we proceed by induction. By definition we have that $\hat{N}|0\rangle = 0$. Next we consider, assuming that $\hat{N}|n-1\rangle = (n-1)|n-1\rangle$,

$$\begin{aligned} \hat{N}|n\rangle &= \frac{N_n}{N_{n-1}}\hat{N}\hat{a}^\dagger|n-1\rangle \\ &= \frac{N_n}{N_{n-1}}\hat{a}^\dagger\hat{a}\hat{a}^\dagger|n-1\rangle \\ &= \frac{N_n}{N_{n-1}}(\hat{a}^\dagger[\hat{a}, \hat{a}^\dagger] + \hat{a}^\dagger\hat{a}^\dagger\hat{a})|n-1\rangle \\ &= \frac{N_n}{N_{n-1}}\hat{a}^\dagger(1 + \hat{N})|n-1\rangle \\ &= n\frac{N_n}{N_{n-1}}\hat{a}^\dagger|n-1\rangle \\ &= n|n\rangle \end{aligned} \quad (6.15)$$

This can also be seen from the following commutators

$$\begin{aligned} [\hat{N}, \hat{a}] &= [\hat{a}^\dagger\hat{a}, \hat{a}] \\ &= \hat{a}^\dagger\hat{a}\hat{a} - \hat{a}\hat{a}^\dagger\hat{a} \\ &= \hat{a}^\dagger\hat{a}\hat{a} - [\hat{a}, \hat{a}^\dagger]\hat{a} - \hat{a}^\dagger\hat{a}\hat{a} \\ &= -\hat{a} \end{aligned} \quad (6.16)$$

and

$$\begin{aligned}
[\hat{N}, \hat{a}^\dagger] &= [\hat{a}^\dagger \hat{a}, \hat{a}^\dagger] \\
&= \hat{a}^\dagger \hat{a} \hat{a}^\dagger - \hat{a}^\dagger \hat{a}^\dagger \hat{a} \\
&= \hat{a}^\dagger \hat{a} \hat{a}^\dagger + \hat{a}^\dagger [\hat{a}, \hat{a}^\dagger] - \hat{a}^\dagger \hat{a} \hat{a}^\dagger \\
&= \hat{a}^\dagger
\end{aligned} \tag{6.17}$$

Hence it follows that

$$\hat{N} \hat{a} |n\rangle = ([\hat{N}, \hat{a}] + \hat{a} \hat{N}) |n\rangle = (-\hat{a} + n \hat{a}) |n\rangle = (n-1) \hat{a} |n\rangle \tag{6.18}$$

and

$$\hat{N} \hat{a}^\dagger |n\rangle = ([\hat{N}, \hat{a}^\dagger] + \hat{a}^\dagger \hat{N}) |n\rangle = (\hat{a}^\dagger + n \hat{a}^\dagger) |n\rangle = (n+1) \hat{a}^\dagger |n\rangle \tag{6.19}$$

The operators \hat{a}^\dagger and \hat{a} are therefore known as raising and lowering operators respectively. From this we see that

$$\hat{H} |n\rangle = \sqrt{\frac{\omega \hbar^2}{m}} \left(\hat{N} + \frac{1}{2} \right) |n\rangle = \sqrt{\frac{\omega \hbar^2}{m}} \left(n + \frac{1}{2} \right) |n\rangle \tag{6.20}$$

i.e. the energy of the state $|n\rangle$ is

$$E_n = \sqrt{\frac{\omega \hbar^2}{m}} \left(n + \frac{1}{2} \right) \tag{6.21}$$

Thus we have deduced the entire spectrum of the Hamiltonian without ever having had to find explicit solutions of the Schrödinger equation. The interpretation of $|n\rangle$ is that corresponds to a state with n particles. The operator \hat{a} is interpreted as destroying a particle with energy $\sqrt{\frac{\omega \hbar^2}{m}}$ and its Hermitian conjugate creates a particle with energy $\sqrt{\frac{\omega \hbar^2}{m}}$. The operator \hat{N} counts the number of particles and the Hamiltonian \hat{H} is, up to an additive constant, given by the sum over the individual particle energies. However the ground state, the state with no particles, does not have zero energy but rather a non-zero vacuum energy due to the quantum oscillations of the field.

This vacuum energy is in fact deeply problematic. Quantum theory predicts that the energy of the vacuum, *i.e.* the lowest energy state, is non-zero. This is an irrelevant constant in theories without gravity, where one can simply redefine the vacuum energy to be zero. However in gravity the vacuum energy will gravitate and affect the large scale structure of the universe. Furthermore, on dimensional grounds, this vacuum energy should be so large and positive that the universe should curl up into a size no larger than a football.

We can also calculate various expectation values in this formalism. For example, in the state $|n\rangle$, we have

$$\langle \hat{x} \rangle = \langle n | \hat{x} | n \rangle$$

$$\begin{aligned}
&= \frac{i}{2} \left(\frac{2\hbar}{\sqrt{m\omega}} \right)^{\frac{1}{2}} \langle n | (\hat{a} - \hat{a}^\dagger) n \rangle \\
&= \frac{i}{2} \left(\frac{2\hbar}{\sqrt{m\omega}} \right)^{\frac{1}{2}} (\langle n | \hat{a} n \rangle - \langle n | \hat{a}^\dagger n \rangle)
\end{aligned} \tag{6.22}$$

Now $\hat{a}|n\rangle \propto |n-1\rangle$ and $\hat{a}^\dagger|n\rangle \propto |n+1\rangle$ so that

$$\langle \hat{x} \rangle \propto \langle n | n+1 \rangle - \langle n | n-1 \rangle = 0 \tag{6.23}$$

since the $|n\rangle$ are orthonormal. On the other hand

$$\begin{aligned}
\langle \hat{x}^2 \rangle &= \langle n | \hat{x}^2 n \rangle \\
&= -\frac{1}{4} \left(\frac{2\hbar}{\sqrt{m\omega}} \right) \langle n | (\hat{a} - \hat{a}^\dagger)^2 n \rangle \\
&= -\frac{1}{4} \left(\frac{2\hbar}{\sqrt{m\omega}} \right) (\langle n | (\hat{a})^2 n \rangle + \langle n | (\hat{a}^\dagger)^2 n \rangle - \langle n | \hat{a} \hat{a}^\dagger n \rangle - \langle n | \hat{a}^\dagger \hat{a} n \rangle)
\end{aligned} \tag{6.24}$$

Now again $(\hat{a})^2|n\rangle \propto |n-2\rangle$ and $(\hat{a}^\dagger)^2|n\rangle \propto |n+2\rangle$ so that the first two terms give zero. However the last two terms give

$$\langle \hat{x}^2 \rangle = \frac{1}{4} \left(\frac{2\hbar}{\sqrt{m\omega}} \right) (\langle \hat{a}^\dagger n | \hat{a}^\dagger n \rangle + \langle \hat{a} n | \hat{a} n \rangle) \tag{6.25}$$

Now we have already seen that

$$|n+1\rangle = \sqrt{\frac{1}{n+1}} \hat{a}^\dagger |n\rangle, \tag{6.26}$$

which implies that

$$\hat{a}|n+1\rangle = \sqrt{\frac{1}{n+1}} \hat{a} \hat{a}^\dagger |n\rangle = \sqrt{\frac{1}{n+1}} (\hat{N} + [\hat{a}, \hat{a}^\dagger]) |n\rangle = \sqrt{\frac{1}{n+1}} (n+1) |n\rangle \tag{6.27}$$

Thus

$$|n\rangle = \sqrt{\frac{1}{n+1}} \hat{a} |n+1\rangle \tag{6.28}$$

and finally

$$\langle \hat{x}^2 \rangle = \frac{1}{4} \left(\frac{2\hbar}{\sqrt{m\omega}} \right) (2n+1) \tag{6.29}$$

You can also evaluate $\langle \hat{p} \rangle$ and $\langle \hat{p}^2 \rangle$ this way and check that the Heisenberg uncertainty relation is satisfied (*see the problem sets*).

Furthermore this approach also allows us to explicitly construct the corresponding energy eigenstate wavefunctions. For example the ground state satisfies

$$\hat{a}|0\rangle = \left(\left(\frac{1}{2\hbar\sqrt{m\omega}} \right)^{\frac{1}{2}} \hat{p} - i \left(\frac{\sqrt{m\omega}}{2\hbar} \right)^{\frac{1}{2}} \hat{x} \right) |0\rangle = 0 \quad (6.30)$$

Recalling the definitions of \hat{x} and \hat{p} we obtain a first order differential equation

$$\left(-i\hbar \left(\frac{1}{2\hbar\sqrt{m\omega}} \right)^{\frac{1}{2}} \frac{\partial}{\partial x} - i \left(\frac{\sqrt{m\omega}}{2\hbar} \right)^{\frac{1}{2}} x \right) \Psi_0 = 0 \quad (6.31)$$

where $\Psi_0(x, t) = e^{-iE_0 t/\hbar} \psi_0(x)$ is the explicit element of $L^2(\mathbb{R})$ that represents $|0\rangle$. Rewriting this equation gives

$$\frac{d\psi_0}{dx} = -\frac{\sqrt{m\omega}}{\hbar} \psi_0 \quad (6.32)$$

The solution to this differential equation is simply $\psi \propto e^{-\sqrt{m\omega}x^2/2\hbar}$ and hence

$$\Psi = N_0 e^{-\frac{itE_0}{\hbar}} e^{-\frac{\sqrt{m\omega}}{2\hbar}x^2} \quad (6.33)$$

where N_0 is a normalisation and we have included the time-dependent phase factor $e^{-\frac{itE_0}{\hbar}}$.

Furthermore we can obtain explicit solutions for the higher energy states $\Psi_n(x, t)$ by acting with \hat{a}^\dagger

$$\begin{aligned} \Psi_n(x, t) &= \sqrt{\frac{1}{n!}} e^{-\frac{itE_n}{\hbar}} \hat{a}^\dagger \psi_0(x) \\ &= \sqrt{\frac{1}{n!}} e^{-\frac{itE_n}{\hbar}} \left(-i\hbar \left(\frac{1}{2\hbar\sqrt{m\omega}} \right)^{\frac{1}{2}} \frac{\partial}{\partial x} + i \left(\frac{\sqrt{m\omega}}{2\hbar} \right)^{\frac{1}{2}} x \right)^n \psi_0(x) \\ &= \sqrt{\frac{1}{n!}} N_0 e^{-\frac{itE_n}{\hbar}} \left(-i\hbar \left(\frac{1}{2\hbar\sqrt{m\omega}} \right)^{\frac{1}{2}} \frac{\partial}{\partial x} + i \left(\frac{\sqrt{m\omega}}{2\hbar} \right)^{\frac{1}{2}} x \right)^n e^{-\frac{\sqrt{m\omega}}{2\hbar}x^2} \end{aligned} \quad (6.34)$$

These have the form of a polynomial times the Gaussian $e^{-\frac{\sqrt{m\omega}}{2\hbar}x^2}$. We can explicitly construct the first few wavefunctions, for example,

$$\begin{aligned} \Psi_1(x, t) &= N_0 e^{-\frac{itE_1}{\hbar}} \left(-i\hbar \left(\frac{1}{2\hbar\sqrt{m\omega}} \right)^{\frac{1}{2}} \frac{\partial}{\partial x} + i \left(\frac{\sqrt{m\omega}}{2\hbar} \right)^{\frac{1}{2}} x \right) e^{-\frac{\sqrt{m\omega}}{2\hbar}x^2}, \\ &= N_0 e^{-\frac{itE_1}{\hbar}} \left(+i \left(\frac{\sqrt{m\omega}}{2\hbar} \right)^{\frac{1}{2}} x + i \left(\frac{\sqrt{m\omega}}{2\hbar} \right)^{\frac{1}{2}} x \right) e^{-\frac{\sqrt{m\omega}}{2\hbar}x^2} \\ &= 2iN_0 \left(\frac{\sqrt{m\omega}}{2\hbar} \right)^{\frac{1}{2}} e^{-\frac{itE_1}{\hbar}} x e^{-\frac{\sqrt{m\omega}}{2\hbar}x^2} \end{aligned} \quad (6.35)$$

It would have been very difficult to find all these solutions to the Schrödinger equation by brute force and also to use them to calculate quantities such as $\langle \hat{x} \rangle$ and $\langle \hat{x}^2 \rangle$.

6.2 symmetries in quantum mechanics

The first requirement of symmetry of a quantum mechanical system is that it preserves the probabilities. Thus there is an operator \hat{Q} such that

$$|\langle \hat{Q}\phi | \hat{Q}\psi \rangle|^2 = |\langle \phi | \psi \rangle|^2 \quad (6.36)$$

note that this is slightly more general than the condition that \hat{Q} is unitary. For example $\hat{Q}^\dagger \hat{Q}$ could be the identity operator times a phase, *i.e.* a complex number with unit length. However this condition contains no information about the dynamics of a particular system. Therefore in addition we require that a symmetry that commutes with the Hamiltonian

$$[\hat{H}, \hat{Q}] = 0 \quad (6.37)$$

This ensures that if $|\Psi\rangle$ is an eigenstate of the Hamiltonian with energy E then $\hat{Q}|\Psi\rangle$ is also an eigenstate with the same value of the energy:

$$\hat{H}(\hat{Q}|\Psi\rangle) = \hat{Q}\hat{H}|\Psi\rangle = \hat{Q}E|\Psi\rangle = E(\hat{Q}|\Psi\rangle) \quad (6.38)$$

An important example of a symmetry is time-reversal. Provided that the potential is independent of time the Schrödinger equation is invariant under $t \rightarrow -t$ and $|\psi\rangle \rightarrow |\psi\rangle^*$. Thus in terms of wave functions in $L^2(\mathbb{R})$ we have $\hat{Q}\Psi(t) = \Psi^*(-t)$. To see that this commutes with \hat{H} we note that

$$\hat{H}\hat{Q}\Psi(t) = \hat{H}\Psi^*(-t) = (\hat{H}\Psi)^*(-t) = \hat{Q}\hat{H}\Psi(t) \quad (6.39)$$

where we used the fact that reversing the time does not alter the Hamiltonian. Thus $[\hat{H}, \hat{Q}] = 0$. To check the conservation of probability we evaluate

$$\langle \hat{Q}\phi | \hat{Q}\psi \rangle = (\langle \phi |)^*(|\psi\rangle)^* = \langle \phi | \psi \rangle^* \quad (6.40)$$

Thus the probability is conserved even though the inner-product isn't. In particular \hat{Q} is not unitary.

On the other hand most symmetries are unitary. Let us assume that \hat{Q} is unitary and furthermore comes from exponentiating a Hermitian operator \hat{S}

$$\hat{Q} = e^{i\theta\hat{S}} \quad (6.41)$$

where θ is a real number. For example this arises if \hat{Q} represents a continuous symmetry which is generated infinitesimally by \hat{S} . We note that \hat{Q} will commute with \hat{H} if and only if \hat{S} does.

Now \hat{S} is an observable and it follows that its expectation value is constant

$$\begin{aligned}
\frac{d}{dt}\langle\hat{S}\rangle &= \langle\frac{\partial\Psi}{\partial t}|\hat{S}\Psi\rangle + \langle\Psi|\hat{S}\frac{\partial\Psi}{\partial t}\rangle \\
&= \langle\frac{-i}{\hbar}\hat{H}\Psi|\hat{S}\Psi\rangle + \langle\Psi|\hat{S}\frac{-i}{\hbar}\hat{H}\Psi\rangle \\
&= \frac{i}{\hbar}\langle\Psi|\hat{H}\hat{S}\Psi\rangle - \frac{i}{\hbar}\langle\Psi|\hat{S}\hat{H}\Psi\rangle \\
&= -\frac{i}{\hbar}\langle[\hat{S},\hat{H}]\rangle = 0
\end{aligned} \tag{6.42}$$

It also follows that if at $t = 0$ $|\Psi(0)\rangle$ is an eigenstate of \hat{S} with eigenvalue λ , then $|\Psi(t)\rangle$ is an eigenstate (with the same eigenvalue) for all t

$$\begin{aligned}
\hat{S}|\Psi(t)\rangle &= \hat{S}e^{-\frac{it}{\hbar}\hat{H}}|\Psi(0)\rangle \\
&= e^{-\frac{it}{\hbar}\hat{H}}\hat{S}|\Psi(0)\rangle \\
&= e^{-\frac{it}{\hbar}\hat{H}}\lambda|\Psi(0)\rangle \\
&= \lambda|\Psi(t)\rangle
\end{aligned} \tag{6.43}$$

Here we used the fact that if \hat{S} commutes with \hat{H} then it commutes with \hat{H}^n for any n and hence also with any function of \hat{H} .

For example let us consider the free Schrödinger equation

$$\hat{H} = \frac{\hat{p}^2}{2m} \tag{6.44}$$

We immediately see that $[\hat{p},\hat{H}] = 0$. Thus \hat{p} generates a symmetry. Indeed this is just translational invariance

$$\begin{aligned}
\hat{Q}\Psi(x,t) &= e^{i\theta\hat{p}}\Psi(x) \\
&= \sum_{n=0}^{\infty} \frac{(i\theta)^n}{n!} \left(-i\hbar\frac{\partial}{\partial x}\right)^n \Psi(x,t) \\
&= \sum_{n=0}^{\infty} \frac{(\hbar\theta)^n}{n!} \frac{\partial^n\Psi(x,t)}{\partial x^n} \\
&= \Psi(x + \hbar\theta)
\end{aligned} \tag{6.45}$$

where we used Taylor's theorem

$$\Psi(x + a, t) = \sum_{n=0}^{\infty} \frac{a^n}{n!} \frac{\partial^n\Psi(x, t)}{\partial x^n} \tag{6.46}$$

And then conserved quantity is momentum $\langle\hat{p}\rangle$, which we can explicitly check is constant (*c.f.* Erhenfest's theorem)

$$\frac{\partial\langle\hat{p}\rangle}{\partial t} = -i\hbar\frac{\partial}{\partial t}\int\Psi^*(x)\frac{\partial\Psi}{\partial x}dx$$

$$\begin{aligned}
&= \int \left(\hat{H}\Psi^* \frac{\partial\Psi}{\partial x} - \Psi^* \frac{\partial\hat{H}\Psi}{\partial x} \right) \\
&= \int \left(-\frac{\hbar^2}{2m} \frac{\partial^2\Psi^*}{\partial x^2} \frac{\partial\Psi}{\partial x} + \frac{\hbar^2}{2m} \Psi^* \frac{\partial^2\Psi}{\partial x^2} \right) dx \\
&= -\frac{\hbar^2}{2m} \int \frac{\partial}{\partial x} \left(\frac{\partial\Psi^*}{\partial x} \frac{\partial\Psi}{\partial x} - \Psi^* \frac{\partial\Psi}{\partial x} \right) dx \\
&= 0
\end{aligned} \tag{6.47}$$

where we used the usual fact that the integral of a total derivative vanishes (for functions in $L^2(\mathbb{R})$).

6.3 the Heisenberg picture

Up to now we have formulated quantum mechanics in what is usually referred to as the *Schrödinger picture*. This is to say, the wave-functions are time-dependent and they satisfy the time-dependent Schrödinger equation

$$i\hbar \frac{\partial|\Psi\rangle}{\partial t} = \hat{H}|\Psi\rangle, \tag{6.48}$$

whereas the observables (such as the position, the momentum, *etc.*) are typically independent of time. As we have just seen, the ‘time-evolution’ operator \hat{U}_t is a unitary operator, and we can therefore also consider a unitarily equivalent description in which the states are time-independent, but the observables are time-dependent. This second point of view is usually called the *Heisenberg picture*.

In order to understand more precisely what the Heisenberg picture is, let us consider the expectation value of an observable in the Schrödinger picture

$$\langle \hat{A} \rangle_t = \langle \Psi(t) | \hat{A} \Psi(t) \rangle. \tag{6.49}$$

Because of (4.137) we can now write this as

$$\langle \hat{A} \rangle_t = \langle \hat{U}_t \Psi(0) | \hat{A} \hat{U}_t \Psi(0) \rangle = \langle \Psi(0) | \hat{U}_t^\dagger \hat{A} \hat{U}_t \Psi(0) \rangle. \tag{6.50}$$

Thus if we define

$$\hat{A}(t) = \hat{U}_t^\dagger \hat{A} \hat{U}_t, \tag{6.51}$$

we now have

$$\langle \hat{A} \rangle_t = \langle \Psi(0) | \hat{A}(t) \Psi(0) \rangle. \tag{6.52}$$

This is the Heisenberg picture in which the states are time-independent (they are evaluated at a fixed time), and the time dependence has been shifted into the operators.

Instead of the (time-dependent) Schrödinger equation that determines the time evolution of the states, in the Heisenberg picture the operators depend on time, and their dependence on time is governed by the equation

$$\frac{d}{dt}\hat{A}(t) = \frac{i}{\hbar}[\hat{H}, \hat{A}](t). \quad (6.53)$$

Indeed, since $\hat{A}(t)$ is defined by (6.51) we have

$$\begin{aligned} \frac{d}{dt}\hat{A}(t) &= \left(\frac{\partial}{\partial t}\hat{U}^\dagger\right)\hat{A}\hat{U}_t + \hat{U}_t^\dagger\hat{A}\frac{\partial}{\partial t}\hat{U}_t \\ &= \frac{i}{\hbar}\hat{U}_t^\dagger\hat{H}\hat{A}\hat{U}_t - \frac{i}{\hbar}\hat{U}_t^\dagger\hat{A}\hat{H}\hat{U}_t \\ &= \frac{i}{\hbar}\hat{U}_t^\dagger(\hat{H}\hat{A} - \hat{A}\hat{H})\hat{U}_t \\ &= \frac{i}{\hbar}[\hat{H}, \hat{A}](t), \end{aligned} \quad (6.54)$$

where we have used that

$$\hat{U}_t = e^{-i\frac{t}{\hbar}\hat{H}}, \quad \hat{U}_t^\dagger = e^{+i\frac{t}{\hbar}\hat{H}}. \quad (6.55)$$

In particular, it follows from this analysis that the expectation value of operators that commute with the Hamilton operator is time independent

6.4 a definition of quantum mechanics

Let us end with a set of axioms that “define” quantum mechanics in the Schrödinger picture. These can be compared and contrasted with the corresponding axioms of classical mechanics (for a brief review of classical mechanics in the Hamiltonian formalism see appendix C). Hopefully this also serves to sum up what we have discussed.

We consider a Hilbert space \mathcal{H} along with a Hermitian operator \hat{H} called the Hamiltonian whose eigenvalues E are interpreted as energies. Time evolution is determined by the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}|\Psi\rangle = \hat{H}|\Psi\rangle \quad (6.56)$$

where $|\Psi\rangle \in \mathcal{H}$ is the state of the system at time t .

Experiments correspond to observations of certain Hermitian linear operators \hat{A} on \mathcal{H} called observables. The eigenvalues λ_n of \hat{A} are the possible outcomes of any measurement. The probability that the system is found in an experiment to have the value λ_n is

$$|\langle\psi_n|\Psi\rangle|^2 \quad (6.57)$$

where $|\psi_n\rangle$ is the eigenvector of \hat{A} with eigenvalue λ_n . The average outcome over many observations is the expectation value

$$\langle\hat{A}\rangle = \langle\Psi|\hat{A}\Psi\rangle \quad (6.58)$$

After a measurement the wave function collapses to the corresponding eigenstate wave function of \hat{A} , but will then evolve according to the Schrödinger equation 6.56.

Erhenfest's theorem ensures that these expectation values satisfy the classical equations of motion. The correspondence principle further asserts that in a large quantum system, with many or large quantum numbers, the effects of uncertainty go away and the classical predictions are recovered to a high degree of accuracy. However we also proved the Heisenberg uncertainty relation which says that for a pair of observables \hat{A} and \hat{B} , the product of the two uncertainties is bounded from below by $\frac{1}{2}|\langle i[\hat{A}, \hat{B}] \rangle|$. Thus if \hat{A} and \hat{B} do not commute then the more accurately one knows the value of \hat{A} the less accurately one can know the value of \hat{B} .

The operators corresponding to observables are obtained from their classical analogues by the condition that the Poisson bracket is mapped to the commutator, with a factor of $i\hbar$. More specifically if A and B are two classical quantities (that is functions on the classical phase space) which are conjugate to each other, *i.e.*

$$\{A, B\} = 1 \tag{6.59}$$

then in the quantum theory these are represented by Hermitian linear operators \hat{A} and \hat{B} acting on \mathcal{H} such that

$$[\hat{A}, \hat{B}] = i\hbar. \tag{6.60}$$

The primary example of this are the position and momentum x, p . In classical Hamiltonian mechanics these are functions on the phase space, indeed they are usually taken to be the coordinates on phase space, that satisfy $\{x, p\} = 1$. In the quantum theory they are therefore represented by operators \hat{x} and \hat{p} which act on \mathcal{H} in such a way that $[\hat{x}, \hat{p}] = i\hbar$. The Hamiltonian which was a function of x and p in the classical theory now becomes an operator $H(x, p) \rightarrow \hat{H}(\hat{x}, \hat{p})$.

The process of replacing the phase space of classical mechanics by a Hilbert space and a replacement of the functions on classical phase space by non-commuting Hermitian operators is called quantisation. In general this process is not unique. For example in the classical theory one might have an expression in the Hamiltonian of the form $x^2 p^2$. In the quantum theory this could be taken to be $\hat{x}^2 \hat{p}^2$ or $\hat{x} \hat{p} \hat{x} \hat{p}$ or various other inequivalent operator expressions.

Appendix A: conventions

In these notes an arrow over a symbol, *e.g.* \vec{p} , denotes a three-dimensional vector. Its norm is then denoted by $p = |\vec{p}| = \sqrt{\vec{p} \cdot \vec{p}}$, where \cdot is the vector inner product;

$$\vec{a} \cdot \vec{b} = (\vec{a})^T \vec{b} = (a_x, a_y, a_z) \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} = a_x b_x + a_y b_y + a_z b_z . \quad (6.1)$$

We generally use the symbol $\Psi = \Psi(\vec{x}, t)$ for a solution to the time-dependent Schrödinger equation and $\psi = \psi(\vec{x})$ for a solution to the time-independent Schrödinger equation.

When we talk about abstract Hilbert spaces then we use a bold faced letter to denote a vector in the space. Operators acting on a Hilbert are denoted by a hat.

The Kronecker symbol δ_i^j is defined by

$$\delta_i^j = 1 \text{ if } i = j , \quad \delta_i^j = 0 \text{ if } i \neq j \quad (6.2)$$

Sometimes the Kronecker symbol is also written as δ_{ij} or δ^{ij} , which has the same definition (in this course).

If i, j, k label the three dimensions of space, $i, j, k = 1, 2, 3$, there is also the ϵ_{ijk} -symbol. This is defined to be antisymmetric under interchanging any two adjacent indices: $\epsilon_{ijk} = -\epsilon_{jik} = -\epsilon_{ikj} = \dots$. It is then fixed by the condition that $\epsilon_{123} = 1$. One can then show that

$$\sum_{k=1}^3 \epsilon_{ijk} \epsilon_{lmk} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl} \quad (6.3)$$

and hence

$$\sum_{j,k=1}^3 \epsilon_{ijk} \epsilon_{ljk} = 2\delta_{il} . \quad (6.4)$$

Once again ϵ_{ijk} can appear with some or all of its indices upstairs with the same definition (in this course).

Appendix B: some numerical values

Planck originally introduced the constant h although it is $\hbar \equiv h/2\pi$ that is almost exclusively used in practice. The experimental values are approximately

$$h \sim 6.63 \times 10^{-27} \text{erg} \cdot \text{s} \quad \hbar \sim 1.05 \times 10^{-27} \text{erg} \cdot \text{s} \quad (6.1)$$

where s denotes seconds and erg denotes ergs. In atomic and particle physics more common units of energy are the electron Volt (eV) and Joule (J), which are related to ergs by

$$1\text{eV} = 1.6 \times 10^{-19} \text{J} = 1.6 \times 10^{-12} \text{ergs} \quad (6.2)$$

The speed of light is approximately

$$c \sim 3 \times 10^{10} \text{cm/s} \quad (6.3)$$

where cm denotes centimeters. The charge on an electron is

$$e \sim -4.8 \times 10^{-10} \text{esu} \quad (6.4)$$

where esu is an electrostatic unit, $\text{esu} = \sqrt{\text{erg} \cdot \text{cm}}$. In these units the force between two electrons at a distance r is simply e^2/r^2 .

It is important to note that a fundamental constant that is not dimensionless really reflects some kind of unit of measurement. That is to say that the numerical value of \hbar and c are not in themselves meaningful since they would change if we should choose to measure time in hours or energy in eV. In fact one can simply choose a system of measurement so that $\hbar = c = 1$. This is called "natural units". In it time is measured in the same units as space (since $c = 1$). In effect one unit of space is the distance light travels in one unit of time. Action (that is energy times time) also has no unit (since $\hbar = 1$), so we see that energy is measured in terms of inverse units of space. This is perhaps less easy to visualize but it reflects the fact that to probe shorter and shorter distances requires performing experiments at higher and higher momentum, which in turn is a result of the Heisenberg uncertainty relation $\Delta x \Delta p \geq \hbar/2$.

From the constants listed above one can form a dimensionless number known as the fine structure constant

$$\alpha = \frac{e^2}{\hbar c} \sim 1/137 \quad (6.5)$$

Of course the value of $1/137$ is only approximate, it is not exactly the inverse of an integer¹. The fine structure constant naturally appears in equations resulting from considering the electromagnetic interactions of charged particles. For example (*see the problem sets*) we saw that the typical speed of an electron in a Bohr orbit was about αc and hence is much less than the speed of light. Since it is dimensionless it is physically meaningful that $\alpha \ll 1$ and this tells us that the electromagnetic interaction between two electrons is not very powerful.

On the other hand you will often read that classical mechanics arises as the limit $\hbar \rightarrow 0$. Given the above comments, this is a meaningless statement (especially so in natural units where $\hbar = 1$). What is really meant is that if the *ratio* of the action of a physical process to \hbar is large, then it can be well described by classical physics.

Appendix C: classical mechanics in the Hamiltonian formalism

Consider the classical mechanics of a particle with a coordinate $\vec{x} = \{x^i\}$ and momentum $\vec{p} = \{p_i\}$, $i = 1, 2, 3$. The space labeled by \vec{x} and \vec{p} is known as phase space. The Hamil-

¹It also turns out that in relativistic quantum theory, so-called quantum electrodynamics (QED), α is not really a constant but depends on the energy scale of the experiment.

tonian formulation takes as its starting point a function $H(\vec{x}, \vec{p}, t)$, the Hamiltonian. The dynamics are described by Hamilton's equations

$$\frac{d\vec{x}}{dt} = \frac{\partial H}{\partial \vec{p}}, \quad \frac{d\vec{p}}{dt} = -\frac{\partial H}{\partial \vec{x}} \quad (6.1)$$

Often the Hamiltonian takes the form

$$H(\vec{x}, \vec{p}) = \frac{p^2}{2m} + V(\vec{x}). \quad (6.2)$$

It is easy to see that Hamilton's equations are

$$\frac{d\vec{x}}{dt} = \frac{\vec{p}}{m}, \quad \frac{d\vec{p}}{dt} = -\frac{\partial V}{\partial \vec{x}}. \quad (6.3)$$

Here the first equation is the familiar expression for the momenta in terms of the velocity

$$\vec{p} = m \frac{d\vec{x}}{dt}. \quad (6.4)$$

Using this the second equation becomes the second order force equation of Newtonian mechanics

$$m \frac{d^2 \vec{x}}{dt^2} = -\frac{\partial V}{\partial \vec{x}} \quad (6.5)$$

The system is therefore described by giving the positions and momenta of all the coordinates as a function of time. Furthermore since Hamilton's equations are first order in time, once the position and momenta are given at some initial time, the dynamics are uniquely determined for all time. It is implicit in classical mechanics that $\vec{x}(t)$ and $\vec{p}(t)$ are continuous variables and can in principle take on any real values (although they may of course be restricted to certain open domains of \mathbb{R}^3 in some cases). In particular it is possible to simultaneously know $\vec{x}(t)$ and $\vec{p}(t)$ to an arbitrary accuracy.

It is useful to introduce the Poisson bracket

$$\{A, B\} \equiv \frac{\partial A}{\partial \vec{x}} \cdot \frac{\partial B}{\partial \vec{p}} - \frac{\partial A}{\partial \vec{p}} \cdot \frac{\partial B}{\partial \vec{x}} \quad (6.6)$$

where A and B are any two function on phase space, which may also have an explicit time dependence. The Poisson bracket has some useful properties:

$$\begin{aligned} \{A, B\} &= -(B, A) \\ \{A, \lambda B + \mu C\} &= \lambda \{A, B\} + \mu \{A, C\} \\ \{AB, C\} &= A \{B, C\} + \{A, C\} B \end{aligned} \quad (6.7)$$

which can be checked explicitly (λ and μ are real numbers). For example one finds that

$$\begin{aligned} \{x_i, x_j\} &= 0 \\ \{p^i, p^j\} &= 0 \\ \{x_i, p^j\} &= \delta_i^j \end{aligned} \quad (6.8)$$

With this definition the time evolution of any function $f(\vec{x}, \vec{p}, t)$ is given by

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \{f, H\} . \quad (6.9)$$

To see this we simply expand out the right hand side

$$\begin{aligned} \frac{\partial f}{\partial t} + \{f, H\} &= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \vec{x}} \cdot \frac{\partial H}{\partial \vec{p}} - \frac{\partial f}{\partial \vec{p}} \cdot \frac{\partial H}{\partial \vec{x}} \\ &= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \vec{x}} \cdot \frac{d\vec{x}}{dt} + \frac{\partial f}{\partial \vec{p}} \cdot \frac{d\vec{p}}{dt} \\ &= \frac{df}{dt} \end{aligned} \quad (6.10)$$

where we used Hamilton's equations in the second line.

Note that

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} \quad (6.11)$$

so that if H does not have an explicit time dependence, it's value along any solution is constant and is interpreted as the energy, E . It is easy to see that any function Q , which does not depend explicitly on time, will be constant along solutions provided that $\{Q, H\} = 0$.

References

- [1] A. Messiah, *Quantum Mechanics, Volume I*, Dover (1999).
- [2] C. Cohen-Tanouji, B. Diu and F. Lalöe ,*Quantum Mechanics, Volumes I and II*,
- [3] R.P. Feynman, R.B. Leighton and M. Sands, *The Feynman Lectures on Physics, Volume III*, Addison-Wesley, (1966).

There are many good books on quantum mechanics. This course is not primarily based on any particular one. However the books listed above maybe particularly useful. A good way to learn quantum mechanics is simply to do as many problems as possible so it is worthwhile having a look at these books and others.