Introductory Quantum Theory Notes

Francesco Chotuck

Abstract

This is KCL undergraduate module 5CCM234A, instructed by Dr Petr Kravchuk. The formal name for this class is "Introductory Quantum Theory".

Contents

1 The wave function

1.1 The Schrödinger equation

Definition 1.1. The 'position' of a particle is determined by the complex-valued function $\Psi(x,t)$, called the **wave function**.

> **Note 1.2.** We can think of $\Psi(x,t)$ as the analogous to the position of a particle in classical dynamics given by $x(t)$.

Theorem 1.3. Since $\Psi \in \mathbb{C}$ we can write $|\Psi|^2 = \Psi^*\Psi$ where Ψ^* is the complex conjugate of Ψ.

Definition 1.4. We define a fundamental quantity

$$\hbar = \frac{h}{2\pi}$$

where h is the Planck's constant. It has units $\mathrm{J\,s}$

> **Theorem 1.5**
>
> We obtain the wave function from solving the (time-dependent) **Schrödinger equation**:
> $$i\hbar\frac{\partial}{\partial t}\Psi(x,t) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x,t) + U(x)\Psi(x,t),$$
> where $U(x)$ is the potential energy function of the particle.

> **Remark 1.6.** The equation above is for a particle moving in one direction (the x-axis) but this holds for a particle moving in $\mathbb{R}^3$ by changing $x \mapsto \boldsymbol{x}$ (or sometimes written $\vec{x}$).

> **Note 1.7.** Recall the total energy of a particle in motion in a dynamical system is given by
> $$E = \frac{p^2}{2m} + U(x)$$
> Therefore, we can think of the Schrödinger equation as describing the total energy of the system:
> $$\underbrace{i\hbar\frac{\partial}{\partial t}\Psi(x,t)}_{E\Psi} = \underbrace{-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x,t)}_{\frac{p^2}{2m}\Psi} + U(x)\Psi(x,t),$$

> **Remark 1.8.** If the particle in motion does not have any potential energy the Schrödinger equation can be adjusted by removing the 'potential energy' term i.e. the $U(x)$ term.

1.2 Statistical interpretation of the wave function

> **Theorem 1.9**
>
> The **probability density** of a wave function is given by
>
> $|\Psi(x,t)|^2 \, dx = \{$probability of finding the particle between x and $(x+dx)$, at time $t\}$
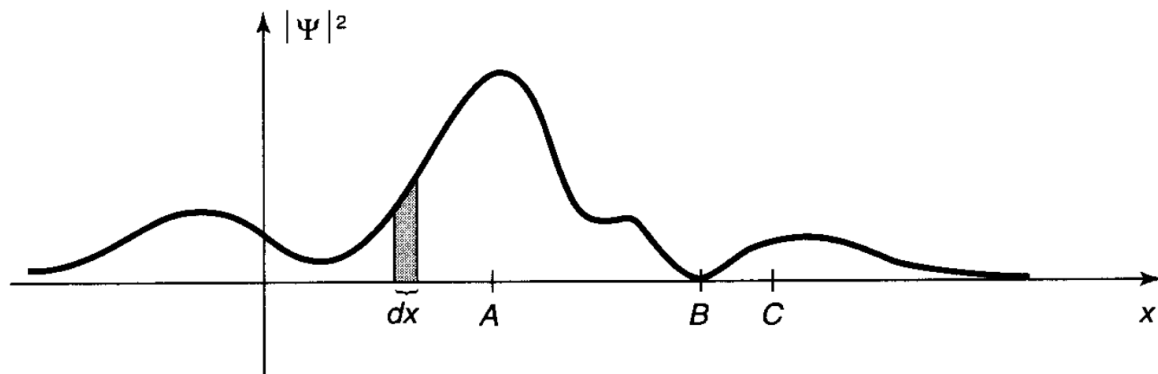
Suppose the graph a wave function is as the one above. The particle would be relatively likely to be found near A, and unlikely to be found near B. The shaded area represents the probability of finding the particle in the range dx.

1.2.1 Normalisation

> **Theorem 1.10**
>
> $$\int_{-\infty}^{\infty} |\Psi(x,t)|^2 = 1.$$

> **Remark 1.11.** This is a required condition otherwise all the statistical interpretation of the wave function would be nonsense.

Corollary 1.12. The value

$$\int_{a}^{b} |\Psi(x,t)|^2 \, dx = \{\text{probability of finding the particle between } a \text{ and } b, \text{ at time } t\}$$

Definition 1.13. The process of altering the wave function such that it satisfies the theorem above is called **normalisation** of the wave function.

> **Remark 1.14.** If Ψ is a solution to the Schrödinger equation then so is $A\Psi$ where $A \in \mathbb{C}$ is a scalar. Thus, to normalise a wave function we multiply by a complex scalar which makes the integral equal to 1.

Definition 1.15. Solution to the Schrödinger equation with integral of $|\Psi|^2$ is infinite or when $\Psi = 0$ are said to be **non-normalisable** solutions.

> **Note 1.16.** Physically these solutions cannot exist so, they must be rejected.

Theorem 1.17

If the integral

$$\int_{-\infty}^{\infty} |\Psi(x,t)|^2 \, dx$$

is finite then the wave function Ψ can be normalised.

1.3   Operators and expected values

Definition 1.18. A linear operator $\widehat{\mathcal{O}}$ is a rule that takes a wave function Ψ and produces a new wave function that we denote $\widehat{\mathcal{O}}\Psi$. Linearity means that it satisfies the property

$$\widehat{\mathcal{O}}\left(\alpha\Psi_1 + \beta\Psi_2\right) = \alpha\widehat{\mathcal{O}}\Psi_1 + \beta\widehat{\mathcal{O}}\Psi_2.$$

> **Remark 1.19.** Linear operators in this module are defined $\widehat{\mathcal{O}} : V \to V$ where V is the vector space of complex-valued functions.

Theorem 1.20. The notation $\widehat{\mathcal{O}}_1\widehat{\mathcal{O}}_2$ means composition of linear operators i.e. $(\widehat{\mathcal{O}}_1\widehat{\mathcal{O}}_2)(x) = \widehat{\mathcal{O}}_1(\widehat{\mathcal{O}}_2(x))$.

Definition 1.21. The linear **operator of position**, $\widehat{x}$, for a fixed time t is defined as

$$(\widehat{x}\,\Psi)(x) = x\Psi(x).$$

> **Remark 1.22.** This is terrible notation. The operator, $\widehat{x}$, is acting on the function Ψ and yields a new function $(\widehat{x}\,\Psi)$. We can think of $(\widehat{x}\,\Psi)$ as a function f. Therefore, f acting on x, written $f(x)$, in this notation is written as $(\widehat{x}\,\Psi)(x)$.

> **Note 1.23.** Think of it as $\widehat{\mathcal{O}}(\text{function}) = \text{function}$.

Definition 1.24. The linear **operator of momentum**, $\widehat{p}$, for a fixed time t is defined as

$$(\widehat{p}\,\Psi)(x) = -i\hbar\frac{\partial}{\partial x}\Psi(x).$$

Theorem 1.25

If Ψ is normalised then the expectation value of a general linear operator $\widehat{\mathcal{O}}$, at a fixed time t is given by

$$\langle\widehat{\mathcal{O}}\rangle = \int_{-\infty}^{\infty} \Psi^*(x)(\widehat{\mathcal{O}}\Psi)(x)\, dx.$$

Definition 1.26. For a particle in a normalised state Ψ, the **expectation value of the position**, x, for a fixed time t is

$$\langle \widehat{x} \rangle = \int_{-\infty}^{\infty} x |\Psi(x)|^2 \, dx$$
$$= \int_{-\infty}^{\infty} \Psi^*(x) \, x \, \Psi(x) \, dx.$$

Definition 1.27. For a particle in a normalised state Ψ, the **expectation value the momentum**, p, for a fixed time t is

$$\langle \widehat{p} \rangle = \int_{-\infty}^{\infty} \Psi^*(x) \left(-i\hbar \frac{\partial}{\partial x} \Psi(x) \right) dx$$
$$= -i\hbar \int_{-\infty}^{\infty} \Psi^*(x) \left(\frac{\partial}{\partial x} \Psi(x) \right) dx.$$

Theorem 1.28

For a general operator $\widehat{\mathcal{O}}$ acting on $\widehat{x}$ we have

$$(\widehat{\mathcal{O}}(\widehat{x})\Psi)(y) = \widehat{\mathcal{O}}(y)\Psi(y).$$

1.4   The Hamiltonian operator

Definition 1.29. The **Hamiltonian operator** is defined as

$$\widehat{H} = \frac{\widehat{p}^2}{2m} + U(\widehat{x}),$$

where $U(\widehat{x})$ is defined as

$$(U(\widehat{x})\Psi)(y) = U(y)\Psi(y).$$

Theorem 1.30

The Schrödinger equation with the Hamiltonian operator can be written as

$$i\hbar \frac{\partial}{\partial t} \Psi(x,t) = \widehat{H}\Psi(x,t).$$

Definition 1.31. An operator $\widehat{\mathcal{O}}$ for which

$$\int_{-\infty}^{\infty} \Psi^*(x)(\widehat{\mathcal{O}}\Phi)(x) \, dx = \int_{-\infty}^{\infty} \left[(\widehat{\mathcal{O}}\Psi)(x) \right]^* \Phi(x) \, dx,$$

for any Φ and Ψ is said to be **Hermitian**.

Theorem 1.32

The Hamiltonian operator is Hermitian.

Proof. We will be using the following facts for this proof:

1. The sum of Hermitian operators is Hermitian.

2. The product of Hermitian operators is Hermitian.

3. A Hermitian operator multiplied by a scalar $\alpha \in \mathbb{R}$ is Hermitian.

Using these facts, we have to prove that $\widehat{p}$ and $U(\widehat{x})$ are Hermitian operators.

- Proof of $U(\widehat{x})$ being Hermitian.
  Suppose $U(\widehat{x})$ is Hermitian then

$$\int_{-\infty}^{\infty} \Psi^*(x)[U(x)\Phi(x)]\,dx = \int_{-\infty}^{\infty} [U(x)\Psi(x)]^*\Phi(x)\,dx.$$

This can only happen if $U(x) = U^*(x)$ that is, $U(x) \in \mathbb{R}$. In this course the potential is always real.

- Proof of $\widehat{p}$ being Hermitian.
  Suppose $\widehat{p}$ is Hermitian, we check that the condition holds.

$$\begin{aligned}
\int_{-\infty}^{\infty} [\widehat{p}\,\Psi(x)]^*\Phi(x) &= \int_{-\infty}^{\infty} \left(-i\hbar\frac{\partial}{\partial x}\Psi(x)\right)^* \Phi(x) \\
&= i\hbar \int_{-\infty}^{\infty} \left(\frac{\partial}{\partial x}\Psi^*(x)\right)\Phi(x) \\
&= -i\hbar \int_{-\infty}^{\infty} \Psi^*(x)\frac{\partial}{\partial x}\Phi(x) \quad \text{(integration by parts)} \\
&= \int_{-\infty}^{\infty} \Psi^*(x)\,\widehat{p}\,\Phi(x).
\end{aligned}$$

$\square$

1.5   Conservation of probability

> **Theorem 1.33**
> The Schrödinger equation preserves the normalisation of the wave function.

> **Note 1.34.** Suppose we have normalised a function at $t = 0$, as time goes on and Ψ evolves we need to ensure that Ψ remains normalised; luckily the Schrödinger equation has the property illustrated above

Proof. If the integral $\int_{-\infty}^{\infty} |\Psi(x,t)|^2 = 1$ at one moment in time then it is 1 for all other t. Thus, we want to prove that

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{-\infty}^{\infty} |\Psi(x,t)|^2 = 0.$$

We have that

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \int_{-\infty}^{\infty} |\Psi(x,t)|^2 \, dx &= \int_{-\infty}^{\infty} \frac{\partial}{\partial t} \Psi(x,t) \Psi^*(x,t) \, dx \\
&= \int_{-\infty}^{\infty} \Psi^*(x,t) \frac{\partial}{\partial t} \Psi(x,t) \, dx + \int_{-\infty}^{\infty} \Psi(x,t) \frac{\partial}{\partial t} \Psi^*(x,t) \, dx \\
&= \int_{-\infty}^{\infty} \Psi^*(x,t) \frac{1}{i\hbar} (\widehat{H}\Psi)(x,t) \, dx - \int_{-\infty}^{\infty} \frac{1}{i\hbar} [(\widehat{H}\Psi)(x,t)]^* \Psi(x,t) \, dx \\
&= \frac{1}{i\hbar} \left[\int_{-\infty}^{\infty} \Psi^*(x,t)(\widehat{H}\Psi)(x,t) \, dx - \int_{-\infty}^{\infty} [(\widehat{H}\Psi)(x,t)]^* \Psi(x,t) \, dx \right] \\
&= 0.
\end{aligned}
$$

Since $\widehat{H}$ is Hermitian the last two integrals are equal and hence, cancel each other out. $\qquad\square$

1.6   Evolution in time of expectation values

Definition 1.35. A **commutator** is defined as

$$
[\widehat{A}, \widehat{B}] = \widehat{A}\widehat{B} - \widehat{B}\widehat{A}.
$$

Theorem 1.36. Properties of commutator. For any $\alpha, \beta \in \mathbb{C}$ and any operator we have:

- $\left[\alpha\widehat{A} + \beta\widehat{B}, \widehat{C} \right] = \alpha \left[\widehat{A}, \widehat{C} \right] + \beta \left[\widehat{B}, \widehat{C} \right]$.

- $\left[\widehat{A}, \alpha\widehat{B} + \beta\widehat{C} \right] = \alpha \left[\widehat{A}, \widehat{B} \right] + \beta \left[\widehat{A}, \widehat{C} \right]$.

- $\left[\widehat{A}, \widehat{B} \right] = - \left[\widehat{B}, \widehat{A} \right]$.

> **Note 1.37.** The first two properties imply that the commutator is bilinear.

Theorem 1.38

For any observable $\widehat{\mathcal{O}}$, the time derivative of its expectation value is given by

$$
\frac{\partial}{\partial t} \langle \widehat{\mathcal{O}} \rangle = \frac{i}{\hbar} \left\langle \left[\widehat{H}, \widehat{\mathcal{O}} \right] \right\rangle,
$$

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

Remark 1.39. An equivalent restatement is

$$
\frac{\partial}{\partial t} \langle \widehat{\mathcal{O}} \rangle = \frac{1}{i\hbar} \left\langle \left[\widehat{\mathcal{O}}, \widehat{H} \right] \right\rangle.
$$

Proof. Recall,

$$
\langle \widehat{\mathcal{O}} \rangle = \int_{-\infty}^{\infty} \Psi^*(x)(\widehat{\mathcal{O}}\Psi)(x).
$$

Taking its time derivative we have,

$$\frac{\partial}{\partial t}\langle\widehat{\mathcal{O}}\rangle = \int_{-\infty}^{\infty} \frac{\partial}{\partial t}\left[\Psi^*(x)(\widehat{\mathcal{O}}\Psi)(x)\right]dx$$

$$= \int_{-\infty}^{\infty}\left[\frac{\partial}{\partial t}\Psi(x)\right]^*(\widehat{\mathcal{O}}\Psi)(x)\,dx + \int_{-\infty}^{\infty}\Psi^*(x)\left(\widehat{\mathcal{O}}\frac{\partial}{\partial t}\Psi\right)(x)\,dx$$

$$= \frac{1}{\hbar}\left(\int_{-\infty}^{\infty}\left[\widehat{H}\Psi(x)\right]^*(\widehat{\mathcal{O}}\Psi)(x)\,dx - \int_{-\infty}^{\infty}\Psi^*(x)(\widehat{\mathcal{O}}\widehat{H}\Psi)(x)\right)dx.$$

Since $\widehat{H}$ is Hermitian we can write (with $\Psi = \Psi$ and $\Phi = \widehat{\mathcal{O}}\Psi$),

$$\int_{-\infty}^{\infty}\left[\widehat{H}\Psi(x)\right]^*\widehat{\mathcal{O}}\Psi(x)\,dx = \int_{-\infty}^{\infty}\Psi^*(x)(\widehat{H}\widehat{\mathcal{O}}\Psi)(x)\,dx.$$

Plugging this into the above we obtain

$$\frac{\partial}{\partial t}\langle\widehat{\mathcal{O}}\rangle = \frac{i}{\hbar}\left(\int_{-\infty}^{\infty}\Psi^*(x)\left(\widehat{H}\widehat{\mathcal{O}}\right)\Psi(x)\,dx - \int_{-\infty}^{\infty}\Psi^*(x)\left(\widehat{\mathcal{O}}\widehat{H}\right)\Psi(x)\right)$$

$$= \frac{i}{\hbar}\left(\int_{-\infty}^{\infty}\Psi^*(x)(\widehat{H}\widehat{\mathcal{O}} - \widehat{\mathcal{O}}\widehat{H})\Psi(x)\right)dx$$

$$= \frac{i}{\hbar}\left\langle\widehat{H}\widehat{\mathcal{O}} - \widehat{\mathcal{O}}\widehat{H}\right\rangle$$

$$= \frac{i}{\hbar}\left[\widehat{H},\widehat{\mathcal{O}}\right].$$

$\square$

Theorem 1.40 (Ehrenfest theorem)

We have the following relation

$$\frac{\partial}{\partial t}\langle\widehat{x}\rangle = \frac{\langle\widehat{p}\rangle}{m},$$

$$\frac{\partial}{\partial t}\langle\widehat{p}\rangle = -\left\langle\frac{\partial}{\partial x}U(\widehat{x})\right\rangle.$$

Remark 1.41. Consider the Classical Dynamics form of the statement above:

$$\frac{\partial}{\partial t}x(t) = \frac{p(t)}{m}$$

$$\frac{\partial}{\partial t}p(t) = -\frac{\partial}{\partial t}U(x(t)).$$

The Quantum and Classical statements are equivalent if $p(t) = \langle\widehat{p}\rangle$ and $x(t) = \langle\widehat{x}\rangle$. This is clearly not the case: in Quantum we have

$$\frac{\partial}{\partial t}\langle\widehat{p}\rangle = \langle F(\widehat{x})\rangle,$$

whereas, in Classical we have

$$\frac{\partial}{\partial t}\langle\widehat{p}\rangle = F(\langle\widehat{x}\rangle).$$

Therefore, $\langle F(\widehat{x}) \rangle \neq F(\langle \widehat{x} \rangle)$ always; in only certain cases the equality is achieved.

2 Simple quantum mechanical systems

2.1 Separation of variables

Definition 2.1. Solutions to the Schrödinger equation which are of the form

$$\Psi(x,t) = \psi(x)\Phi(t)$$

are called **separable solutions**.

Recall the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \Psi(x,t) = \widehat{H} \Psi(x,t).$$

Assuming $\Psi(x,t)$ is separable we substitute to obtain

$$i\hbar \psi(x) \frac{\partial \Phi(t)}{\partial t} = \Phi(t)(\widehat{H}\psi)(x)$$

$$\underbrace{\frac{i\hbar}{\Phi(t)} \frac{\partial \Phi(t)}{\partial t}}_{t \text{ dependence}} = \underbrace{\frac{(\widehat{H}\psi)(x)}{\psi(x)}}_{x \text{ dependence}}.$$

Since, the LHS is only dependent on t and the RHS is only dependent on x we must have that both sides are equal to a constant – otherwise by varying t, we could change the LHS without touching the RHS and the two would no longer be equal. So, we set

$$\frac{i\hbar}{\Phi(t)} \frac{\partial \Phi(t)}{\partial t} = \frac{(\widehat{H}\psi)(x)}{\psi(x)} = E \in \mathbb{R},$$

i.e. the energy.

Analysing the LHS:

$$i\hbar \frac{\partial \Phi(t)}{\partial t} = E\Phi(t)$$

$$\Rightarrow \frac{\partial \Phi(t)}{\partial t} = -\frac{iE}{\hbar} \Phi(t)$$

$$\Rightarrow \Phi(t) = A \exp\left(-\frac{iE}{\hbar}t\right)$$

$$= Ae^{-\frac{iE}{\hbar}t}.$$

Remark 2.2. In some texts the partial derivatives become total derivatives as the wave function $\Psi(x,t)$ is no longer multivariable. This allows us to treat the PDEs as ODEs and solve them with techniques we learnt previously.

Analysing the RHS we have the following theorem.

> **Theorem 2.3**
>
> The **time-independent** Schrödinger equation is given by
>
> $$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) + U(x)\psi(x) = \psi(x)E.$$
>
> Equivalently, it can be written as
>
> $$(\widehat{H}\psi)(x) = E\psi(x).$$

> **Note 2.4.** Due to separation of variables the partial derivative above can be written as a total derivative and use standard techniques to solve a second order ODE.

The time-dependent Schrödinger equation can be solved for a given potential thus, leading to the following.

> **Corollary 2.5**
>
> The general solution of the Schrödinger equation is given by
>
> $$\Psi(x,t) = \psi(x)e^{-\frac{iE}{\hbar}t}.$$

> **Remark 2.6.** The constant A from $\Phi(t)$ is incorporated into the normalisation constant of $\Psi(x,t)$.

2.2   Particle in a box

Definition 2.7. Suppose

$$U(x) = \begin{cases} 0 & \text{if } 0 \leq x \leq L \\ \infty & \text{otherwise.} \end{cases}$$

We call this potential the **infinite square well** or a **particle moving a box**.
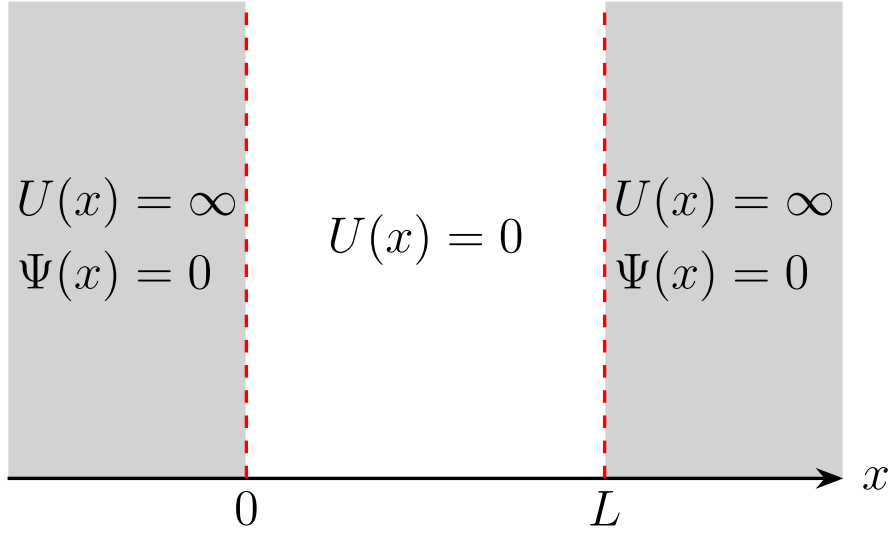
Figure 1: Figure of the infinite square well

Since $U(x) = 0$ inside the 'box' the Schrödinger equation becomes

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

this takes the form of **simple harmonic motion** thus, the solution is given by

$$\Psi(x) = A\cos(kx) + B\sin(kx) \quad \text{where } k \equiv \frac{\sqrt{2mE}}{\hbar},$$

where A and B are arbitrary constants. We assume $E > 0$ (since $E < 0$ does not work if tried) which forces $k > 0$. Continuity of the potential function requires that

$$\Psi(0) = \Psi(L) = 0,$$

which allows us to evaluate the arbitrary constants.

$$\Psi(0) = 0 \quad \Rightarrow A = 0$$
$$\Psi(L) = 0 \quad \Rightarrow B\sin(kL) = 0.$$

The second equation implies $B = 0$ or $\sin(kL) = 0$; if $B = 0$ then $\Psi(x) = 0$ which is not normalisable hence cannot be used. Therefore, $\sin(kL) = 0$ which means that

$$kL = n\pi \quad \text{for } n \in \mathbb{Z}.$$

However, $k = 0$ would again imply $\Psi(x) = 0$ and the negative solutions give nothing new, since $\sin(-\theta) = -\sin(\theta)$ and we can absorb the minus sign into B. Hence, $n \in \mathbb{Z}_{\geq 0}$ and the distinct solutions are

$$k_n = \frac{\pi n}{L} \quad \text{for } n \in \mathbb{Z}_{\geq 0}.$$

Since, k_n depends on E we can work out the possible values of E:

$$\boxed{E_n = \frac{\pi^2 n^2 \hbar^2}{2mL^2}.}$$

12

> **Note 2.8.** The existence of E_n implies that 'quantum' particles have discrete energy levels.

Now, we need to find B, we do this by normalising the wave function $\Psi = B \sin(kx)$:

$$
\begin{aligned}
1 = \int_{-\infty}^{\infty} |\Psi(x)|^2 \, dx &= \int_{-\infty}^{\infty} |B|^2 \sin^2(kx) \, dx \\
&= \int_{-\infty}^{0} 0 \, dx + \int_{0}^{L} |B|^2 \sin^2(kx) \, dx + \int_{L}^{\infty} 0 \, dx \\
&= \int_{0}^{L} |B|^2 \sin^2(kx) \, dx \\
&= |B|^2 \frac{L}{2}.
\end{aligned}
$$

We conclude $B = \sqrt{\frac{2}{L}}$. The wave function for a particle under the influence of this potential is given by

$$
\boxed{\Psi_n(x,t) = \sqrt{\frac{2}{L}} \sin\left(\frac{\pi n}{L} x\right) \exp\left(\frac{-i E_n t}{\hbar}\right)}
$$

$$
= \sqrt{\frac{2}{L}} \sin\left(\frac{\pi n}{L} x\right) e^{-i \frac{E_n}{\hbar} t}.
$$

> **Remark 2.9.** Since $B \in \mathbb{C}$ we can be $B = \sqrt{\frac{2}{L}} e^{i\theta}$, which we call the phase of the wave function, but we set this to 1.

2.2.1 Superposition of states

Definition 2.10. Superposition is the term physicists use for a linear combination or the continuous analogue thereof, namely an integral.

> **Remark 2.11.** This is due to the properties of the Schrödinger equation being a differential equation: any linear combination of some solutions is also a solution.

2.3 Potential well

Definition 2.12. Suppose

$$
U(x) = \begin{cases} -U_0 & \text{if } -\frac{L}{2} \leq x \leq \frac{L}{2} \\ 0 & \text{if } |x| > \frac{L}{2}, \end{cases}
$$

where $U_0 > 0$. We call this potential the **finite square well** or the **potential well**.
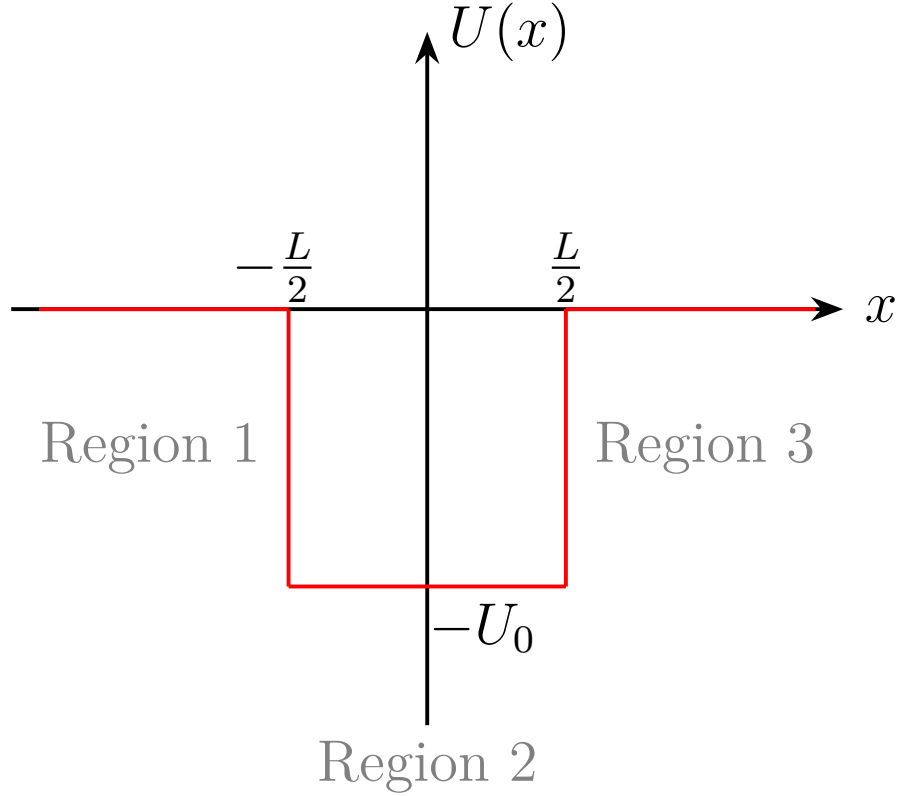
Figure 2: A diagram of the potential well

We split the potential well into three regions as shown by the Figure above. The Schrödinger equation for Region 1 and Region 3 is the time-independent equation with no potential term i.e.

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

where $E < 0$. Therefore,

$$\text{Region 1}: \quad \Psi(x) = A_1 e^{kx} + B_1 e^{-kx}$$
$$\text{Region 3}: \quad \Psi(x) = A_3 e^{kx} + B_3 e^{-kx}$$

for

$$k \equiv \frac{\sqrt{-2mE}}{\hbar} > 0.$$

Remark 2.13. We use Euler's formula to write the trigonometric functions in terms of exponentials.

Since $\Psi(x)$ must be normalisable to be an appropriate solution we have that the terms $A_3 = 0$ and $B_1 = 0$ otherwise, when taking the integral over from $-\infty$ to ∞ these terms would diverge. Region 2 is the only region with a potential thus, the Schrödinger equation becomes:

$$-\frac{\hbar}{2m}\frac{\partial^2\Psi(x)}{\partial x^2} - U_0\Psi(x) = E\Psi(x)$$
$$-\frac{\hbar}{2m}\frac{\partial^2\Psi(x)}{\partial x^2} = (E + U_0)\Psi(x).$$

14

Although, $E < 0$ we must have $E \geq U_0$ which implies, $E + U_0 \geq 0$. The general solution is

$$\Psi(x) = A_2 \cos(k_2 x) + B_2 \sin(k_2 x)$$

where

$$k_2 \equiv \frac{\sqrt{2m(E + U_0)}}{\hbar}.$$

The next step is to impose boundary conditions: Ψ and $\dfrac{\partial \Psi(x)}{\partial x}$ must be continuous at $-\frac{L}{2}$ and $\frac{L}{2}$. Note that the potential function is an even function $(U(x) = U(-x))$ i.e. symmetric about the line $y = x$. This implies: if $\Psi(x)$ is a solution of the Schrödinger equation, then $\Psi(-x)$ is a solution. Furthermore, the Schrödinger equation is linear and homogeneous in Ψ i.e. we can add and subtract solutions thus, we define

$$\Psi_+(x) := \Psi(x) + \Psi(-x)$$
$$\Psi_-(x) := \Psi(x) - \Psi(-x),$$

which are also solutions. These new solutions have the property that the original solution can be recovered from them,

$$\Psi(x) = \frac{\Psi_+(x) + \Psi_-(x)}{2}$$

and they are even and odd respectively.

$$\Psi_\pm(x) = \pm \Psi(x).$$

The even case

Suppose that the solution we are after, $\Psi(x)$, is even. Using the solutions to Region 1 and Region 2 and applying continuity:

$$A_1 e^{-k\frac{L}{2}} = A_2 \cos\left(\frac{k_2 L}{2}\right) + B_2 \sin\left(\frac{-k_2 L}{2}\right)$$
$$A_1 k e^{-k\frac{L}{2}} = A_2 k_2 \sin\left(\frac{k_2 L}{2}\right) + B_2 k_2 \cos\left(\frac{k_2 L_2}{2}\right).$$

Imposing that $\Psi(x)$ is even we must have $B_2 = 0$ so, the equations above become

$$A_1 e^{-k\frac{L}{2}} = A_2 \cos\left(\frac{k_2 L}{2}\right)$$
$$A_1 k e^{-k\frac{L}{2}} = A_2 k_2 \sin\left(\frac{k_2 L}{2}\right).$$

Taking the ration of these two equations, we find

$$k = k_2 \tan\left(\frac{k_2 L}{2}\right).$$

Expressing k in terms of k_2: recall that

$$k = \sqrt{\frac{-2mE}{\hbar^2}} \quad \text{and} \quad k_2 = \sqrt{\frac{2m(U_0 + E)}{\hbar^2}}$$

$$\Rightarrow k = \sqrt{\frac{2mU_0}{\hbar^2} - k_2^2}.$$

Since, $E < 0$ it implies that $k \in \mathbb{R}$ and that

$$0 \leq k_2 \leq \sqrt{\frac{2mU_0}{\hbar^2}}.$$

Finally, we have an expression in k_2 only,

$$\sqrt{\frac{2mU_0}{\hbar^2} - k_2^2} = k_2 \tan\left(\frac{k_2 L}{2}\right).$$

This equation allows us to determine k_2 (and thus E) in terms of the parameters of the problem.

The odd case

Suppose now that the solution, $\Psi(x)$ is odd. Doing a similar process with the even case we have the ratio

$$\sqrt{\frac{2mU_0}{\hbar^2} - k_2^2} = k_2 \tan\left(\frac{k_2 L + \pi}{2}\right).$$

Graphical representation
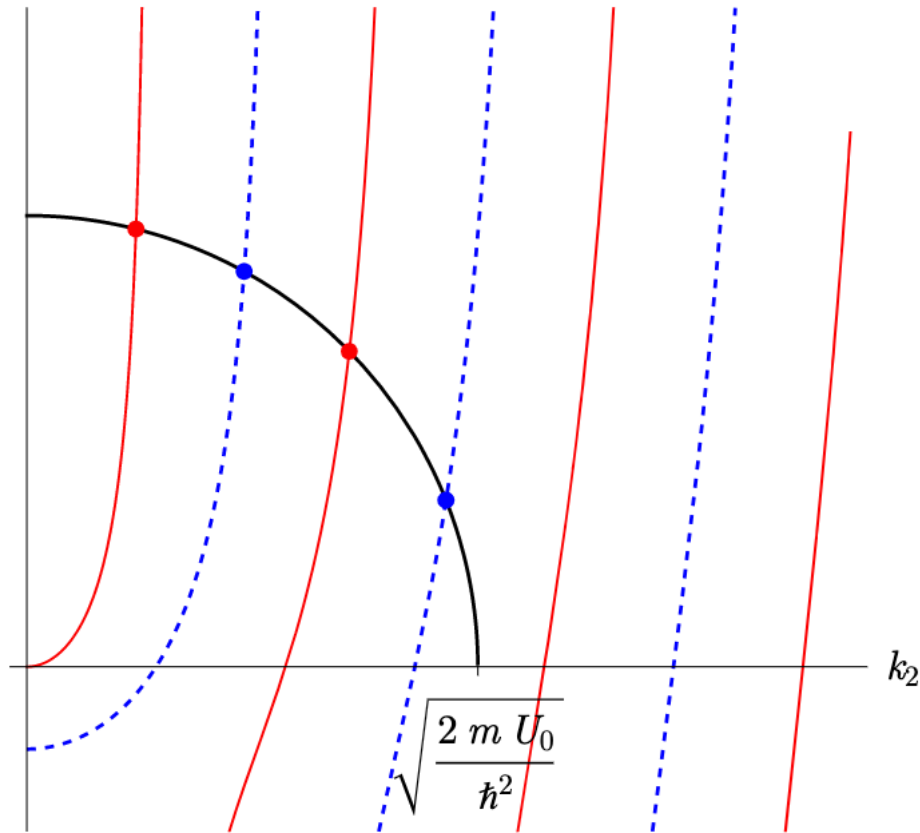
Figure 3: A graphical solution of the even and odd case. The LHS is plotted in solid black whereas, the RHS, if even, is plotted in solid red and if odd in dashed blue. The red and blue circle marks represent the solution to the even and odd case; the number of solutions depends on the parameters of the problem.

2.4   Quantum tunnelling and reflection

We now study the solutions to the problem in the previous section with $E > 0$. We define $V_0 = -U_0$ and imagine V_0 being positive (thus, U_0 being negative).

Definition 2.14. Suppose

$$U(x) = \begin{cases} V_0 & \text{if } -\frac{L}{2} \leq x \leq \frac{L}{2} \\ 0 & \text{if } |x| > \frac{L}{2} \end{cases}$$

where $V_0 > 0$. We call this potential the **potential barrier**.
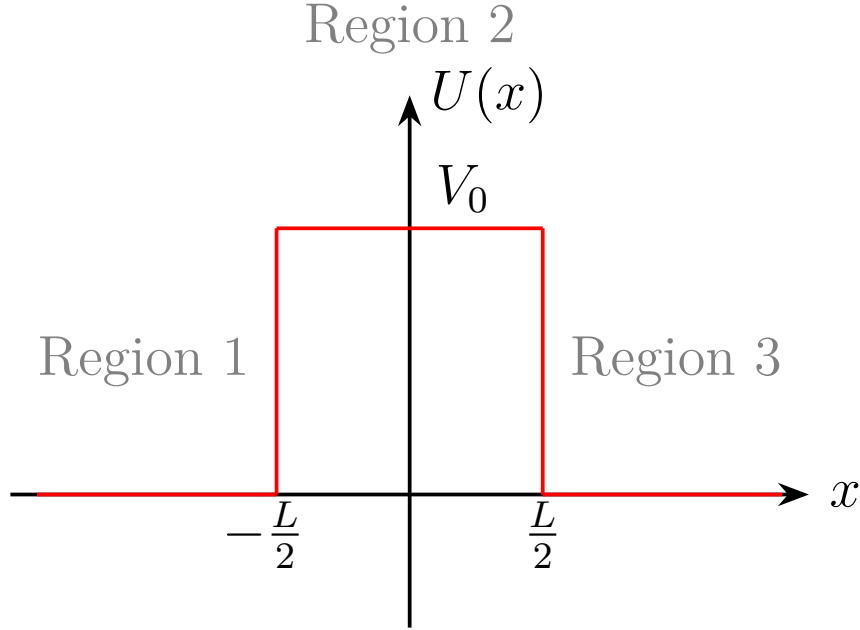
We solve the time-independent Schrödinger equation for each region, and we have

$$\Psi(x) = \begin{cases} A_+e^{\frac{ipx}{\hbar}} + A_-e^{-\frac{ipx}{\hbar}} & \text{for } x < -\frac{L}{2} \\ B_+e^{\frac{ip_2x}{\hbar}} + B_-e^{-\frac{ip_2x}{\hbar}} & \text{for } -\frac{L}{2} \le x \le \frac{L}{2} \\ C_+e^{\frac{ipx}{\hbar}} + C_-e^{-\frac{ipx}{\hbar}} & \text{for } x > \frac{L}{2}, \end{cases}$$

where p represents momentum and is defined as $p = \sqrt{2mE}$ whereas, p_2 is determined by $p_2 = \sqrt{2m(E - V_0)}$. Imposing the condition that $\Psi(x)$ must be continuous we can evaluate some constants, so we have

$$A_-(p) = \text{Stuff}(p) \cdot A_+$$
$$C_+(p) = \text{Stuff}(p) \cdot A_-.$$

However, with these constants the solution $\Psi(x)$ is not normalisable thus, we cannot fix A_+ with the normalisable condition; we choose $A_+ = 1$, and we have

$$A_-(p) = \frac{p^2 - p_2^2}{(p + p_2)^2} \cdot \frac{e^{\frac{-ipL}{\hbar}} - e^{\frac{i(2p_2 - p)L}{\hbar}}}{1 - \left(\frac{p - p_2}{p + p_2}\right)^2 e^{\frac{i2p_2L}{\hbar}}}$$

$$C_+(p) = \frac{4pp_2}{(p + p_2)^2} \cdot \frac{e^{\frac{i(p_2 - p)L}{\hbar}}}{1 - \left(\frac{p - p_2}{p + p_2}\right)^2 e^{\frac{i2p_2L}{\hbar}}}$$

We have written $A_-(p)$ and $C_+(p)$ to stress that these coefficients depend on the value of p which is connected E via $E = \frac{p^2}{2m}$.

In this thought experiment we want the particle coming from the left and travelling to the right. Analysing the solution of the Schrödinger equation:

- the term $A_+e^{\frac{ipx}{\hbar}}$ describes a particle moving to the right in Region 1, in the direction of the positive x axis towards the potential barrier in Region 2.

- The term $A_- e^{-\frac{ipx}{\hbar}}$ describes a particle moving away from the potential barrier, we can think of this as a **reflected particle**.

- The term $C_+ e^{\frac{ipx}{\hbar}}$ describes a particle travelling away from the potential barrier, and we can think of it as the **transmitted particle**.

- The term $C_- e^{-\frac{ipx}{\hbar}}$ describes a particle travelling from the right towards the potential barrier. As we set up the thought experiment for a particle coming from the left in Region 1 we set this term to 0 i.e. set $C_- = 0$.

Therefore, the full time-independent solution for a given p (equivalently E) is

$$\Psi_p(x,t) = e^{\frac{-i\frac{p^2}{2m}t}{\hbar}} \cdot \begin{cases} e^{\frac{ipx}{\hbar}} + A_- e^{-\frac{ipx}{\hbar}} & \text{for } x < -\frac{L}{2} \\ B_+ e^{\frac{ip_2 x}{\hbar}} + B_- e^{-\frac{ip_2 x}{\hbar}} & \text{for } -\frac{L}{2} \leq x \leq \frac{L}{2} \\ C_+ e^{\frac{ipx}{\hbar}} & \text{for } x > \frac{L}{2}. \end{cases}$$

2.4.1 Interpreting the scattering amplitudes

In the 'quantum realm' we deal with statistical interpretation of the happenings i.e. the state of the particle is not definite, until the particle is observed the particle is both transmitted and reflected by the quantum barrier.

Theorem 2.15

For a particle travelling in a potential barrier:

- the **transmission probability** is given by

$$T(p) = |C_+(p)|^2,$$

- the **reflection probability** is given by

$$R(p) = |A_-(p)|^2,$$

with the condition that

$$R(p) + T(p) = 1.$$

---

3   Hilbert spaces

3.1   Vector spaces

Definition 3.1. A **vector space** V, over a field K, is a non-empty set (the elements of which are called vectors), together with two binary operations $+ : V \times V \to V$ and $\cdot : K \times V \to V$ which we call *vector addition* and *scalar multiplication* respectively.

Definition 3.2. A vector space, V, is said to be **finite-dimensional** if there exists a finite set of vectors $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ such that any vector $\boldsymbol{v} \in V$ can be written as

$$\boldsymbol{v} = \alpha_1 \boldsymbol{e}_1 + \cdots + \alpha_n \boldsymbol{e}_n$$

for some $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$. Such a set is called **complete**.

> **Remark 3.3.** That is, a spanning set.

Definition 3.4. The smallest possible size n of a complete set is called the **dimension** of the vector space V and is denoted $\dim(V)$. If no complete finite set exists, V is called **infinite-dimensional**.

> **Remark 3.5.** For convenience, the dimension of the vector space $V = \{0\}$ is defined to be 0.

Theorem 3.6. In a finite dimensional vector space V the size of the linearly independent set cannot be larger than $\dim(V)$.

Corollary 3.7. Any linearly independent set with $\dim(V)$ elements is complete.

Definition 3.8. A finite set of vectors $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is called **linearly independent** is there exists $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$, for which the following holds

$$\alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_n \boldsymbol{v}_n = 0 \iff \alpha_1 = \cdots = \alpha_n = 0.$$

Otherwise, if not all α_i are zero then we call such set **linearly dependent**.

Example 3.9 (The vector space of all functions)

Let X be a set, and let $V = \mathbb{C}^X$ be the set of all functions from $X \to \mathbb{C}$. We define the addition and scalar multiplication for $f, f_1, f_2 \in V$ and $a \in \mathbb{C}$ as

$$(f_1 + f_2)(x) \equiv f_1(x) + f_2(x)$$
$$(af)(x) \equiv af(x).$$

If $X = \{1, 2, \ldots n\}$ we can identify V with the vector space $\mathbb{C}^n$. Whereas, if X has infinitely many elements, then V is not a finite-dimensional space.

Proof. Choose $m \in \mathbb{Z}_{\geq 0}$ and select m distinct elements $x_1, \ldots, x_m \in X$ (which is possible for any m since X has infinitely many elements). Define functions $f_1, \ldots, f_m$ so that $f_i(x) = 1$ if $x = x_i$ and $f_i(x) = 0$ otherwise. Consider the function

$$f = \alpha_1 f_1 + \cdots + \alpha_m f_m,$$

where $\alpha_i \in \mathbb{C}$ are arbitrary constants. By construction, $f(x_i) = \alpha_i$ therefore, if $f = 0$ then $\alpha_i = 0$ and thus $\{f_1, \ldots, f_m\}$ is a linearly independent set with m elements. Since, m was arbitrary, we deduce that V cannot have a finite dimension – otherwise m would have to be bounded by $\dim(V)$. $\qquad\square$

3.2 Inner products

Definition 3.10. A **norm** on V is a function $\|\cdot\| : V \to \mathbb{C}$ with the following properties:

- $\|\boldsymbol{v}\| \geq 0$ for all $\boldsymbol{v} \in V$,

- $\|\boldsymbol{v}\| = 0 \iff \boldsymbol{v} = 0$,

- **Homogeneity:** $\|\alpha \boldsymbol{v}\| = |\alpha| \|\boldsymbol{v}\|$ for all $\alpha \in \mathbb{C}$ and $\boldsymbol{v} \in V$,

- **Triangle inequality:** $\|\boldsymbol{v}_1 + \boldsymbol{v}_2\| \leq \|\boldsymbol{v}_1\| + \|\boldsymbol{v}_2\|$ for all $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$.

Remark 3.11. As the norm defines the notion of distance in a vector space it also defines a topology (i.e. a notion of a limit) in on V. For example, we say that a sequence $\boldsymbol{v}_n \in V$ of vectors converges to $v \in V$ and write

$$\lim_{n \to \infty} \boldsymbol{v}_n = \boldsymbol{v},$$

if for all $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that $\|\boldsymbol{v}_n - \boldsymbol{v}\| < \varepsilon$ for all $n \geq N$.

Definition 3.12. A vector space equipped with a norm is called a **normed vector space**.

Definition 3.13. A **Hermitian inner product** on V is a function $(\cdot, \cdot) : V \times V \to \mathbb{C}$ which satisfies the following properties

1. Positive definite:
$$(\boldsymbol{v}, \boldsymbol{v}) \geq 0 \quad \forall \boldsymbol{v} \in V$$

   and $(\boldsymbol{v}, \boldsymbol{v}) \iff \boldsymbol{v} = 0$.

2. Hermiticity:
$$(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{v}, \boldsymbol{u})^* \quad \forall \boldsymbol{u}, \boldsymbol{v} \in V,$$

3. linearity in the second argument,

$$(\boldsymbol{u}, \alpha \boldsymbol{v} + \beta \boldsymbol{w}) = \alpha(\boldsymbol{u}, \boldsymbol{v}) + \beta(\boldsymbol{u}, \boldsymbol{w}) \quad \forall \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V \text{ and } \forall \alpha, \beta \in \mathbb{C},$$

4. anti-linearity in the first argument

$$(\alpha \boldsymbol{u} + \beta \boldsymbol{v}, \boldsymbol{w}) = \alpha^*(\boldsymbol{u}, \boldsymbol{w}) + \beta^*(\boldsymbol{v}, \boldsymbol{w}) \quad \forall \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V \text{ and } \forall \alpha, \beta \in \mathbb{C}.$$

Theorem 3.14

One can define a norm by
$$\|\boldsymbol{v}\| = \sqrt{(\boldsymbol{v}, \boldsymbol{v})}.$$

Lemma 3.15 (Cauchy-Schwarz inequality)

Let $\boldsymbol{u}, \boldsymbol{v} \in V$ and $(\cdot, \cdot)$ be a Hermitian inner product on V. Then

$$|(\boldsymbol{u}, \boldsymbol{v})|^2 \leq (\boldsymbol{u}, \boldsymbol{u})(\boldsymbol{v}, \boldsymbol{v}) = \|\boldsymbol{u}\|^2 \|\boldsymbol{v}\|^2.$$

Remark 3.16. The inequality is an **equality** when $\boldsymbol{u}$ is a scalar multiple of $\boldsymbol{v}$.

3.3   Hilbert spaces

To recap the previous section we have that

Vector spaces $\supseteq$ normed vector spaces $\supseteq$ Hermitian inner product vector spaces.

With a new property this will turn into

Vector spaces $\supseteq$ Banach spaces $\supseteq$ Hilbert spaces.

Definition 3.17. A <u>normed vector space</u> is called **complete** if every Cauchy sequence in this space has a limit.

Definition 3.18. A **Banach space** is a complete normed vector space.

Definition 3.19. A **Hilbert space** is a complete Hermitian inner product space (where the topology is derived from the norm defined by the inner product). The Hilbert space is denoted by $\mathcal{H}$.

Example 3.20

Some examples of Hilbert space:

- $\mathbb{C}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\}$ with inner product $\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right) = a_1^* b_1 + \cdots + a_n^* b_n.$

- $L^2(\mathbb{R}) = \{$the vector space of square integrable functions from $\mathbb{R} \to \mathbb{C}\}.$

Remark 3.21. We need to refine the definition of $L^2(\mathbb{R})$ as we can choose a function for which $(f, f) = 0$ but $f \neq 0$, for example define $f(x) = 1$ for $x = 1$ and $f(x) = 0$ otherwise.

Definition 3.22. The vector space $L^2(\mathbb{R})$ is defined to be the set of equivalence classes of square integral functions $\mathbb{R} \to \mathbb{C}$ (two square integrable functions, f_1 and f_2, are considered to be equivalent if $\|f_1 - f_2\| = 0$ 'almost everywhere'). The vector addition and scalar multiplication are defined on $L^2(\mathbb{R})$ are

$$(f_1 + f_2)(x) \equiv f_1(x) + f_2(x),$$
$$(af)(x) \equiv af(x),$$

for $f, f_1, f_2 \in L^2(\mathbb{R})$ and for all $a \in \mathbb{C}$.

Remark 3.23. By "square integrable" we mean that for $f \in L^2(\mathbb{R})$ the integral

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

i.e. it is finite.

Proof. We prove that $f + g \in L^2(\mathbb{R})$. Suppose $f, g \in L^2(\mathbb{R})$ then

$$|f(x) + g(x)|^2 \leq (|f(x)| + |g(x)|)^2$$
$$= |f(x)|^2 |g(x)|^2 + 2|f(x)||g(x)|.$$

By the AM-GM (AM $\geq$ GM) inequality we have that

$$|f(x)||g(x)| = \sqrt{|f(x)|^2 |g(x)|^2}$$
$$\leq \frac{|f(x)|^2 + |g(x)|^2}{2}.$$

Therefore,
$$2|f(x)||g(x)| \leq |f(x)|^2 |g(x)|^2,$$

which implies
$$|f(x) + g(x)|^2 \leq 2 \left(|f(x)|^2 |g(x)|^2 \right).$$

We know the integral of the RHS is finite then, clearly the integral of the LHS is also finite. $\qquad\square$

Definition 3.24. We define the Hermitian inner product on $L^2(\mathbb{R})$ as:

$$(f, g) = \int_{-\infty}^{\infty} f^*(x) g(x) \, dx.$$

Remark 3.25. With this definition we can write

$$\langle \widehat{\mathcal{O}} \rangle = \int_{-\infty}^{\infty} \Psi^*(x) (\widehat{\mathcal{O}} \Psi)(x) \, dx = (\Psi, \widehat{\mathcal{O}} \Psi).$$

Proof. This inner product is well-defined:

$$\left| \int_{-\infty}^{\infty} f^*(x) g(x) \, dx \right| \leq \int_{-\infty}^{\infty} |f(x)||g(x)| \, dx$$
$$\leq \frac{1}{2} \int_{-\infty}^{\infty} |f(x)| + |g(x)| \, dx$$
$$< \infty.$$

$\qquad\square$

Theorem 3.26

All finite dimensional vector spaces and $L^2(\mathbb{R})$ are complete, and they are all Hilbert spaces.

Remark 3.27. In particular $L^2(X)$ is complete and hence, Hilbert for any metric space X.

3.4   Dual vector space

Definition 3.28. A **linear functional**, α, on a complex vector space V is a function $\alpha : V \to \mathbb{C}$ which is linear i.e.

1. for all $c \in \mathbb{C}$ and $\boldsymbol{v} \in V$, $\alpha(c\boldsymbol{v}) = c\alpha(\boldsymbol{v})$,

2. for all $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$ we have $\alpha(\boldsymbol{v}_1 + \boldsymbol{v}_2) = \alpha(\boldsymbol{v}_1) + \alpha(\boldsymbol{v}_2)$.

> **Note 3.29.** We can think of linear functionals as linear maps which take in vectors and output scalars in $\mathbb{C}$.

Definition 3.30. Let $\mathcal{H}$ be a Hilbert space. The **(continuous) dual space** $\mathcal{H}^*$ is the vector space of all continuous linear functionals on $\mathcal{H}$.

> **Remark 3.31.** For a general vector space V we do not need continuity in the definition.

Example 3.32

Suppose $\mathcal{H} = \mathbb{C}^n$, if we have $\alpha \in (\mathbb{C}^n)^*$ and $\boldsymbol{v} \in \mathbb{C}^n$ such that

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

Then, we can define the linear functional α as the row vector

$$\alpha = (\alpha_1, \cdots, \alpha_n)$$

so that

$$\alpha(\boldsymbol{v}) = (\alpha_1, \cdots, \alpha_n) \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \alpha_1 v_1 + \cdots \alpha_n v_n.$$

That is, $\alpha : \mathbb{C}^n \to (\mathbb{C}^n)^*$. In conclusion, elements of $(\mathbb{C}^n)^*$ are row vectors i.e.

$$(\mathbb{C}^n)^* = \{(\alpha_1, \ldots, \alpha_n) : \alpha_1, \ldots, \alpha_n \in \mathbb{C}\}.$$

Theorem 3.33

Let $\mathcal{H}$ be a Hilbert space. Then there is a canonical anti-linear bijection from $\mathcal{H} \to \mathcal{H}^*$. This map is given by the inner product: given a vector $\boldsymbol{u} \in \mathcal{H}$, define $\alpha_{\boldsymbol{u}} \in \mathcal{H}^*$ by

$$\alpha_{\boldsymbol{u}}(\boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v}).$$

> **Example 3.34**
>
> Suppose $\mathcal{H} = \mathbb{C}^n$ and let
>
> $$\boldsymbol{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix},$$
>
> then, $\alpha_{\boldsymbol{u}} \in \mathcal{H}^*$ is given by the row vector
>
> $$\alpha_{\boldsymbol{u}} = \boldsymbol{u}^\dagger$$
> $$= (u_1^*, \ldots, u_n^*),$$
>
> where '$\dagger$' means 'conjugate transpose'. Indeed, for any $\boldsymbol{v} \in \mathcal{H} = \mathbb{C}^n$ we have
>
> $$\alpha_{\boldsymbol{u}}(\boldsymbol{v}) = u_1^* v_1 + \cdots + u_n^* v_n$$
> $$= (\boldsymbol{u}, \boldsymbol{v}).$$

Example 3.35. In the case of $\mathcal{H} = L^2(\mathbb{R})$ the theorem says that any linear functional $\alpha \in \mathcal{H}^*$ can be represented as

$$\alpha_g(f) = \int_{-\infty}^{\infty} g^*(x) f(x) \, dx$$

for some $g \in L^2(\mathbb{R})$.

3.5   Dirac's notation

Definition 3.36. Elements of the Hilbert space, $\mathcal{H}$, are denoted by $|v\rangle \in \mathcal{H}$, we call this '**ket**'.

Definition 3.37. Linear functionals in the dual of the Hilbert space, $\mathcal{H}^*$ are denoted by $\langle \alpha| \in \mathcal{H}^*$, we call this '**bra**'.

Theorem 3.38. The value of a linear functional $\langle \alpha|$ on a vector $|v\rangle$ is denoted by

$$\langle \alpha|v\rangle \in \mathbb{C}.$$

> **Theorem 3.39** (Relation between bra-ket)
>
> The bra vector is related to the ket vector by conjugate transpose:
>
> $$\langle u| = (|u\rangle)^\dagger.$$
>
> Therefore, the inner product can be written as
>
> $$(|u\rangle, |v\rangle) = \langle u|v\rangle.$$

Proof. The bra vector $\langle u|$ (for any label) denotes the linear functional that corresponds to $|u\rangle$, i.e. $\langle u| = \alpha_{|u\rangle}$. Recall, the relation $\alpha_{|u\rangle}^\dagger = (|u\rangle)^\dagger$ so, $\langle u| = (|u\rangle)^\dagger$. Furthermore, for any $|v\rangle \in \mathcal{H}$ we have

$$\alpha_{|u\rangle}(|v\rangle) = (|u\rangle, |v\rangle)$$
$$= \langle u|v\rangle.$$

$\square$

3.6   Orthonormal basis in $\mathcal{H}$

Definition 3.40. Let $\mathcal{H}$ be an infinite dimensional Hilbert space and let $\{|e_1\rangle, |e_2\rangle, \dots\} \subset \mathcal{H}$ be a sequence of vectors. We say that the vectors $|e_k\rangle$ are **orthonormal** if

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

The δ_{ij} represents the Kronecker δ symbol.

Theorem 3.41. Let $\{a_k\}_{k=1}^{\infty}$ be a sequence of complex numbers and $|e_i\rangle \in \mathcal{H}$. The infinite sum

$$\sum_{k=1}^{\infty} a_k |e_k\rangle = \lim_{N \to \infty} \sum_{k=1}^{N} a_k |e_k\rangle,$$

exists and converges in $\mathcal{H}$ if and only if the sum

$$\sum_{k=1}^{\infty} |a_k|^2 < \infty$$

(that is, the sum above converges and is finite).

> **Note 3.42.** This theorem allows us to write $|v\rangle$ as a linear combination of the basis of the vector space.

Proof. This proof is NOT examinable hence, we present a sketch of the proof. It is enough to prove $\sum_{k=1}^{\infty} a_k |e_k\rangle$ is a Cauchy sequence then, since $\mathcal{H}$ is complete, by definition all Cauchy sequences converge to a limit. Thus, the sum must exist. $\qquad\square$

Definition 3.43. Hilbert spaces, $\mathcal{H}$, with have elements which can be written in terms of an orthonormal basis are called **separable**.

> **Remark 3.44.** In this course, all Hilbert spaces will be separable.

3.6.1   Computing coefficients of a basis

Suppose V is a finitely dimensional vector space with an orthonormal basis $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$, that is

$$\boldsymbol{v}_i \cdot \boldsymbol{v}_j = \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

Then, any vector $\boldsymbol{x} \in V$ can be expressed in terms of the basis vectors:

$$\boldsymbol{x} = a_1 \boldsymbol{v}_1 + \dots + a_i \boldsymbol{v}_i + \dots + a_n \boldsymbol{v}_n.$$

Since the basis is orthonormal if we consider $\boldsymbol{v}_i \cdot \boldsymbol{x}$ we will obtain the following

$$\boldsymbol{v}_i \cdot \boldsymbol{x} = a_1 \underbrace{\boldsymbol{v}_i \cdot \boldsymbol{v}_1}_{0} + \cdots + a_i \boldsymbol{v}_i \cdot \boldsymbol{v}_i + \cdots + a_n \underbrace{\boldsymbol{v}_i \cdot \boldsymbol{v}_n}_{0}$$

$$= a_i \underbrace{\boldsymbol{v}_i \cdot \boldsymbol{v}_i}_{1}$$

$$= a_i.$$

We now extend this idea to Hilbert spaces. Suppose that $\mathcal{H}$ is separable then,

$$|v\rangle = \sum_{k=1}^{\infty} a_k |e_k\rangle.$$

Consider, the inner product

$$\langle e_i | v \rangle = \langle e_i | \sum_{k=1}^{\infty} a_k |e_k\rangle$$

$$= \sum_{k=1}^{\infty} a_k \langle e_i | e_k \rangle$$

$$= \sum_{k=1}^{\infty} a_k \delta_{ik}$$

$$= a_i;$$

since $\langle e_i | \in \mathcal{H}^*$ is a *continuous* linear functional, we can exchange the sum with the action of $\langle e_i |$.

Theorem 3.45

If $\mathcal{H}$ has an orthonormal basis then every $|v\rangle \in \mathcal{H}$ can be written as

$$|v\rangle = \sum_{k=1}^{\infty} a_k |e_k\rangle,$$

where the coefficients a_i are given by

$$a_i = \langle e_i | v \rangle.$$

Corollary 3.46. If $\mathcal{H}$ has an orthonormal basis then the vectors $|v\rangle \in \mathcal{H}$ are also written as

$$|v\rangle = \sum_{k=1}^{\infty} \langle e_k | v \rangle |e_k\rangle$$

$$= \sum_{k=1}^{\infty} |e_k\rangle \langle e_k | v \rangle.$$

Remark 3.47. The expression $|e_k\rangle \langle e_k|$ can be understood as a linear operator via the map

$$|v\rangle \mapsto (|e_k\rangle \langle e_k|) |v\rangle$$

$$\mapsto |e_k\rangle \langle e_k | v \rangle.$$

Therefore, the expression

$$\sum_{k=1}^{\infty} |e_k\rangle \langle e_k| = 1;$$

the identity operator of this space.

3.7   Delta-normalisable states

3.7.1   Motivation

In this section we introduce a new piece of technology that we will need. In the previous section we discussed a basis composed of vector $|e_k\rangle$, with k as a discrete label. In this section, we will discuss the case when k is a continuous label. A motivating example: consider $\mathcal{H} = L^2(\mathbb{R})$ with the functions

$$e_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx}$$

where $k \in \mathbb{R}$. These functions are not actually elements of $\mathcal{H}$ because they are not square-integrable as,

$$\int_{-\infty}^{\infty} |e_k(x)|^2 \, dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx = \infty.$$

Nevertheless, we pretend that these functions are elements of $\mathcal{H}$ and try to understand what their inner product is:

$$\langle e_{k_1} | e_{k_2} \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(e^{ik_1 x} \right)^* e^{ik_2 x} \, dx$$
$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i(k_2 - k_1)x} \, dx.$$

Clearly, this integral does not converge in the usual sense. However, we can define

$$\int_{-\infty}^{\infty} e^{ikx} \, dx = 2\pi \delta(k),$$

where $\delta(k)$ is the Dirac delta function. Therefore,

$$\langle e_{k_1} | e_{k_2} \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i(k_2 - k_1)x} \, dx$$
$$= \delta(k_2 - k_1)$$
$$= \delta(k_1 - k_2).$$

Strictly speaking $|e_k\rangle \notin \mathcal{H}$, but it is possible to form elements of $\mathcal{H}$ using them. By defining

$$|\Psi\rangle = \int_{-\infty}^{\infty} a(k) \, |e_k\rangle \, dk$$

for some function $f(k)$, we have that $|\Psi\rangle \in \mathcal{H}$ if $f(k)$ is square-integrable. In fact

$$\langle \Psi | \Psi \rangle = \left(\langle \Psi | \right)^\dagger | \Psi \rangle$$

$$= \left(\int_{-\infty}^{\infty} a(k') \, |e_{k'}\rangle \, dk' \right)^\dagger \int_{-\infty}^{\infty} a(k) \, |e_k\rangle \, dk$$

$$= \int_{-\infty}^{\infty} a^*(k') \, \langle e_{k'}| \, dk' \int_{-\infty}^{\infty} a(k) \, |e_k\rangle \, dk$$

$$= \int_{-\infty}^{\infty} a^*(k') \, dk' \int_{-\infty}^{\infty} a(k) \, \langle e_{k'}|e_k\rangle \, dk$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} a^*(k')a(k)\delta(k' - k) \, dk' \, dk.$$

Consider the following identity

$$\int_{-\infty}^{\infty} f(k')\delta(k' - k) \, dk' = f(k'),$$

by changing the variable $k' \to k' + k$ the integral becomes

$$\int_{-\infty}^{\infty} f(k' + k)\delta(k') \, dk' = f(k),$$

as such we can write

$$\int_{\infty}^{\infty} a^*(k')\delta(k' - k) \, dk' \to \int_{\infty}^{\infty} a^*(k' + k)\delta(k') \, dk' = a^*(k).$$

Therefore,

$$\langle \Psi | \Psi \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} a^*(k')a(k)\delta(k' - k) \, dk' \, dk$$

$$= \int_{-\infty}^{\infty} a(k)a^*(k) \, dk$$

$$= \int_{-\infty}^{\infty} |a(k)|^2 \, dk,$$

which is finite if $a(k)$ is square integrable.

3.7.2   The Dirac delta function

Definition 3.48. The **Dirac delta function** is defined as

$$\delta(k) = \begin{cases} \infty & \text{if } k = 0 \\ 0 & \text{if } k \neq 0 \end{cases}$$

with the property that

$$\int_{-\infty}^{\infty} \delta(k) \, dk = 1.$$

Remark 3.49. Although, it is called a 'function' it is not truly a function, rather a distribution.

Theorem 3.50

Properties of the δ-function:

- $\delta(x) \in \mathbb{R}$,

- $\delta(x) \geq 0$,

- $\delta(x) = \delta(-x)$ (even).

- $\int_{-\infty}^{\infty} \delta(k) f(k)\, dk = f(0)$ (sometimes we use the notation $\delta(f) = f(0)$);

- $\int_{-\infty}^{\infty} \delta(k - a) f(k)\, dk = f(a)$;

- $\int_{-\infty}^{\infty} e^{ikx}\, dx = 2\pi \delta(k)$;

- $\langle e_x | e_y \rangle = \delta(x - y)$.

3.7.3   Stocktaking

Definition 3.51. We say that $|e_x\rangle$ and $|e_y\rangle$ are **delta-normalisable** if

$$\langle e_x | e_y \rangle = \delta(x - y).$$

Definition 3.52. We say that $|e_k\rangle$ for $k \in \mathbb{R}$ form a **delta-normalisable basis** of $\mathcal{H}$ if any $|\Psi\rangle \in \mathcal{H}$ can be represented in the form

$$|\Psi\rangle = \int_{-\infty}^{\infty} a(k) |e_k\rangle\, dk,$$

for some function square-integrable function $a(k)$ (i.e. $a(k) \in L^2(\mathbb{R})$).

Remark 3.53. Strictly, speaking $|e_k\rangle \notin \mathcal{H}$ as they are not square integrable.

Note 3.54. We can think of $a(k)$ as the analogous to the coefficients a_k.

Example 3.55

Vectors of the form

$$|e_k\rangle = \frac{1}{\sqrt{2\pi}} e^{ikx}$$

form a delta-normalisable basis.

Theorem 3.56

If $\mathcal{H}$ has a delta-normalisable basis the 'coefficients' $a(k)$ are given by

$$a(k) = \langle e_k | \Psi \rangle.$$

Proof. Consider a $|\Psi\rangle \in \mathcal{H}$ where $\mathcal{H}$ has a delta-normalisable basis, $\{|e_j\rangle\}$, then we can write

$$|\Psi\rangle = \int_{-\infty}^{\infty} a(j) |e_j\rangle \, dj.$$

We have that

$$\begin{aligned}
\langle e_k|\Psi\rangle &= \langle e_k| \int_{-\infty}^{\infty} a(j) |e_j\rangle \, dj \\
&= \int_{-\infty}^{\infty} \langle e_k| \, a(j) |e_j\rangle \, dj \\
&= \int_{-\infty}^{\infty} a(j) \langle e_k|e_j\rangle \, dj \\
&= \int_{-\infty}^{\infty} a(j)\delta(k-j) \, dj \\
&= a(k).
\end{aligned}$$

$\square$

Corollary 3.57. If $\mathcal{H}$ has a delta-normalisable basis then every $|\Psi\rangle \in \mathcal{H}$ can be represented as

$$\begin{aligned}
|\Psi\rangle &= \int_{-\infty}^{\infty} \langle e_k|\Psi\rangle |e_k\rangle \, dk \\
&= \int_{-\infty}^{\infty} |e_k\rangle \langle e_k|\Psi\rangle \, dk.
\end{aligned}$$

Remark 3.58. Thus, the expression

$$\int_{-\infty}^{\infty} |e_k\rangle \langle e_k| \, dk = 1.$$

4 Linear operators and observables

4.1 Linear operators

Definition 4.1. A **linear operator** $\widehat{\mathcal{O}}$ on a vector space V is a linear map $\widehat{\mathcal{O}} : V \to V$.

Note 4.2. By linear we mean the usual linear conditions are met:

$$\widehat{\mathcal{O}} \left(\alpha |v\rangle + \beta |u\rangle \right) = \alpha\widehat{\mathcal{O}} |v\rangle + \beta\widehat{\mathcal{O}} |u\rangle.$$

Definition 4.3. The **identity operator**, denote by $\widehat{1}$, is defined as

$$\widehat{1} |v\rangle = |v\rangle.$$

Corollary 4.4. Scalars can be written as operators. For example, the operator $\widehat{\mathcal{O}} + 2$ acts on vectors by

$$\left(\widehat{\mathcal{O}} + 2 \right) |v\rangle = \widehat{\mathcal{O}} |v\rangle + 2 \cdot \widehat{1} |v\rangle.$$

Theorem 4.5. Operators form a vector space.

Proof. Addition and scalar multiplication is closed i.e.

$$\left(\widehat{\mathcal{O}}_1 + \widehat{\mathcal{O}}_2\right)|v\rangle = \widehat{\mathcal{O}}_1|v\rangle + \widehat{\mathcal{O}}_2|v\rangle,$$

$$\left(\alpha\widehat{\mathcal{O}}\right)|v\rangle = \alpha\widehat{\mathcal{O}}|v\rangle.$$

$\square$

Theorem 4.6. The product of operators $\widehat{\mathcal{O}}_1\widehat{\mathcal{O}}_2$ is defined as the composition of operators (maps)

$$\left(\widehat{\mathcal{O}}_1\widehat{\mathcal{O}}_2\right)|v\rangle = \widehat{\mathcal{O}}_1\left(\widehat{\mathcal{O}}_2|v\rangle\right).$$

Remark 4.7. Since, composition of maps is not always commutative the same applies for composition of operators.

Remark 4.8. Recall if $|\Psi\rangle \in L^2(\mathbb{R})$ is the function $\Psi(x)$ then, $\widehat{p}|\Psi\rangle$ is the function $-i\hbar\dfrac{\partial}{\partial x}\Psi(x)$. However, not every function in $L^2(\mathbb{R})$ is differentiable, and even if they are their derivative is not necessarily in $L^2(\mathbb{R})$. The operator $\widehat{p}$, as well as most other operator in QM are not well-defined linear operators, they are called **densely-defined** linear operators.

A densely defined linear operator $\widehat{\mathcal{O}}$ on a Hilbert space $\mathcal{H}$ with domain D is a linear function $D \to \mathcal{H}$, where $D \subseteq \mathcal{H}$ is a dense linear subspace. Note that the domain D need to be supplied as a part of the definition of $\widehat{\mathcal{O}}$. Thus, just saying $\widehat{p}$ does not strictly make sense.

4.2   Expectation values and matrix elements

Theorem 4.9

Let $\widehat{\mathcal{O}} : \mathcal{H} \to \mathcal{H}$ be a linear operator on a Hilbert space $\mathcal{H}$. Then, the **expectation value** of $\widehat{\mathcal{O}}$ in a state $|\Psi\rangle \in L^2(\mathbb{R})$ is defined as

$$\langle\widehat{\mathcal{O}}\rangle = \frac{\langle\Psi|\,\widehat{\mathcal{O}}\,|\Psi\rangle}{\langle\Psi|\Psi\rangle}.$$

Proof. Recall the expectation of an operator is defined as

$$\langle\widehat{\mathcal{O}}\rangle = \int_{-\infty}^{\infty} \Psi^*(x)\left(\widehat{\mathcal{O}}\Psi\right)(x)\,dx.$$

Furthermore, if $\mathcal{H} = L^2(\mathbb{R})$ the inner product is defined as

$$(f,g) = \int_{-\infty}^{\infty} f^*(x)g(x)\,dx.$$

If Ψ is not normalised then, we can normalise it by dividing by $|\Psi|^2 = \langle\Psi|\Psi\rangle$ thus, the definition of $\langle\widehat{\mathcal{O}}\rangle$ follows as above. $\square$

Corollary 4.10. If the state $|\Psi\rangle$ is normalised, $\langle\Psi|\Psi\rangle = 1$ then, the expectation value is

$$\langle\widehat{\mathcal{O}}\rangle = \langle\Psi|\,\widehat{\mathcal{O}}\,|\Psi\rangle.$$

4.2.1 Matrix elements

Suppose that $\{|e_1\rangle, |e_2\rangle, \ldots\} \subset \mathcal{H}$ is an orthonormal of $\mathcal{H}$. Then, every vector (by the previous discussion) can be written as

$$|\Psi\rangle = \sum_{k=1}^{\infty} a_k |e_k\rangle$$

$$= \sum_{k=1}^{\infty} \langle e_k|\Psi\rangle |e_k\rangle.$$

Since $|\Psi\rangle$ is a (column) vector, we can think of the coefficients $a_k = \langle e_k|\Psi\rangle$ as the co-ordinates (components) of $|\Psi\rangle$ in the basis $\{|e_1\rangle, |e_2\rangle, \ldots\}$. Applying the same argument to $\widehat{\mathcal{O}}|\Psi\rangle$ we get

$$\widehat{\mathcal{O}}|\Psi\rangle = \sum_{k=1}^{\infty} b_k |e_k\rangle$$

$$= \sum_{k=1}^{\infty} \langle e_k|\widehat{\mathcal{O}}|\Psi\rangle |e_k\rangle.$$

The coordinates of $\widehat{\mathcal{O}}|\Psi\rangle$ in the same basis are $\langle e_k|\widehat{\mathcal{O}}|\Psi\rangle$. Using the previous expression we can express

$$\langle e_k|\widehat{\mathcal{O}}|\Psi\rangle = \langle e_k| \sum_{j=1}^{\infty} \langle e_j|\widehat{\mathcal{O}}|\Psi\rangle |e_j\rangle$$

$$= \sum_{j=1}^{\infty} \langle e_k| \langle e_j|\widehat{\mathcal{O}}|\Psi\rangle |e_j\rangle$$

$$= \sum_{j=1}^{\infty} \langle e_k|\widehat{\mathcal{O}}|\Psi\rangle \underbrace{\langle e_j|e_j\rangle}_{1}.$$

Now substituting the expression for $|\Psi\rangle$ we have

$$\langle e_k|\widehat{\mathcal{O}}|\Psi\rangle = \sum_{j=1}^{\infty} \langle e_k|\widehat{\mathcal{O}} \langle e_j|\Psi\rangle |e_j\rangle$$

$$= \sum_{j=1}^{\infty} \langle e_k|\widehat{\mathcal{O}}|e_j\rangle \langle e_j|\Psi\rangle.$$

The coordinates of the vector can be arranged into a column vector:

$$\begin{pmatrix} \langle e_1|\Psi\rangle \\ \langle e_2|\Psi\rangle \\ \vdots \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \langle e_1|\widehat{\mathcal{O}}|\Psi\rangle \\ \langle e_2|\widehat{\mathcal{O}}|\Psi\rangle \\ \vdots \end{pmatrix}.$$

Then we can write the identity for $\langle e_k|\widehat{\mathcal{O}}|\Psi\rangle$ as the following matrix product:

$$\begin{pmatrix} \langle e_1|\widehat{\mathcal{O}}|\Psi\rangle \\ \langle e_2|\widehat{\mathcal{O}}|\Psi\rangle \\ \vdots \end{pmatrix} = \begin{pmatrix} \langle e_1|\widehat{\mathcal{O}}|e_1\rangle & \langle e_1|\widehat{\mathcal{O}}|e_2\rangle & \cdots \\ \langle e_2|\widehat{\mathcal{O}}|e_1\rangle & \langle e_2|\widehat{\mathcal{O}}|e_2\rangle & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \langle e_1|\Psi\rangle \\ \langle e_2|\Psi\rangle \\ \vdots \end{pmatrix}.$$

Remark 4.11. If $\mathcal{H}$ is a finite vector space then clearly, the matrix become finite too.

4.3 Adjoint operator

Definition 4.12. Let $\widehat{\mathcal{O}} : \mathcal{H} \to \mathcal{H}$ be a linear operator. The **adjoint** of $\widehat{\mathcal{O}}$, denoted by $\widehat{\mathcal{O}}^\dagger$, is an operator such that

$$\left(|u\rangle, \widehat{\mathcal{O}} |v\rangle \right) = \left(\widehat{\mathcal{O}}^\dagger |u\rangle, |v\rangle \right) \quad \forall |u\rangle, |v\rangle \in \mathcal{H}.$$

Note 4.13. Emphasis: the brackets denote the inner product.

Theorem 4.14

The **Hermitian conjugate**, denoted by $\dagger$, can act on expressions involving numbers, linear operator and bra-ket vectors. It turns

- numbers into their complex conjugate;

- bra into ket vectors and vice versa;

- operators into their adjoint;

Furthermore, the Hermitian conjugate reverses the order of all products.

Example 4.15

Consider the number $\langle u| \widehat{\mathcal{O}}_1 \widehat{\mathcal{O}}_2 |v\rangle$ then,

$$\left(\langle u| \widehat{\mathcal{O}}_1 \widehat{\mathcal{O}}_2 |v\rangle \right)^* = \left(\langle u| \widehat{\mathcal{O}}_1 \widehat{\mathcal{O}}_2 |v\rangle \right)^\dagger$$
$$= (|v\rangle)^\dagger \, \widehat{\mathcal{O}}_2^\dagger \widehat{\mathcal{O}}_1^\dagger \, (\langle u|)^\dagger$$
$$= \langle v| \widehat{\mathcal{O}}_2^\dagger \widehat{\mathcal{O}}_1^\dagger |u\rangle.$$

Proposition 4.16

Properties of the adjoint operator:

- $\left(\widehat{\mathcal{O}}^\dagger \right)^\dagger = \widehat{\mathcal{O}}$;

- $\left(\alpha \widehat{\mathcal{O}}_1 + \beta \widehat{\mathcal{O}}_2 \right)^\dagger = \alpha^* \widehat{\mathcal{O}}_1^\dagger + \beta^* \widehat{\mathcal{O}}_2^\dagger$;

- $\left(\widehat{\mathcal{O}}_1 \widehat{\mathcal{O}}_2 \right)^\dagger = \widehat{\mathcal{O}}_2^\dagger \widehat{\mathcal{O}}_1^\dagger$.

4.4   Self-adjoint (Hermitian) operators

Definition 4.17. Let $\mathcal{H}$ be a Hilbert space. A linear operator $\widehat{\mathcal{O}} : \mathcal{H} \to \mathcal{H}$ is called **Hermitian** (or **self-adjoint**) if $\widehat{\mathcal{O}} = \widehat{\mathcal{O}}^\dagger$.

> **Remark 4.18.** The domain of a Hermitian operator $\widehat{\mathcal{O}}^\dagger$ is the same as the domain of the operator $\widehat{\mathcal{O}}$.

Definition 4.19. An observable on $\mathcal{H}$ is a Hermitian operator on $\mathcal{H}$.

Theorem 4.20

Let $\widehat{\mathcal{O}}$ be a Hermitian operator. The properties of Hermitian operators are as follows:

- matrix elements of $\widehat{\mathcal{O}}$ form a Hermitian matrix i.e. $M = M^\dagger$;

- the expectation values $\widehat{\mathcal{O}}$ are real i.e. $\widehat{\mathcal{O}} \in \mathbb{R}$;

- the eigenvalues of $\widehat{\mathcal{O}}$ are real;

- the eigenvectors with distinct eigenvalues are orthogonal to each other i.e. if $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are eigenvectors of with eigenvalues λ_1 and λ_2 respectively with $\lambda_1 \neq \lambda_2$ then $\langle \Psi_1 | \Psi_2 \rangle = 0$.

Proof. We prove each statement in turn.

- The matrix elements of $\widehat{\mathcal{O}}$ are the numbers $\langle e_k | \widehat{\mathcal{O}} | e_j \rangle$ so, consider

$$
\begin{aligned}
\left(\langle e_k | \widehat{\mathcal{O}} | e_j \rangle \right)^* &= \left(\langle e_k | \widehat{\mathcal{O}} | e_j \rangle \right)^\dagger \\
&= (|e_j\rangle)^\dagger \widehat{\mathcal{O}}^\dagger (\langle e_k|)^\dagger \\
&= \langle e_j | \widehat{\mathcal{O}} | e_k \rangle .
\end{aligned}
$$

We can exchange $* \leftrightarrow \dagger$ since $\langle e_k | \widehat{\mathcal{O}} | e_j \rangle$ is just a number.

- The expectation of $\langle \widehat{\mathcal{O}} \rangle = \langle \Psi | \widehat{\mathcal{O}} | \Psi \rangle$ then, consider

$$
\begin{aligned}
\langle \widehat{\mathcal{O}} \rangle^* &= \left(\langle \Psi | \widehat{\mathcal{O}} | \Psi \rangle \right)^\dagger \\
&= \langle \Psi | \widehat{\mathcal{O}} | \Psi \rangle .
\end{aligned}
$$

Therefore, $\langle \widehat{\mathcal{O}} \rangle = \langle \widehat{\mathcal{O}} \rangle^*$.

- Recall what is an eigenvector: $|\Psi\rangle$ of $\widehat{\mathcal{O}}$ is an eigenvector if

$$
\widehat{\mathcal{O}} | \Psi \rangle = \lambda | \Psi \rangle \quad \text{for } |\Psi\rangle \neq \mathbf{0}.
$$

Taking the Hermitian conjugate (†) of this equation we have

$$
\begin{aligned}
\left(\widehat{\mathcal{O}} | \Psi \rangle \right)^\dagger &= (\lambda | \Psi \rangle)^\dagger \\
\Rightarrow \langle \Psi | \widehat{\mathcal{O}}^\dagger &= \lambda^* \langle \Psi | \\
\Rightarrow \langle \Psi | \widehat{\mathcal{O}} &= \lambda^* \langle \Psi | .
\end{aligned}
$$

Thus, if $|\Psi\rangle$ is a 'ket eigenvector' of $\widehat{\mathcal{O}}$ with eigenvalue λ then, $\langle\Psi|$ is a 'bra eigenvector' of $\widehat{\mathcal{O}}$ with eigenvalue λ^*. Now consider the equation

$$\left(\langle\Psi|\,\widehat{\mathcal{O}}\right)|\Psi\rangle = \langle\Psi|\left(\widehat{\mathcal{O}}\,|\Psi\rangle\right),$$

and apply the definition of an eigenvector, we obtain

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

If $|\Psi\rangle \neq 0$ then $\langle\Psi|\Psi\rangle \neq 0$ which implies

$$\lambda^* = \lambda.$$

- Suppose $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are eigenvectors of the Hermitian operator $\widehat{\mathcal{O}}$ with distinct eigenvalues $\lambda_1 \neq \lambda_2$. Consider,

$$\left(\langle\Psi_1|\,\widehat{\mathcal{O}}\right)|\Psi_2\rangle = \langle\Psi_1|\left(\widehat{\mathcal{O}}\,|\Psi_2\rangle\right).$$

Using the definition of an eigenvector we obtain

$$\lambda_1\langle\Psi_1|\Psi_2\rangle = \lambda_2\langle\Psi_1|\Psi_2\rangle$$
$$\Rightarrow (\lambda_1 - \lambda_2)\langle\Psi_1|\Psi_2\rangle = 0.$$

If $\lambda_1 \neq \lambda_2$, we can divide by $(\lambda_1 - \lambda_2)$ and conclude

$$\langle\Psi_1|\Psi_2\rangle = 0.$$

$\square$

Example 4.21

The following are examples of Hermitian operators i.e. observables:

- the position operator, $\widehat{x}$;

- the momentum operator, $\widehat{p}$;

- the Hamiltonian operator, $\widehat{H}$.

---

5 The spectral theorem

Theorem 5.1 (The spectral theorem)

Let $\mathcal{H}$ be a Hilbert space and let $\widehat{\mathcal{O}}$ be a self-adjoint linear operator on $\mathcal{H}$. Then

- (Any two eigenvectors of $\widehat{\mathcal{O}}$ with distinct eigenvalues are orthogonal to each other);

- the set of all eigenvalues of $\widehat{\mathcal{O}}$, which we call the **spectrum**, is a subset of $\mathbb{R}$ i.e. all the eigenvalues of $\widehat{\mathcal{O}}$ are real;

- the eigenvectors of $\widehat{\mathcal{O}}$ form a complete basis of the Hilbert space $\mathcal{H}$.

> **Note 5.2.** In practice, we need to distinguish between a *discrete* and *continuous* spectrum of $\widehat{\mathcal{O}}$.

5.1   Discrete spectrum

Suppose $\widehat{\mathcal{O}}$ is a Hermitian operator with eigenvalues (with discrete labels) $\{\lambda_1, \lambda_2, \ldots\}$. By the spectral theorem $\{\lambda_1, \lambda_2, \ldots\}$. Let us denote the corresponding eigenvectors by

$$|\lambda_1\rangle, |\lambda_2\rangle, \ldots$$

(assuming for simplicity that each eigenvalue has just one corresponding eigenvector). Furthermore, imposing the normalisation condition on $|\lambda_k\rangle$ we have $\langle\lambda_k|\lambda_k\rangle = 1$ thus, by the spectral theorem the set of $|\lambda_k\rangle$ form an orthonormal basis of the Hilbert space. This means that $\langle\lambda_k|\lambda_j\rangle = \delta_{jk}$ and any vector $|\Psi\rangle \in \mathcal{H}$ can be written as

$$|\Psi\rangle = \sum_{k=1}^{\infty} c_k |\lambda_k\rangle \quad \text{where } c_k = \langle\lambda_k|\Psi\rangle.$$

Example 5.3

The Hamiltonian operator, $\widehat{H} = \frac{\widehat{p}}{2m}$ acting $L^2([0, L])$ has eigenvectors

$$|\Psi_n\rangle \to \Psi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{\pi x}{L} n\right)$$

with corresponding eigenvalues

$$|E_n\rangle \to E_n = \frac{\hbar^2 \pi^2 n^2}{2L^2 m}.$$

Therefore, by the spectral theorem the eigenvectors $|\Psi_n\rangle$ form an orthonormal basis of $\widehat{H}$.

Remark 5.4. Above, we assume the multiplicity of each eigenvalue to be 1, (by multiplicity we mean how many eigenvectors are associated to each eigenvalue). If the number of eigenvectors with eigenvalue λ_k is n_k then, we can denote them by $|\lambda_k, a\rangle$ where $a \in \{1, \ldots, n_k\}$. In general, we have

$$\langle\lambda_i, a|\lambda_k, b\rangle = 0 \quad \text{for } i \neq k.$$

If $i = k$, these inner products are not constrained and depend on the choice of $|\lambda_k, a\rangle$ for fixed k. We can choose an orthonormal basis of eigenvectors for a given eigenvalue λ_k,

$$\langle\lambda_k, a|\lambda_k, b\rangle = \delta_{ab}$$

so, all together we have

$$\langle\lambda_i, a|\lambda_k, b\rangle = \delta_{ik}\delta_{ab}.$$

Thus, any vector can be written as

$$|\Psi\rangle = \sum_{i=1}^{\infty} \sum_{a=1}^{n_i} c_{i,a} |\lambda_i, a\rangle \quad \text{where } c_{i,a} = \langle \lambda_i, a | \Psi \rangle.$$

5.2 Continuous spectrum

The spectrum of $\widehat{\mathcal{O}}$ can also be continuous. For example, it could be the interval

$$[\lambda_{\min}, \lambda_{\max}] \subset \mathbb{R}.$$

Again, denote the corresponding eigenvectors by λ with $\lambda \in [\lambda_{\min}, \lambda_{\max}]$. Imposing the normalisation condition we have

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

hence, by the spectral theorem each $|\lambda\rangle$ forms delta-normalisable basis of $\mathcal{H}$ so, any $|\Psi\rangle \in \mathcal{H}$ can be written as

$$|\Psi\rangle = \int_{\lambda_{\min}}^{\lambda_{\max}} c(\lambda) |\lambda\rangle \, d\lambda \quad \text{where } c(\lambda) = \langle \lambda | \Psi \rangle.$$

Example 5.5

Consider the momentum operator, $\widehat{p} = -i\hbar\dfrac{\partial}{\partial x}$ acting on $L^2(\mathbb{R})$. We evaluate the spectrum of $\widehat{p}$:

$$\widehat{p}\ket{\Psi} = \lambda\ket{\Psi} \iff -i\hbar\frac{\partial\Psi(x)}{\partial x} = \lambda\Psi(x)$$
$$\Rightarrow \Psi(x) = C_\lambda e^{\frac{i\lambda}{\hbar}x}.$$

We have two ways of determining the spectrum:

- we know $\widehat{p}$ is a Hermitian operator, so it has real eigenvalues;

- if $\lambda \in \mathbb{C}$ then the exponential term diverges when imposing the normalisation condition thus, it is not useful on the other hand, when $\lambda \in \mathbb{R}$ $\Psi(x)$ oscillates and can be delta-normalised.

In conclusion, the eigenvalues of $\widehat{p}$ must be real. Therefore, the spectrum of $\widehat{p}$ is whole real line i.e.
$$\text{spectrum} = \mathbb{R}.$$

We know, check the delta-normalisation condition:

$$\braket{\lambda|\lambda'} = \int_{-\infty}^{\infty} \left(C_\lambda e^{\frac{i\lambda}{\hbar}x}\right)^* \left(C_{\lambda'} e^{\frac{i\lambda}{\hbar}x}\right) dx$$
$$= \int_{-\infty}^{\infty} C_\lambda^* C_{\lambda'} e^{ix\frac{(\lambda'-\lambda)}{\hbar}} dx.$$

Using the identity $\int_{-\infty}^{\infty} e^{ikx} dx = 2\pi\delta(k)$ and making the change of variable $x = \hbar y$ we have that

$$\braket{\lambda|\lambda'} = C_\lambda^* C_{\lambda'} \int_{-\infty}^{\infty} e^{iy(\lambda'-\lambda)} dy$$
$$= C_\lambda^* C_{\lambda'}\,\hbar\,2\pi\,\delta(\lambda'-\lambda).$$

Therefore, we set
$$\ket{\lambda} = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i\lambda}{\hbar}x}$$

so that it is delta-normalised.

Remark 5.6. It is common to denote the eigenvector of an operator by the label of the operator. For example, the eigenvectors of $\widehat{p}$ would be denoted by $\ket{p}$; similarly the eigenvalues of $\widehat{p}$ are denoted by p. As such we have expression as these:

- $\widehat{p}\ket{p} = p\ket{p}$,

- $\braket{p|p'} = \delta(p-p')$.

Therefore, a basis of the momentum operator can be written as

$$\ket{\Psi} = \int_{-\infty}^{\infty} f(p)\ket{p}\,dp \quad \forall\,\ket{p} \in \mathcal{H},$$

as a function

$$\Psi(x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\hbar}} f(p) e^{\frac{ip}{\hbar} x} \, dp.$$

Remark 5.7. Similarly to the discrete case, we could have multiple eigenvectors for a given eigenvalue λ, which we could label as $|\lambda, a\rangle$ where $a \in \{1, \ldots, n_\lambda\}$. In this case, the above would generalise to

$$\langle \lambda, a | \lambda', b \rangle = \delta_{ab} \delta(\lambda - \lambda'),$$

$$|\Psi\rangle = \int_{\min}^{\max} \sum_{a=1}^{n_\lambda} c(\lambda, a) \, |\lambda\rangle \, d\lambda, \quad \text{where } c(\lambda, a) = \langle \lambda, a | \Psi \rangle.$$

Remark 5.8. The spectrum can be a mixture of discrete and continuous eigenvalues. For example,

$$\text{spectrum} = [\lambda_{\min}, \lambda_{\max}] \cup \{\lambda_1, \lambda_2, \ldots\}.$$

In this case we have

$$|\Psi\rangle = \sum_{i=1}^{\infty} \sum_{a=1}^{n_i} c_{i,a} \, |\lambda_i, a\rangle + \int_{\lambda_{\min}}^{\lambda_{\max}} \sum_{a=1}^{n_\lambda} c(\lambda, a) \, |\lambda\rangle \, d\lambda.$$

5.3 Measurements

A key axiom of quantum theory is that if a state $|\Psi\rangle$ is an eigenvector of an operator $\widehat{\mathcal{O}}$ with eigenvalue λ,

$$\widehat{\mathcal{O}} |\Psi\rangle = \lambda |\Psi\rangle$$

then, measuring the observable described by $\widehat{\mathcal{O}}$ in the state $|\Psi\rangle$ will always give the result λ. Thus, we say $\widehat{\mathcal{O}}$ has the definite value λ in the state $|\Psi\rangle$.

Suppose, we carry out a measurement in a more general state $|\Psi\rangle$ which is not an eigenvector of $\widehat{\mathcal{O}}$. In general, the measurement will cause the system to transition from $|\Psi\rangle$ to some new state $|\Phi\rangle$. Assume, the outcome of this measurement is some number λ. We employ the following physical intuition: if *immediately* after performing the measurement of $\widehat{\mathcal{O}}$ we perform a second measurement of $\widehat{\mathcal{O}}$, we should always get the same result λ (that is, the result of the first measurement may be random, but the result of the second measurement must be the same as the first). The first measurement happens in the state $|\Psi\rangle$, which transitions into a state $|\Phi\rangle$, as such the second measurement happens in the state $|\Phi\rangle$. By the above, we expect that $\widehat{\mathcal{O}}$ has a definite value in the state $|\Phi\rangle$ thus, we expect $|\Phi\rangle$ to be an eigenvector of $\widehat{\mathcal{O}}$ with eigenvalue λ, i.e.

$$\widehat{\mathcal{O}} |\Phi\rangle = \lambda |\Phi\rangle.$$

In conclusion: the only possible outcomes of measuring $\widehat{\mathcal{O}}$ are the eigenvalues of $\widehat{\mathcal{O}}$; if we measure $\widehat{\mathcal{O}}$ in a state $|\Psi\rangle$ and get a result λ then, the system will have transitioned to a state $|\Phi\rangle$ which is an eigenvector of $\widehat{\mathcal{O}}$ with the eigenvalue λ in the process of the measurement.

5.4 Probabilities

Let $\widehat{\mathcal{O}}$ be a Hermitian operator that describes a physical observable. Since $\widehat{\mathcal{O}}$ is a Hermitian operator, by the spectral theorem the eigenvectors of $\widehat{\mathcal{O}}$ form a complete basis. Assume for simplicity that the spectrum is discrete and that each eigenvalue has multiplicity 1. Thus, we can write any $|\Psi\rangle$ as

$$|\Psi\rangle = \sum_{i=1}^{\infty} c_i \, |\lambda_i\rangle \,,$$

where $|\lambda_i\rangle$ are the normalised eigenvectors of $\widehat{\mathcal{O}}$ corresponding to the eigenvalues λ_i. If $|\Psi\rangle$ is normalised i.e. $\langle\Psi|\Psi\rangle = 1$, implies that

$$\begin{aligned}
1 &= \langle\Psi|\Psi\rangle \\
&= \sum_{\substack{i=1 \\ j=1}}^{\infty} c_j^* c_i \, \langle\lambda_j|\lambda_i\rangle \\
&= \sum_{i=1}^{\infty} |c_i|^2.
\end{aligned}$$

As such, we interpret the numbers $|c_i|^2$ as the probabilities of obtaining the result λ_i when measuring $\widehat{\mathcal{O}}$ in the state $|\Psi\rangle$.

Theorem 5.9

If $|\Psi\rangle$ is normalised we can write $|\Psi\rangle = \sum_{i=1}^{\infty} c_i \, |\lambda_i\rangle$, and we interpret the numbers

$$|c_i|^2$$

as the probabilities of obtaining the result λ_i when measuring $\widehat{\mathcal{O}}$ in the state $|\Psi\rangle$. The numbers c_i are called the **amplitudes** of $|\Psi\rangle$.

Theorem 5.10

In the case of a continuous spectrum, the coefficients represent the probability density.

Example 5.11. Consider the position operator $\widehat{x}$; we evaluate the eigenvalues and eigenvectors. Recall, the action of $\widehat{x}$ is

$$(\widehat{x}\Psi)(y) = y\Psi(y).$$

We solve the eigenvalue equation

$$\begin{aligned}
\widehat{x}\,|\Psi\rangle &= \lambda\,|\Psi\rangle \\
\Rightarrow x\Psi(x) &= \lambda\Psi(x) \\
\Rightarrow (x-\lambda)\Psi(x) &= 0.
\end{aligned}$$

If $\lambda \in \mathbb{C}$ and $\lambda \notin \mathbb{R}$ then $(x-\lambda) \neq 0$ which implies $\Psi(x) = 0$. Similarly, if $\lambda \in \mathbb{R}$ and $x \neq \lambda$ then, $\Psi(x) = 0$. We conclude, $\lambda \in \mathbb{R}$ and $x \neq \lambda$ so, we set

$$\Psi(x) = \delta(x - \lambda).$$

Denoting the corresponding vector by $|\lambda\rangle$, we compute

$$\langle\lambda|\lambda\rangle = \int_{-\infty}^{\infty} \delta(x-\lambda)\delta(x-\lambda')\,dx.$$

To evaluate this integral we recall the identity

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

so, setting $f(x) = \delta(x-\lambda')$ we have

$$\langle\lambda|\lambda\rangle = \int_{-\infty}^{\infty} \delta(x-\lambda)\delta(x-\lambda')\,dx$$
$$= \delta(\lambda-\lambda').$$

Therefore, $|\lambda\rangle$ form a delta-normalisable basis. Since, we could choose λ to be any real number, the spectrum of $\widehat{x}$ is $\mathbb{R}$.

Now that we have a basis each vector $|\Psi\rangle \in \mathcal{H}$ can be expressed as

$$|\Psi\rangle = \int_{-\infty}^{\infty} c(\lambda)\,|\lambda\rangle\,d\lambda.$$

In terms of the wave function this becomes

$$\Psi(x) = \int_{-\infty}^{\infty} c(\lambda)\delta(x-\lambda)\,d\lambda$$
$$= c(\lambda).$$

Therefore, the original expression in terms of vectors can be written as

$$|\Psi\rangle = \int_{-\infty}^{\infty} \Psi(\lambda)\,|\lambda\rangle$$

however, it is common to rename the integration variable $\lambda \to x$, so that we have

$$|\Psi\rangle = \int_{-\infty}^{\infty} \Psi(x)\,|x\rangle\,dx.$$

We conclude the coefficients, $c(x)$, of the expansion of $|\Psi\rangle$ into the basis of the eigenvectors of $\widehat{x}$ are given by the wave function, i.e.

$$c(x) = \Psi(x).$$

By the theorem above we have that

$$|c(x)|^2 = |\Psi(x)|^2$$

gives the probability density of obtaining the result x after measuring $\widehat{x}$, which is indeed how we interpreted the wave function before.

Example 5.12. Suppose $\mathcal{H} = \mathbb{C}^2$, since this is a finite dimensional Hilbert space we can represent the operators as matrices. Consider the operator

$$\widehat{\mathcal{O}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

this operator is clearly Hermitian, as $\widehat{\mathcal{O}}^\dagger = \widehat{\mathcal{O}}$. As per our discussion above we are interested in the eigenvalues and corresponding eigenvectors of a Hermitian operator.

- Finding the eigenvalues of $\widehat{\mathcal{O}}$:

$$\det \begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} = 0$$
$$\Rightarrow \lambda^2 - 1 = 0$$
$$\Rightarrow \lambda = \pm 1.$$

Therefore, the spectrum of $\widehat{\mathcal{O}}$ is given by $\{-1, +1\}$.

- Finding the eigenvectors of $\widehat{\mathcal{O}}$.

  - For $\lambda = 1$:
    recall the eigenvector equation

$$\widehat{\mathcal{O}} \left| \Psi \right\rangle = \lambda \left| \Psi \right\rangle,$$

we choose our ansatz as

$$\left| \Psi \right\rangle = \begin{pmatrix} a \\ b \end{pmatrix}.$$

Substituting into the ansatz into eigenvector equation we have

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = 1 \cdot \begin{pmatrix} a \\ b \end{pmatrix}$$
$$\begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}.$$

We conclude that $a = b$ thus,

$$\left| \Psi \right\rangle = \begin{pmatrix} a \\ a \end{pmatrix}.$$

To fix a we impose that $\left| \Psi \right\rangle$ must be normalised so,

$$1 = \left\langle \Psi | \Psi \right\rangle$$
$$= \left(\left| \Psi \right\rangle \right)^\dagger \left| \Psi \right\rangle$$
$$= \begin{pmatrix} a^* & a^* \end{pmatrix} \begin{pmatrix} a \\ a \end{pmatrix}$$
$$= 2|a|^2.$$

We choose $a = \frac{1}{\sqrt{2}}$ (we could have chosen any value $a \in \mathbb{C}$ such that $2|a|^2 = 1$). We conclude that

$$\left| \Psi \right\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

By the discussion above, we recall that the eigenvector with eigenvalue λ is denoted by $|\lambda\rangle$ so, this eigenvector is denoted as

$$|+1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

– For $\lambda = -1$: By a similar process as above we conclude

$$|-1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

By the spectral theorem we know the eigenvectors, $|+1\rangle$ and $|-1\rangle$ form a basis of $\widehat{\mathcal{O}}$:

$$|\Psi\rangle = c_{+1} |+1\rangle + c_{-1} |-1\rangle.$$

We recall that the coefficients are given as follows:

$$\begin{aligned} c_{+1} &= \langle +1|\Psi\rangle \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \frac{1}{\sqrt{2}}, \end{aligned}$$

and

$$\begin{aligned} c_{-1} &= \langle -1|\Psi\rangle \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} \end{aligned}$$

so,

$$|\Psi\rangle = \frac{1}{\sqrt{2}} |+1\rangle + \frac{1}{\sqrt{2}} |-1\rangle.$$

Suppose we measure $\widehat{\mathcal{O}}$ in the state $|\Psi\rangle$, and we obtain $\lambda = +1$ then, the state changes from

$$|\Psi\rangle \to |+1\rangle,$$

and the probability of obtaining $\lambda = +1$ is

$$\left| \frac{1}{\sqrt{2}} \right|^2 = \frac{1}{2}.$$

6   Unitary and time evolution

Recall the time-dependent Schrödinger equation, for a wave function which depends on both x and t:

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \Psi(x,t) &= \left(\widehat{H} \Psi \right)(x,t) \\ &= -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi(x,t) + U(x)\Psi(x,t). \end{aligned}$$

We now want to interpret this equation with the bra-ket notation, for this we need a Hilbert space.

> **Theorem 6.1**
>
> The time-dependent Schrödinger equation for any $\mathcal{H}$ is given by
>
> $$i\hbar \frac{\partial}{\partial t} |\Psi(t)\rangle = \widehat{H} |\Psi(t)\rangle.$$

Remark 6.2. We can write the Schrödinger equation in this form for any $\mathcal{H}$ in which the operator $\widehat{H}$ is defined.

The Schrödinger is powerful as by just an initial value for $|\Psi(t)\rangle$ we can determine $|\Psi\rangle$ at a (small) later time. Consider $|\Psi(t + \delta t)\rangle$ for small δt then, by the Taylor expansion we have

$$|\Psi(t + \delta t)\rangle \approx |\Psi(t)\rangle + \delta t \frac{\partial}{\partial t} |\Psi(t)\rangle$$

$$\approx |\Psi(t)\rangle - \delta t \frac{i}{\hbar} \widehat{H} |\Psi(t)\rangle.$$

Example 6.3. Suppose $\mathcal{H} = \mathbb{C}$ then, $|\Psi\rangle \to \Psi(t) \in \mathbb{C}$ (i.e. it is a number) and similarly $\widehat{H}$ becomes a 1×1 matrix (again just a number in $\mathbb{C}$). As such we can solve the Schrödinger equation:

$$i\hbar \frac{\partial}{\partial t} \Psi(t) = H\Psi(t),$$

and we have

$$\Psi(t) = Ce^{\frac{-iH}{\hbar}t}.$$

To fix the constant C we need a condition for Ψ so, we suppose that at $t = 0$ we have $\Psi(0)$. Then, $C = \Psi(0)$ and we have

$$\Psi(t) = \Psi(0)e^{\frac{-iH}{\hbar}t}$$

> **Theorem 6.4**
>
> The time-dependent Schrödinger equation is solved by
>
> $$|\Psi(t)\rangle = e^{-\frac{i\widehat{H}}{\hbar}t} |\Psi(0)\rangle.$$

Remark 6.5. We can think of the $e^{-\frac{i\widehat{H}}{\hbar}t}$ term as adding time t to $|\Psi(0)\rangle$.

7 Unitary operators and time evolution

7.1 Exponential of operators

Definition 7.1. If $\widehat{\mathcal{O}}$ is an operator then we define

$$e^{\widehat{\mathcal{O}}} = \sum_{n=0}^{\infty} \frac{\widehat{\mathcal{O}}^n}{n!}$$

$$= \mathbb{1} + \widehat{\mathcal{O}} + \frac{\widehat{\mathcal{O}}^2}{2!} + \cdots + \frac{\widehat{\mathcal{O}}^n}{n!} + \cdots,$$

where $\mathbb{1}$ represents the identity operator.

> **Note 7.2.** This definition is obtained from the Taylor series of e^x.

Example 7.3. Suppose $\mathcal{H} = \mathbb{C}^2$ and let $\widehat{\mathcal{O}} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ then,

$$
\begin{aligned}
e^{\widehat{\mathcal{O}}} &= \mathbb{1} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \cdots \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.
\end{aligned}
$$

Since, we are in a finite dimensional vector space we can identify operators with matrices thus, $\mathbb{1}$ represents the identity matrix. Furthermore, we notice that $\widehat{\mathcal{O}}^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ which implies $\widehat{\mathcal{O}}^n = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ as well.

Example 7.4

Suppose $\mathcal{H} = \mathbb{C}^2$ and let $\widehat{\mathcal{O}} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ then

$$
\begin{aligned}
e^{t\widehat{\mathcal{O}}} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + t \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} + \cdots + \frac{t^n}{n!} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}^n + \cdots \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} t & 0 \\ 0 & 2t \end{pmatrix} + \cdots + \begin{pmatrix} \frac{t^n}{n!} & 0 \\ 0 & \frac{(2t)^n}{n!} \end{pmatrix} + \cdots
\end{aligned}
$$

Consider the first entry of $e^{t\widehat{\mathcal{O}}}$:

$$
1 + t + \frac{t^2}{2!} + \cdots \frac{t^n}{n!} + \cdots = e^t
$$

therefore,

$$
e^{t\widehat{\mathcal{O}}} = \begin{pmatrix} e^t & 0 \\ 0 & e^{2t} \end{pmatrix}.
$$

Generally, if $\widehat{\mathcal{O}} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ then,

$$
e^{t\widehat{\mathcal{O}}} = \begin{pmatrix} e^{at} & 0 \\ 0 & e^{bt} \end{pmatrix}.
$$

Theorem 7.5

Let $\widehat{\mathcal{O}}$ be a Hermitian operator and let $\{\lambda_n\}_{n=1}^{\infty}$ be the eigenvector basis of $\widehat{\mathcal{O}}$. We define $e^{\alpha\widehat{\mathcal{O}}}$ by its action on all $|\Psi\rangle \in \mathcal{H}$: for all

$$|\Psi\rangle = \sum_{n=1}^{\infty} c_n |\lambda_n\rangle$$

with

$$e^{\alpha\widehat{\mathcal{O}}} |\Psi\rangle = \sum_{n=1}^{\infty} c_n e^{\alpha\lambda_n} |\lambda_n\rangle.$$

Proof. First notice that $\widehat{\mathcal{O}}^n |v_k\rangle = \lambda_k^n |\lambda_k\rangle$ then, using the Taylor series definition we have:

$$e^{\alpha\widehat{\mathcal{O}}} = \left(\mathbb{1} + \alpha\widehat{\mathcal{O}} + \frac{(\alpha\widehat{\mathcal{O}})^2}{2!} + \cdots\right) |\Psi\rangle$$

$$= \left(\mathbb{1} + \alpha\widehat{\mathcal{O}} + \frac{(\alpha\widehat{\mathcal{O}})^2}{2!} + \cdots\right) \sum_{n=1}^{\infty} c_n |\lambda_n\rangle$$

$$= \sum_{n=1}^{\infty} c_n \left(\mathbb{1} + \alpha\widehat{\mathcal{O}} + \frac{(\alpha\widehat{\mathcal{O}})^2}{2!} + \cdots\right) |\lambda_n\rangle$$

$$= \sum_{n=1}^{\infty} c_n \left(1 + \alpha\lambda_n + \frac{(\alpha\lambda_n)^2}{2!} + \cdots\right) |\lambda_n\rangle$$

$$= \sum_{n=1}^{\infty} c_n e^{\alpha\lambda_n} |\lambda_n\rangle.$$

$\square$

Remark 7.6. For simplicity, we have assumed the spectrum of $\widehat{\mathcal{O}}$ is discrete and that all the eigenvalues have multiplicity 1.

Theorem 7.7 (Properties of exponentials of operators)

For any operator $\widehat{\mathcal{O}}$ in a Hilbert space $\mathcal{H}$ and $\alpha \in \mathbb{C}$ we have:

- $\dfrac{\partial}{\partial\alpha} e^{\alpha\widehat{\mathcal{O}}} = \widehat{\mathcal{O}} e^{\alpha\widehat{\mathcal{O}}} = e^{\alpha\widehat{\mathcal{O}}} \widehat{\mathcal{O}}$.

- $e^{\alpha\widehat{\mathcal{O}}} e^{\beta\widehat{\mathcal{O}}} = e^{(\alpha+\beta)\widehat{\mathcal{O}}}$.

Proof. We prove each point.

- By considering the Taylor series of $e^{\alpha\widehat{\mathcal{O}}}$:

$$
\begin{aligned}
\frac{\partial}{\partial\alpha}e^{\alpha\widehat{\mathcal{O}}} &= \frac{\partial}{\partial\alpha}\left(\mathbb{1} + \alpha\widehat{\mathcal{O}} + \frac{(\alpha\widehat{\mathcal{O}})^2}{2!} + \cdots\right) \\
&= 0 + \widehat{\mathcal{O}} + \alpha\widehat{\mathcal{O}}^2 + \cdots \\
&= \widehat{\mathcal{O}}\left(\mathbb{1} + \alpha\widehat{\mathcal{O}} + \frac{(\alpha\widehat{\mathcal{O}})^2}{2!} + \cdots\right) \\
&= \left(\mathbb{1} + \alpha\widehat{\mathcal{O}} + \frac{(\alpha\widehat{\mathcal{O}})^2}{2!} + \cdots\right)\widehat{\mathcal{O}} \\
&= \widehat{\mathcal{O}}e^{\alpha\widehat{\mathcal{O}}} \\
&= e^{\alpha\widehat{\mathcal{O}}}\widehat{\mathcal{O}}.
\end{aligned}
$$

- We present an easier proof. Suppose $\{|\lambda_n\rangle\}$ is a basis of $\widehat{\mathcal{O}}$ then, we compute

$$
\begin{aligned}
e^{\alpha\widehat{\mathcal{O}}}e^{\beta\widehat{\mathcal{O}}}|\Psi\rangle &= e^{\alpha\widehat{\mathcal{O}}}e^{\beta\widehat{\mathcal{O}}}\sum_{n=1}^{\infty}c_n|\lambda_n\rangle \\
&= e^{\alpha\widehat{\mathcal{O}}}\sum_{n=1}^{\infty}c_n e^{\beta\lambda_n}|\lambda_n\rangle \\
&= \sum_{n=1}^{\infty}c_n e^{\beta\lambda_n}e^{\alpha\lambda_n}|\lambda_n\rangle \\
&= \sum_{n=1}^{\infty}c_n e^{(\alpha+\beta)\lambda_n}|\lambda_n\rangle \\
&= e^{(\alpha+\beta)\widehat{\mathcal{O}}}\sum_{n=1}^{\infty}c_n|\lambda_n\rangle \\
&= e^{(\alpha+\beta)\widehat{\mathcal{O}}}|\Psi\rangle.
\end{aligned}
$$

$\square$

7.2   Evolution operator

In this section we verify that

$$
|\Psi(t)\rangle = e^{-\frac{i\widehat{H}}{\hbar}t}|\Psi(0)\rangle
$$

solves the time-dependent Schrödinger equation:

$$
i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = \widehat{H}|\Psi(t)\rangle.
$$

We compute

$$
\begin{aligned}
\frac{\partial}{\partial t}|\Psi(t)\rangle &= \frac{\partial}{\partial t}e^{-\frac{i\widehat{H}}{\hbar}t}|\Psi(0)\rangle \\
&= -\frac{i\widehat{H}}{\hbar}e^{-\frac{i\widehat{H}}{\hbar}t}|\Psi(0)\rangle.
\end{aligned}
$$

Clearly,

$$i\hbar\left(-\frac{i\widehat{H}}{\hbar}e^{-\frac{i\widehat{H}}{\hbar}t}\left|\Psi(0)\right\rangle\right) = -i^2\widehat{H}e^{-\frac{i\widehat{H}}{\hbar}t}\left|\Psi(0)\right\rangle$$

$$= \widehat{H}\left|\Psi\right\rangle.$$

Definition 7.8. We call the operator $e^{-\frac{i\widehat{H}}{\hbar}t}$ the **time evolution operator** and is denoted by

$$\widehat{U}(t) = e^{-\frac{i\widehat{H}}{\hbar}t}.$$

Theorem 7.9

Properties of $\widehat{U}(t)$:

- $\widehat{U}(t_2)\widehat{U}(t_1) = \widehat{U}(t_1 + t_2).$

- $\left(\widehat{U}(t)\right)^\dagger = \widehat{U}(-t).$

Proof. We split the proof in parts:

- By the definition of $\widehat{U}(t)$ the result follows:

$$\widehat{U}(t_2)\widehat{U}(t_1) = e^{-\frac{i\widehat{H}}{\hbar}t_2}e^{-\frac{i\widehat{H}}{\hbar}t_1}$$

$$= e^{-\frac{i\widehat{H}}{\hbar}(t_1+t_2)}$$

$$= \widehat{U}(t_1 + t_2).$$

- Firstly, note that $\widehat{H}$ is Hermitian and $\left(\widehat{\mathcal{O}}^n\right)^\dagger = \left(\widehat{\mathcal{O}}^\dagger\right)^n$ Using the series definition we have:

$$\left(\widehat{U}(t)\right)^\dagger = \left(e^{-\frac{i\widehat{H}}{\hbar}t}\right) = \left(\mathbb{1} + \left(-\frac{i\widehat{H}}{\hbar}\right) + \frac{1}{2}\left(-\frac{i\widehat{H}}{\hbar}\right)^2 + \frac{1}{3!}\left(-\frac{i\widehat{H}}{\hbar}\right)^3 + \cdots\right)^\dagger$$

$$= \mathbb{1} + \left(-\frac{i\widehat{H}}{\hbar}\right)^\dagger + \frac{1}{2}\left(\left(-\frac{i\widehat{H}}{\hbar}\right)^2\right)^\dagger + \cdots$$

$$= \mathbb{1} + \frac{i\widehat{H}}{\hbar} + \frac{1}{2}\left(\frac{i\widehat{H}}{\hbar}\right)^2 + \cdots$$

$$= e^{\frac{i\widehat{H}}{\hbar}t}$$

$$= \widehat{U}(-t).$$

$\square$

7.3   Unitary

Definition 7.10. An operator, $\widehat{A}$, which satisfies the property

$$\widehat{A}^\dagger \widehat{A} = \mathbb{1},$$

where $\mathbb{1}$ is the identity operator, are called **unitary**.

Theorem 7.11

The time evolution operator, $\widehat{U}(t)$ is unitary.

Proof. We know $\left(\widehat{U}(t)\right)^\dagger = \widehat{U}(-t)$ therefore,

$$
\begin{aligned}
\left(\widehat{U}(t)\right)^\dagger \widehat{U}(t) &= \widehat{U}(-t)\widehat{U}(t) \\
&= \widehat{U}(0) \\
&= \mathbb{1}.
\end{aligned}
$$

(We have used $\widehat{U}(0) = e^0 = \mathbb{1}$).   □

Theorem 7.12

Unitary operators preserve the inner product i.e.

$$\left(\widehat{U}(t)\left|\Psi(t)\right\rangle, \widehat{U}(t)\left|\Phi(t)\right\rangle\right) = \left(\left|\Psi(t)\right\rangle, \left|\Phi(t)\right\rangle\right).$$

Proof. Consider two vectors $\left|\Psi\right\rangle, \left|\Phi\right\rangle \in \mathcal{H}$, by computing the inner product we have

$$
\begin{aligned}
\left(\widehat{U}(t)\left|\Psi(t)\right\rangle, \widehat{U}\left|\Phi(t)\right\rangle\right) &= \left([\widehat{U}(t)]^\dagger \widehat{U}(t)\left|\Psi\right\rangle, \left|\Phi\right\rangle\right) \\
&= \left(\left|\Psi(t)\right\rangle, \left|\Phi(t)\right\rangle\right).
\end{aligned}
$$

□

Corollary 7.13. As a result we have

$$\left\|\left|\Psi(t)\right\rangle\right\|^2 = \left\|\left|\Psi(0)\right\rangle\right\|^2;$$

in particular if $\left\|\left|\Psi(0)\right\rangle\right\|^2 = 1$ then $\left\|\left|\Psi(t)\right\rangle\right\|^2 = 1$ for all t.

Proof. By definition

$$
\begin{aligned}
\left\|\left|\Psi(t)\right\rangle\right\|^2 = \left\langle\Psi(t)|\Psi(t)\right\rangle &= \left(\left|\Psi(t)\right\rangle, \left|\Psi(t)\right\rangle\right) \\
&= \left(\widehat{U}(t)\left|\Psi(0)\right\rangle, \widehat{U}(t)\left|\Psi(0)\right\rangle\right) \\
&= \left(\left|\Psi(0)\right\rangle, \left|\Psi(0)\right\rangle\right) \\
&= \left\|\left|\Psi(0)\right\rangle\right\|^2.
\end{aligned}
$$

□

7.4 Evolution in the energy basis

Suppose we are given the Hamiltonian $\widehat{H}$ and some initial state $|\Psi(0)\rangle$, and we want to find the state $|\Psi(t)\rangle$ at some $t \in \mathbb{R}$. We can do so by the following algorithm:

1. Find the eigenvalues, E_n, and the orthonormal basis of eigenvectors $\{|n\rangle\}_{n=0}^{\infty}$ of $\widehat{H}$.

2. Find the expression for $|\Psi(0)\rangle$ in this basis of eigenvectors,

$$|\Psi(0)\rangle = \sum_{n=0}^{\infty} c_n |n\rangle,$$

   where $c_n = \langle n|\Psi(0)\rangle$ if the basis $\{|n\rangle\}$ is orthonormal.

3. Act with $\widehat{U}(t) = e^{-\frac{i\widehat{H}}{\hbar}t}$,

$$\begin{aligned}
|\Psi(t)\rangle &= \widehat{U}(t)\,|\Psi(0)\rangle \\
&= e^{-\frac{i\widehat{H}}{\hbar}t} \sum_{n=0}^{\infty} c_n |n\rangle \\
&= \sum_{n=0}^{\infty} c_n e^{-\frac{iE_n}{\hbar}t} |n\rangle.
\end{aligned}$$

> **Remark 7.14.** This algorithm is for the case where the spectrum of $\widehat{H}$ is discrete and all the eigenvalues have multiplicity 1.

This gives us a concrete expression for $|\Psi(t)\rangle$. Of course, this algorithm is not always simple to carry out: firstly, we require a complete solution of the time-independent Schrödinger equation, which is just the eigenvector equation for $\widehat{H}$,

$$\widehat{H}\,|n\rangle = E_n\,|n\rangle.$$

Example 7.15. In the case of a particle in a box we determined the eigenvalues to be

$$E_n = \frac{\pi^2 n^2 \hbar^2}{2mL^2}$$

and the corresponding eigenvectors

$$|E_n\rangle \sim \Psi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{\pi n}{L}x\right).$$

Therefore,

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

where

$$c_n = \int_0^L \Psi_n^*(x)\Psi(x,0)\,dx.$$

8   Non-commutativity and uncertainty

8.1   Implications for measurements of operators

Let $\widehat{A}$ and $\widehat{B}$ be Hermitian operators with non-vanishing commutator:

$$[\widehat{A}, \widehat{B}] = \widehat{C} \neq 0.$$

Here $\widehat{C}$ is a new operator (not necessarily Hermitian). We now ask: is it possible for $\widehat{A}$ and $\widehat{B}$ to simultaneously take definite values in a state $|\Psi\rangle$? In other words, can $|\Psi\rangle$ be an eigenvector of $\widehat{A}$ and $\widehat{B}$ simultaneously? That is, is the following set of equations possible:

$$\widehat{A}|\Psi\rangle = a|\Psi\rangle$$
$$\widehat{B}|\Psi\rangle = b|\Psi\rangle,$$

for $a, b \in \mathbb{C}$. Suppose this is true then, we can compute

$$\begin{aligned}
\widehat{C}|\Psi\rangle &= [\widehat{A}, \widehat{B}]|\Psi\rangle \\
&= (\widehat{A}\widehat{B} - \widehat{B}\widehat{A})|\Psi\rangle \\
&= \widehat{A}\left(\widehat{B}|\Psi\rangle\right) - \widehat{B}\left(\widehat{A}|\Psi\rangle\right) \\
&= b\widehat{A}|\Psi\rangle - a\widehat{B}|\Psi\rangle \\
&= ba|\Psi\rangle - ab|\Psi\rangle \\
&= 0.
\end{aligned}$$

Therefore, $|\Psi\rangle$ is an eigenvector of $\widehat{A}$ and $\widehat{B}$ simultaneously if and only if $\widehat{C}|\Psi\rangle = 0$, i.e. $\widehat{A}$ and $\widehat{B}$ commute.

8.2   Heisenberg's uncertainty principle

Definition 8.1. The **uncertainty** of a <u>normalised</u> Hermitian operator $\widehat{A}$, denoted by $\Delta\widehat{A}$, is defined by

$$\begin{aligned}
(\Delta\widehat{A})^2 &= \left\langle \left(\widehat{A} - \langle\widehat{A}\rangle\right)^2 \right\rangle \\
&= \langle\Psi|\left(\widehat{A} - \langle\widehat{A}\rangle\right)^2|\Psi\rangle.
\end{aligned}$$

> **Note 8.2.** Operators need to be Hermitian as these are the operators which are used to measure quantities in Quantum Mechanics.

Theorem 8.3. The uncertainty of an operator $\widehat{A}$ is real and non-negative i.e. $\Delta\widehat{A} \geq 0$.

Proof. We have

$$\begin{aligned}
(\Delta\widehat{A})^2 &= \langle\Psi|\left(\widehat{A} - \langle\widehat{A}\rangle\right)^2|\Psi\rangle \\
&= \langle\Psi|(\widehat{A} - \langle\widehat{A}\rangle)(\widehat{A} - \langle\widehat{A}\rangle)|\Psi\rangle \\
&= \left[(\widehat{A} - \langle\widehat{A}\rangle)|\Psi\rangle\right]^{\dagger}(\widehat{A} - \langle\widehat{A}\rangle)|\Psi\rangle \\
&= \left\|\left(\widehat{A} - \langle\widehat{A}\rangle\right)|\Psi\rangle\right\|^2 \\
&\geq 0
\end{aligned}$$

by the axioms of norms. In the last equality we have used the fact that $\widehat{A}$ is a Hermitian operator. $\square$

Corollary 8.4. Let $\widehat{A}$ be an operator then, $\Delta\widehat{A} = 0$ in a state $|\Psi\rangle$ if and only if $|\Psi\rangle$ is an eigenvector of $\widehat{A}$.

Proof. Recall the definition of uncertainty

$$(\Delta\widehat{A})^2 = \langle\Psi| \left(\widehat{A} - \langle\widehat{A}\rangle\right)^2 |\Psi\rangle$$

$$= \left\|\left(\widehat{A} - \langle\widehat{A}\rangle\right)|\Psi\rangle\right\|^2,$$

by the axioms of the norm it is 0 if and only if

$$\left(\widehat{A} - \langle\widehat{A}\rangle\right)|\Psi\rangle = 0.$$

We can write this condition as

$$\widehat{A}|\Psi\rangle = \langle\widehat{A}\rangle|\Psi\rangle,$$

i.e. $|\Psi\rangle$ is an eigenvector (since $\langle\widehat{A}\rangle$ is a real number). $\square$

Theorem 8.5

Let $\widehat{A}$ and $\widehat{B}$ be Hermitian operators and $\Delta\widehat{A}$ and $\Delta\widehat{B}$ their respective uncertainty in a state $|\Psi\rangle$. Then, we have

$$\Delta\widehat{A}\,\Delta\widehat{B} \geq \frac{1}{2}\left|\left\langle\left[\widehat{A}, \widehat{B}\right]\right\rangle\right|.$$

Proof. Define

$$\widehat{A}_0 = \widehat{A} - \langle\widehat{A}\rangle \quad \text{and} \quad \widehat{B}_0 = \widehat{B} - \langle\widehat{B}\rangle.$$

We first compute

$$[\widehat{A}_0, \widehat{B}_0] = [\widehat{A} - \langle\widehat{A}\rangle, \widehat{B} - \langle\widehat{B}\rangle]$$

$$= [\widehat{A}, \widehat{B} - \langle\widehat{B}\rangle] - [\langle\widehat{A}\rangle, \widehat{B} - \langle\widehat{B}\rangle]$$

$$= [\widehat{A}, \widehat{B}] - [\widehat{A}, \langle\widehat{B}\rangle] - \left([\langle\widehat{A}\rangle, \widehat{B}] - [\langle\widehat{A}\rangle, \langle\widehat{B}\rangle]\right)$$

$$= [\widehat{A}, \widehat{B}] + [\langle\widehat{B}\rangle, \widehat{A}] + [\widehat{B}, \langle\widehat{A}\rangle] + [\langle\widehat{A}\rangle, \langle\widehat{B}\rangle].$$

Notice that $\langle\widehat{A}\rangle$ and $\langle\widehat{B}\rangle$ are real numbers since $\widehat{A}$ and $\widehat{B}$ are Hermitian operators thus, their commutators with anything are zero. As such we have that

$$[\widehat{A}_0, \widehat{B}_0] = [\widehat{A}, \widehat{B}].$$

Furthermore, it follows that

$$(\Delta\widehat{A})^2 = \langle\widehat{A}_0^2\rangle \quad \text{and} \quad (\Delta\widehat{B})^2 = \langle\widehat{B}_0^2\rangle.$$

Consider

$$
\begin{aligned}
(\Delta \widehat{A})^2 (\Delta \widehat{B})^2 &= \langle \Psi | \, \widehat{A}_0^2 \, | \Psi \rangle \, \langle \Psi | \, \widehat{B}_0^2 \, | \Psi \rangle \\
&= (\widehat{A}_0 \, | \Psi \rangle)^\dagger (\widehat{A}_0 \, | \Psi \rangle) \cdot (\widehat{B}_0 \, | \Psi \rangle)^\dagger (\widehat{B}_0 \, | \Psi \rangle) \\
&= \left\| \widehat{A}_0 \, | \Psi \rangle \right\|^2 \left\| \widehat{B}_0 \, | \Psi \rangle \right\|^2 .
\end{aligned}
$$

Now, recall the Cauchy-Schwarz inequality

$$
| \langle u | v \rangle |^2 \le \| \, | u \rangle \, \|^2 \, \| \, | v \rangle \, \|^2 .
$$

Using this in the above expression and, we obtain

$$
\begin{aligned}
(\Delta \widehat{A})^2 (\Delta \widehat{B})^2 &= \left\| \widehat{A}_0 \, | \Psi \rangle \right\|^2 \left\| \widehat{B}_0 \, | \Psi \rangle \right\|^2 \\
&\ge \left| \langle \Psi | \, \widehat{A}_0 \widehat{B}_0 \, | \Psi \rangle \right|^2 .
\end{aligned}
$$

Using the fact that for any complex number z we have $|z|^2 \ge |\mathrm{Im}(z)|^2 = \frac{1}{4} |z - z^*|^2$, we get

$$
\begin{aligned}
(\Delta \widehat{A})^2 (\Delta \widehat{B})^2 &= \left\| \widehat{A}_0 \, | \Psi \rangle \right\|^2 \left\| \widehat{B}_0 \, | \Psi \rangle \right\|^2 \\
&\ge \left| \langle \Psi | \, \widehat{A}_0 \widehat{B}_0 \, | \Psi \rangle \right|^2 \\
&\ge \frac{1}{4} \left| \langle \Psi | \, \widehat{A}_0 \widehat{B}_0 \, | \Psi \rangle - \left(\langle \Psi | \, \widehat{A}_0 \widehat{B}_0 \, | \Psi \rangle \right)^* \right|^2 .
\end{aligned}
$$

Using the fact

$$
\begin{aligned}
\left(\langle \Psi | \, \widehat{A}_0 \widehat{B}_0 \, | \Psi \rangle \right)^* &= \left(\langle \Psi | \, \widehat{A}_0 \widehat{B}_0 \, | \Psi \rangle \right)^\dagger \\
&= | \Psi \rangle^\dagger \, \widehat{B}_0^\dagger \widehat{A}_0^\dagger \, \langle \Psi |^\dagger \\
&= \langle \Psi | \, \widehat{B}_0 \widehat{A}_0 \, | \Psi \rangle ,
\end{aligned}
$$

$$
\begin{aligned}
(\Delta \widehat{A})^2 (\Delta \widehat{B})^2 &\ge \frac{1}{4} \left| \langle \Psi | \, \widehat{A}_0 \widehat{B}_0 \, | \Psi \rangle - \langle \Psi | \, \widehat{B}_0 \widehat{A}_0 \, | \Psi \rangle \right|^2 \\
&= \frac{1}{4} \left| \langle \Psi | \left[\widehat{A}_0, \widehat{B}_0 \right] | \Psi \rangle \right|^2 \\
&= \frac{1}{4} \left| \langle \Psi | \left[\widehat{A}, \widehat{B} \right] | \Psi \rangle \right|^2 .
\end{aligned}
$$

Taking square roots of both sides we obtain

$$
\Delta \widehat{A} \, \Delta \widehat{B} \ge \frac{1}{2} \left| \left\langle \left[\widehat{A}, \widehat{B} \right] \right\rangle \right| .
$$

$\square$

Example 8.6

Applying this to $\widehat{A} = \widehat{x}$ and $\widehat{B} = \widehat{p}$ we have

$$
\begin{aligned}
\Delta \widehat{x} \, \Delta \widehat{p} &\ge \frac{1}{2} | \langle [\widehat{x}, \widehat{p}] \rangle | \\
&= \frac{1}{2} | \langle i\hbar \rangle | \\
&= \frac{\hbar}{2} .
\end{aligned}
$$

9 Harmonic oscillator

9.1 Classical solution

Definition 9.1. A system with Hamiltonian

$$H = \frac{p^2}{2m} + \frac{m\omega^2 x}{2}$$

is known as a **harmonic oscillator**.

Example 9.2. A particle attached to a string abides this Hamiltonian.

We now solve for the equation of motion and momentum the harmonic oscillator. The Hamilton equations of motion for this system are

$$\dot{x} = \frac{\partial H}{\partial p} = \frac{p}{m},$$

$$\dot{p} = -\frac{\partial H}{\partial x} = -m\omega^2 x.$$

Differentiating the first equation we can substitute an expression for $\dot{p}$ and so, we have

$$\ddot{x} = -\omega^2 x$$

hence, the general solution is of the form

$$x(t) = Ae^{i\omega t} + A^* e^{-i\omega t}.$$

> **Note 9.3.** We use A^* as the coefficient of the second term to ensure that $x(t)$ is real.

Similarly, we have that

$$p(t) = m\dot{x}$$
$$= im\omega \left(Ae^{i\omega t} + A^* e^{-i\omega t}\right).$$

We now consider a particular linear combination of $x(t)$ and $p(t)$, namely

$$a(t) = \frac{p(t) - im\omega\, x(t)}{\sqrt{2m\omega\hbar}} = -i\left(\sqrt{\frac{2m\omega}{\hbar}}\right) A^* e^{-i\omega t}$$

$$a^*(t) = \frac{p(t) + im\omega\, x(t)}{\sqrt{2m\omega\hbar}} = i\left(\sqrt{\frac{2m\omega}{\hbar}}\right) Ae^{i\omega t}.$$

These quantities satisfy the first-order equations

$$\frac{\partial}{\partial t}a(t) = -i\omega a(t) \quad \text{and} \quad \frac{\partial}{\partial t}a^*(t) = i\omega a^*(t).$$

These quantities play a prominent role in the 'quantum solution' of the Harmonic oscillator.

9.2   Quantum solution

Definition 9.4. A system with Hamiltonian operator

$$\widehat{H} = \frac{\widehat{p}^2}{2m} + \frac{m\omega^2\widehat{x}^2}{2}$$

is called a **harmonic oscillator**.

As alluded to in the previous section, it will be useful to consider the operators:

$$\widehat{a} = \frac{\widehat{p} - im\omega\,\widehat{x}}{\sqrt{2m\omega\hbar}} \quad \text{and} \quad \widehat{a}^\dagger = \frac{\widehat{p} + im\omega\,\widehat{x}}{\sqrt{2m\omega\hbar}}.$$

Proposition 9.5. The value of

$$[\widehat{a}, \widehat{a}^\dagger] = 1.$$

Proof. Use properties of commutators and the fact $[\widehat{x}, \widehat{p}] = i\hbar$.   □

Theorem 9.6

The Hamiltonian operator, $\widehat{H} = \frac{\widehat{p}^2}{2m} + \frac{m\omega^2\widehat{x}^2}{2}$ can be expressed as

$$\widehat{H} = \hbar\omega\left(\widehat{a}^\dagger\widehat{a} + \frac{1}{2}\right).$$

Proof. Consider the following

$$\begin{aligned}
\hbar\omega\widehat{a}^\dagger\widehat{a} &= \hbar\omega\left(\frac{\widehat{p} + im\omega\,\widehat{x}}{\sqrt{2m\omega\hbar}}\right)\left(\frac{\widehat{p} - im\omega\,\widehat{x}}{\sqrt{2m\omega\hbar}}\right) \\
&= \frac{1}{2m}(\widehat{p} + im\omega\,\widehat{x})(\widehat{p} + im\omega\,\widehat{x}) \\
&= \frac{1}{2m}(\widehat{p}^2 - im\omega\widehat{p}\widehat{x} + im\omega\widehat{x}\widehat{p} + m^2\omega^2\widehat{x}^2) \\
&= \frac{1}{2m}(\widehat{p}^2 + im\omega\widehat{x}\widehat{p} - im\omega\widehat{p}\widehat{x} + m^2\omega^2\widehat{x}^2) \\
&= \frac{1}{2m}(\widehat{p}^2 + m^2\omega^2\widehat{x}^2 + im\omega[\widehat{x}, \widehat{p}]) \\
&= \frac{\widehat{p}^2}{2m} + \frac{m\omega^2\widehat{x}^2}{2} - \frac{\hbar\omega}{2}.
\end{aligned}$$

Recall that $\widehat{H} = \frac{\widehat{p}^2}{2m} + \frac{m\omega^2\widehat{x}^2}{2}$ so,

$$\hbar\omega\widehat{a}^\dagger\widehat{a} = \widehat{H} - \frac{\hbar\omega}{2}.$$

As such we can rewrite,

$$\widehat{H} = \hbar\omega\left(\widehat{a}^\dagger\widehat{a} + \frac{1}{2}\right).$$

□

We would like to determine the spectrum and the eigenstates of $\widehat{H}$ i.e., solve the time-independent Schrödinger equation

$$\widehat{H}\left|\Psi\right\rangle = E\left|\Psi\right\rangle.$$

We can write

$$\widehat{H} = \hbar\omega\left(\widehat{N} + \frac{1}{2}\right) \quad \text{where } \widehat{N} = \widehat{a}^\dagger\widehat{a}.$$

Proposition 9.7. Properties of $\widehat{N} = \widehat{a}^\dagger\widehat{a}$.

- $\widehat{N}$ is Hermitian i.e. $\widehat{N}^\dagger = \widehat{N}$;

- $\left|\Psi\right\rangle$ is an eigenvector of $\widehat{H}$ if and only if $\left|\Psi\right\rangle$ is an eigenvector of $\widehat{N}$.

> **Note 9.8.** The eigenvectors are the same, but the eigenvalues are different.

Proof. We prove each bullet point in turn.

- Proof is trivial.

- Suppose $\left|\Psi\right\rangle$ is an eigenvector of $\widehat{N}$ with eigenvalue n i.e. $\widehat{N}\left|\Psi\right\rangle = n\left|\Psi\right\rangle$. Then, consider

$$\widehat{H}\left|\Psi\right\rangle = \hbar\omega\widehat{N}\left|\Psi\right\rangle + \frac{\hbar\omega}{2}\left|\Psi\right\rangle$$
$$= \hbar\omega\left(n + \frac{1}{2}\right)\left|\Psi\right\rangle.$$

Therefore, $\left|\Psi\right\rangle$ is an eigenvector of $\widehat{H}$ with eigenvalue $\hbar\omega\left(n + \frac{1}{2}\right)$. The proof of the converse direction follows similarly.

$\square$

> **Lemma 9.9**
>
> If λ is an eigenvalue of $\widehat{N}$ then $\lambda \geq 0$.

Proof. Let $\left|\Psi\right\rangle$ be an eigenvector of $\widehat{N}$ with eigenvalue λ then,

$$\widehat{N}\left|\Psi\right\rangle = \lambda\widehat{\Psi} \Rightarrow \left\langle\Psi\right|\widehat{N}\left|\Psi\right\rangle = \lambda\left\langle\Psi|\Psi\right\rangle.$$

Thus, we can write

$$\lambda = \frac{\left\langle\Psi\right|\widehat{N}\left|\Psi\right\rangle}{\left\langle\Psi|\Psi\right\rangle} = \frac{\left\langle\Psi\right|\widehat{a}^\dagger\widehat{a}\left|\Psi\right\rangle}{\left\langle\Psi|\Psi\right\rangle}$$
$$= \frac{\left(\widehat{a}\left|\Psi\right\rangle\right)^\dagger\left(\widehat{a}\left|\Psi\right\rangle\right)}{\left\langle\Psi|\Psi\right\rangle}$$
$$= \frac{\left\|\widehat{a}\left|\Psi\right\rangle\right\|^2}{\left\langle\Psi|\Psi\right\rangle}.$$

(We have used the fact $\widehat{a}\left|\Psi\right\rangle$ is a number thus, $\left(\widehat{a}\left|\Psi\right\rangle\right)^\dagger\left(\widehat{a}\left|\Psi\right\rangle\right) = \left\|\widehat{a}\left|\Psi\right\rangle\right\|^2$). Since, $\left\langle\Psi|\Psi\right\rangle = \left\|\left|\Psi\right\rangle\right\|^2 > 0$ (since $\left|\Psi\right\rangle \neq 0$) and $\left\|\widehat{a}\left|\Psi\right\rangle\right\|^2 \geq 0$, we conclude $\lambda \geq 0$. $\square$

> **Lemma 9.10**
>
> Let $|\Psi\rangle$ be an eigenvector of $\widehat{N}$ with the eigenvalue λ. Then exactly one of the statements holds:
>
> 1. $\lambda = 0$ <u>and</u> $\widehat{a}|\Psi\rangle = 0$ OR,
>
> 2. $\lambda \neq 0$, $\widehat{a}|\Psi\rangle \neq 0$ and $\widehat{a}|\Psi\rangle$ is an eigenvector of $\widehat{N}$ with eigenvalue $\lambda - 1$.

Proof. First, we compute

$$
\begin{aligned}
[\widehat{N}, \widehat{a}] = [\widehat{a}^\dagger \widehat{a}, \widehat{a}] &= \widehat{a}^\dagger \widehat{a}\widehat{a} - \widehat{a}\widehat{a}^\dagger \widehat{a} \\
&= \left(\widehat{a}^\dagger \widehat{a} - \widehat{a}\widehat{a}^\dagger\right)\widehat{a} \\
&= [\widehat{a}^\dagger, \widehat{a}]\widehat{a} \\
&= -[\widehat{a}, \widehat{a}^\dagger]\widehat{a} \\
&= -\widehat{a}.
\end{aligned}
$$

Now, suppose $|\Psi\rangle$ is an eigenvector of $\widehat{N}$ with eigenvalue λ i.e.,

$$
\widehat{N}|\Psi\rangle = \lambda|\Psi\rangle.
$$

Notice that $\widehat{N}\widehat{a} - \widehat{a}\widehat{N} = [\widehat{N}, \widehat{a}]$ allows us to write, $\widehat{N}\widehat{a} = \widehat{a}\widehat{N} + [\widehat{N}, \widehat{a}]$. Consider the vector $\widehat{a}|\Psi\rangle$ and act on it with $\widehat{N}$:

$$
\begin{aligned}
\widehat{N}\widehat{a}|\Psi\rangle &= \left(\widehat{a}\widehat{N} + [\widehat{N}, \widehat{a}]\right)|\Psi\rangle \\
&= (\widehat{a}\lambda - \widehat{a})|\Psi\rangle \\
&= (\lambda - 1)\widehat{a}|\Psi\rangle.
\end{aligned}
$$

Thus, we have two possibilities.

1. $\widehat{a}|\Psi\rangle = 0$.
   Supposing this is true then,
   $$
   \begin{aligned}
   \widehat{a}|\Psi\rangle &= 0 \\
   \Rightarrow \widehat{a}^\dagger \widehat{a}|\Psi\rangle &= 0 \\
   \Rightarrow \widehat{N}|\Psi\rangle &= 0.
   \end{aligned}
   $$
   But, since we assumed $\widehat{N}|\Psi\rangle = \lambda|\Psi\rangle$ then, $|\Psi\rangle \neq 0$ (by the definition of an eigenvector) so, $\lambda = 0$. Conversely, if $\lambda = 0$ then, $0 = \lambda|\Psi\rangle = \widehat{N}|\Psi\rangle$ and thus
   $$
   0 = \langle\Psi|\widehat{N}|\Psi\rangle = \langle\Psi|\widehat{a}^\dagger \widehat{a}|\Psi\rangle = \|\widehat{a}|\Psi\rangle\|^2
   $$
   which implies (by the axioms of the norm) $\widehat{a}|\Psi\rangle = 0$.

2. $\widehat{a}|\Psi\rangle$ is an eigenvector of $\widehat{N}$ with eigenvalue $\lambda - 1$.

$\square$

> **Lemma 9.11**
>
> Let $|\Psi\rangle$ be an eigenvector of $\widehat{N}$ with eigenvalue λ. Then, $\lambda \in \mathbb{Z}_{\geq 0}$ and $|\Psi\rangle$, $\widehat{a}\,|\Psi\rangle$, $\widehat{a}^2\,|\Psi\rangle, \ldots, \widehat{a}^\lambda\,|\Psi\rangle$ are eigenstates of $\widehat{N}$ with the eigenvalues $\lambda, \lambda - 1, \lambda - 2, \ldots, 0$ respectively. Finally, $\widehat{a}^{\lambda+1}\,|\Psi\rangle = 0$.

Proof. By iteration of the previous lemma we get the infinite sequence of eigenvalues and their corresponding eigenstates.

We prove $\lambda \in \mathbb{Z}_{\geq 0}$.

For the sake of contradiction, assume λ is not a non-negative integer. Then, none of the numbers $\{\lambda, \lambda - 1, \lambda - 2, \cdots\}$ (which are eigenvalues) are 0. This means, from one eigenvalue we get an infinite sequence of eigenvalues $\{\lambda, \lambda - 1, \lambda - 2, \cdots\}$ however, at eventually we will get negative eigenvalues which is forbidden by a previous lemma. Whereas, if λ is a non-negative integer we get a finite sequence of eigenvalues which terminates at 0. Hence, λ must be a non-negative integer. $\qquad\square$

> **Lemma 9.12**
>
> Let $|\Psi\rangle$ be an eigenvector of $\widehat{N}$ with the eigenvalue λ. Then, $\widehat{a}^\dagger\,|\Psi\rangle \neq 0$ and $\widehat{a}^\dagger\,|\Psi\rangle$ is an eigenvector of $\widehat{N}$ with eigenvalue $\lambda + 1$.

> **Note 9.13.** To remember that $\widehat{a}^\dagger$ increases the eigenvalue by 1 we can think of $\dagger$ as $+$.

Proof. Note that $[\widehat{A}, \widehat{B}]^\dagger = -[\widehat{A}^\dagger, \widehat{B}^\dagger]$ so, applying this identity to $[\widehat{N}, \widehat{a}] = -\widehat{a}$, and we have

$$[\widehat{N}, \widehat{a}]^\dagger = -\widehat{a}^\dagger = -[\widehat{N}^\dagger, \widehat{a}^\dagger]$$
$$= -[\widehat{N}, \widehat{a}^\dagger].$$

We conclude, $[\widehat{N}, \widehat{a}^\dagger] = \widehat{a}^\dagger$. Furthermore, notice that $\widehat{N}\widehat{a} - \widehat{a}\widehat{N} = [\widehat{N}, \widehat{a}]$ which allows us to write, $\widehat{N}\widehat{a} = \widehat{a}\widehat{N} + [\widehat{N}, \widehat{a}]$. Suppose $|\Psi\rangle$ is an eigenvector of $\widehat{N}$ with eigenvalue λ. We can compute

$$\widehat{N}\widehat{a}^\dagger\,|\Psi\rangle = \left(\widehat{a}^\dagger\widehat{N} + [\widehat{N}, \widehat{a}^\dagger]\right)|\Psi\rangle$$
$$= (\widehat{a}^\dagger\widehat{N} + \widehat{a}^\dagger)\,|\Psi\rangle$$
$$= \widehat{a}^\dagger\widehat{N}\,|\Psi\rangle + \widehat{a}^\dagger\,|\Psi\rangle$$
$$= \widehat{a}^\dagger\lambda\,|\Psi\rangle + \widehat{a}^\dagger\,|\Psi\rangle$$
$$= (\lambda + 1)\widehat{a}^\dagger\,|\Psi\rangle.$$

Thus, we have two possibilities.

1. $\widehat{a}^\dagger\,|\Psi\rangle = 0$.
   Supposing this is true then also $\widehat{a}\widehat{a}^\dagger\,|\Psi\rangle = 0$. But we can compute

$$0 = \widehat{a}\widehat{a}^\dagger\,|\Psi\rangle = \left(\widehat{a}^\dagger\widehat{a} + [\widehat{a}, \widehat{a}^\dagger]\right)|\Psi\rangle$$
$$= (\widehat{N} + 1)\,|\Psi\rangle$$
$$= (\lambda + 1)\,|\Psi\rangle.$$

Since λ is an eigenvalue of $\widehat{N}$ by a previous lemma we have that $\lambda \geq 0$ so, $\lambda + 1 \neq 0$. However, this implies that $|\Psi\rangle = 0$, which contradicts the assumption that $|\Psi\rangle$ is an eigenvector (as by definition of an eigenvector $|\Psi\rangle \neq 0$).

2. $\widehat{a}^\dagger |\Psi\rangle$ is an eigenvector of $\widehat{N}$ with eigenvalue $\lambda + 1$.

Therefore, only the second statement is possible. $\square$

Note 9.14. In conclusion, we see that we can lower the eigenvalue of $\widehat{N}$ by applying $\widehat{a}$ to the eigenstate, and similarly we can raise it by applying $\widehat{a}^\dagger$. For this reason $\widehat{a}$ and $\widehat{a}^\dagger$ are called the **lowering** and the **raising** operators.

Lemma 9.15

Let $|\Psi\rangle$ be an eigenvector of $\widehat{N}$ with the eigenvalue λ. Then, the following statements hold

1. $\lambda \in \mathbb{Z}_{\geq 0}$,

2. there exists a state $|\Phi\rangle \equiv \widehat{a}^\lambda |\Psi\rangle \in \mathcal{H}$ which satisfies $\widehat{a} |\Phi\rangle = 0$ (in particular, it is an eigenstate of $\widehat{N}$ with eigenvalue 0),

3. we have $|\Psi\rangle = C(\widehat{a}^\dagger)^\lambda |\Phi\rangle$ for some $C \in \mathbb{C}$ where $C \neq 0$.

Note 9.16. This lemma allows us to undo the action of $\widehat{a}$ with $\widehat{a}^\dagger$ and vice versa. For example consider the state $\widehat{a} |\Psi\rangle$ with eigenvalue $\lambda - 1$, and then apply $\widehat{a}^\dagger$ to get back to the eigenvalue λ. We find the state

$$\widehat{a}^\dagger \left(\widehat{a} |\Psi\rangle \right) = \widehat{a}^\dagger \widehat{a} |\Psi\rangle$$
$$= \widehat{N} |\Psi\rangle$$
$$= \lambda |\Psi\rangle .$$

Thus, unless $\lambda = 0$, we can undo the action of $\widehat{a}$ by acting with $\widehat{a}^\dagger$.

Example 9.17

Suppose that $|\Psi\rangle$ is the eigenstate of $\widehat{N}$ with the eigenvalue $\lambda = 2$. Then by Lemma 9.3 we can define the eigenstates of $\widehat{N}$

$$|\Phi_0\rangle = \widehat{a}^2 |\Psi\rangle \quad \text{and} \quad |\Phi_1\rangle = \widehat{a} |\Psi\rangle,$$

which have eigenvalues 0 and 1 respectively. We have

$$\begin{aligned}
\widehat{a}^\dagger |\Phi_1\rangle &= \widehat{a}^\dagger \widehat{a} |\Psi\rangle \\
&= \widehat{N} |\Psi\rangle \\
&= 2 |\Psi\rangle.
\end{aligned}$$

Moreover,

$$\begin{aligned}
\widehat{a}^\dagger |\Phi_0\rangle &= \widehat{a}^\dagger \widehat{a} |\Psi\rangle \\
&= \widehat{a}^\dagger \widehat{a} \widehat{a} |\Psi\rangle \\
&= \widehat{a}^\dagger \widehat{a} |\Phi_1\rangle \\
&= \widehat{N} |\Phi_1\rangle \\
&= |\Phi_1\rangle.
\end{aligned}$$

In the last two equations we used that Ψ and Φ_1 are eigenstates of $\widehat{N}$ with the eigenvalues 2 and 1 respectively. We get from these equations

$$\begin{aligned}
|\Psi\rangle &= \frac{1}{2} \widehat{a}^\dagger |\Phi_1\rangle \\
&= \frac{1}{2} \widehat{a}^\dagger (\widehat{a}^\dagger |\Phi_0\rangle) \\
&= \frac{1}{2} (\widehat{a}^\dagger)^2 |\Phi_0\rangle.
\end{aligned}$$

Lemma 9.18

Let $|\Phi\rangle$ be such that

$$\widehat{a} |\Phi\rangle = 0 \quad \text{and} \quad \langle \Phi | \Phi \rangle = 1.$$

Then, (up to a possible overall phase) $|\Phi\rangle$ is described by the wave function

$$\Phi(x) = \left(\frac{m\omega}{\pi \hbar} \right)^{\frac{1}{4}} e^{-\frac{m\omega x^2}{2\hbar}}.$$

Note 9.19. This lemma tells us that there exists a unique normalised state $|\Phi\rangle$ which satisfies $\widehat{a} |\Phi\rangle = 0$.

Proof. Recall the definition of $\widehat{a} = \frac{\widehat{p} - im\omega\widehat{x}}{\sqrt{2m\omega\hbar}}$ then we can write

$$\widehat{a} \, |\Phi\rangle = 0 \Rightarrow (\widehat{p} - im\omega\widehat{x}) \, |\Phi\rangle = 0$$

$$\Rightarrow -i\hbar\frac{\partial}{\partial x}\Phi(x) - im\omega x\Phi(x) = 0$$

$$\Rightarrow \frac{\partial}{\partial x}\Phi(x) = -\frac{m\omega}{\hbar} \cdot x\Phi(x)$$

$$\Rightarrow \Phi(x) = Ae^{-\frac{m\omega x^2}{2\hbar}},$$

with $A \in \mathbb{C}$. We can determine A (up to a phase) from the normalisation condition:

$$\langle\Phi|\Phi\rangle = \int_{-\infty}^{\infty} |\Phi(x)|^2 \, dx$$

$$= |A|^2 \int_{-\infty}^{\infty} e^{-\frac{m\omega x^2}{\hbar}}.$$

Using the change of variable $x = y\sqrt{\frac{\hbar}{m\omega}}$ and the fact $\int_{-\infty}^{\infty} e^{-y^2} \, dy = \sqrt{\pi}$, we have

$$\langle\Phi|\Phi\rangle = \sqrt{\frac{\hbar}{m\omega}} \int_{-\infty}^{\infty} e^{-y^2} \, dy$$

$$= |A|^2\sqrt{\frac{\hbar}{m\omega}}.$$

Since, we must have $\langle\Phi|\Phi\rangle = 1$ it follows that

$$A = \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}}.$$

$\square$

Definition 9.20 (Notation)**.** We identify the state $|\Phi\rangle \equiv |0\rangle$ where $|\Phi\rangle$ is defined as above. The state $|0\rangle$ is called the **ground state** (or vacuum state).

Remark 9.21. $|0\rangle \neq 0$ and indeed $|0\rangle$ is an eigenvector.

Note 9.22. We call $|0\rangle$ the *ground state* because it has the lowest possible energy.

Theorem 9.23

From $|0\rangle$ we can construct the state for $n \in \mathbb{Z}_{\geq 0}$,

$$|n\rangle \equiv \frac{1}{\sqrt{n!}}(\widehat{a}^\dagger)^n \, |0\rangle,$$

and $\langle n|n\rangle = 1$.

Proof. Recall that

$$|n\rangle \equiv A_n (\widehat{a}^\dagger)^n |0\rangle \, ;$$

in this proof we illustrate how to determine the constant A_n from the condition $\langle n|n\rangle = 1$. We will do this in two ways to demonstrate two different approaches. First, we choose A_n to be real and positive. Second, we have the following useful results

$$\langle 0| \widehat{a}^\dagger = (\widehat{a} |0\rangle)^\dagger = 0 \qquad \langle 0| \widehat{a}^\dagger |0\rangle = 0$$

$$\langle 0| \widehat{a} = \left(\widehat{a}^\dagger |0\rangle\right)^\dagger \qquad \langle 0| \widehat{a} |0\rangle = 0.$$

Next, we can write

$$\langle n|n\rangle = A_n \langle n| \left(\widehat{a}^\dagger\right)^n |0\rangle$$
$$= A_n^2 \langle 0| \left(\widehat{a}\right)^n \left(\widehat{a}^\dagger\right)^n |0\rangle \, .$$

> **Note 9.24.** The trick is to move $\widehat{a}$ past $\widehat{a}^\dagger$ so, that it becomes $\widehat{N}$, and we know how this acts on $|0\rangle$.

We compute $\langle n|n\rangle$ for certain values of n.

- Suppose $n = 1$ then,

$$\langle 1|1\rangle = A_1^2 \langle 0| \widehat{a}\widehat{a}^\dagger |0\rangle$$
$$= A_1^2 \langle 0| \widehat{a}^\dagger\widehat{a} + [\widehat{a}, \widehat{a}^\dagger] |0\rangle$$
$$= A_1^2 \langle 0| \left(\widehat{a}^\dagger\widehat{a} + 1\right) |0\rangle$$
$$= A_1^2 \langle 0| \widehat{a}^\dagger\widehat{a} |0\rangle + A_1^2 \langle 0| 1 |0\rangle$$
$$= A_1^2.$$

  Therefore, $A_1 = 1$.

- Suppose $n = 2$ then,

$$\langle 2|2\rangle = A_2^2 \langle 0| \widehat{a}\widehat{a}\widehat{a}^\dagger\widehat{a}^\dagger |0\rangle$$
$$= A_2^2 \langle 0| \widehat{a} \left(\widehat{a}^\dagger\widehat{a} + [\widehat{a}, \widehat{a}^\dagger]\right) \widehat{a}^\dagger |0\rangle$$
$$= A_2^2 \langle 0| \widehat{a} \left(\widehat{a}^\dagger\widehat{a} + 1\right) \widehat{a}^\dagger |0\rangle$$
$$= A_2^2 \langle 0| \widehat{a}\widehat{a}^\dagger\widehat{a}\widehat{a}^\dagger |0\rangle + A_2^2 \underbrace{\langle 0| \widehat{a}\widehat{a}^\dagger |0\rangle}_{\langle 1|1\rangle = 1}$$
$$= A_2^2 \langle 0| \widehat{a}\widehat{a}^\dagger \left(\widehat{a}^\dagger\widehat{a} + 1\right) |0\rangle + A_2^2$$
$$= \underbrace{\langle 0| \widehat{a}\,\widehat{a}^\dagger\widehat{a}^\dagger\,\widehat{a} |0\rangle}_{0} + A_2^2$$
$$= 2A_2^2.$$

  Therefore, $A_2 = \frac{1}{2}$.

For the general case n, we first consider

$$\widehat{a}(\widehat{a}^\dagger)^n = \widehat{a}\widehat{a}^\dagger(\widehat{a}^\dagger)^{n-1}$$
$$= \left(\widehat{a}^\dagger\widehat{a} + [\widehat{a}, \widehat{a}^\dagger]\right) (\widehat{a}^\dagger)^{n-1}$$
$$= \widehat{a}^\dagger\widehat{a}(\widehat{a}^\dagger)^{n-1} + (\widehat{a}^\dagger)^{n-1}.$$

Consider the term

$$\widehat{a}^\dagger\widehat{a}(\widehat{a}^\dagger)^{n-1} = \widehat{a}^\dagger\widehat{a}\widehat{a}^\dagger(\widehat{a}^\dagger)^{n-2}$$
$$= \widehat{a}^\dagger(\widehat{a}^\dagger\widehat{a}+1)(\widehat{a}^\dagger)^{n-2}$$
$$= (\widehat{a}^\dagger)^2\,\widehat{a}\,(\widehat{a}^\dagger)^{n-2} + (\widehat{a}^\dagger)^{n-1}.$$

Therefore,

$$\widehat{a}(\widehat{a}^\dagger)^n = (\widehat{a}^\dagger)^2\,\widehat{a}\,(\widehat{a}^\dagger)^{n-2} + 2(\widehat{a}^\dagger)^{n-1}.$$

Iterating this process k times we have

$$\widehat{a}(\widehat{a}^\dagger)^n = (\widehat{a}^\dagger)^k\,\widehat{a}\,(\widehat{a}^\dagger)^{n-1} + k(\widehat{a}^\dagger)^{n-1}.$$

Then iterating the process n times i.e. when $k = n$ we have

$$\widehat{a}(\widehat{a}^\dagger)^n = (\widehat{a}^\dagger)^n\widehat{a} + n(\widehat{a}^\dagger)^{n-1}.$$

With this identity we can compute

$$\langle n|n\rangle = A_n^2\,\langle 0|\,\widehat{a}^n(\widehat{a}^\dagger)^n\,|0\rangle$$
$$= A_n^2\,\langle 0|\,\widehat{a}^{n-1}\widehat{a}(\widehat{a}^\dagger)^n\,|0\rangle$$
$$= A_n^2\,\langle 0|\,\widehat{a}^{n-1}\left((\widehat{a}^\dagger)^n\widehat{a} + n(\widehat{a}^\dagger)^{n-1}\right)|0\rangle$$
$$= A_n^2\,\langle 0|\,\widehat{a}^{n-1}(\widehat{a}^\dagger)^n\,\underbrace{\widehat{a}\,|0\rangle}_{0} + A_n^2 n\,\langle 0|\,\widehat{a}^{n-1}(\widehat{a}^\dagger)^{n-1}\,|0\rangle$$
$$= A_n^2 n\,\langle 0|\,\widehat{a}^{n-1}(\widehat{a}^\dagger)^{n-1}\,|0\rangle$$
$$= A_n^2 n(n-1)\,\langle 0|\,\widehat{a}^{n-2}(\widehat{a}^\dagger)^{n-2}\,|0\rangle$$
$$= \cdots$$
$$= A_n^2 n(n-1)(n-2)\cdots 1\,\langle 0|\,\widehat{a}^0(\widehat{a}^\dagger)^0\,|0\rangle$$
$$= A_n^2 n!\,\langle 0|0\rangle$$
$$= A_n^2 n!.$$

Therefore,

$$A_n = \frac{1}{\sqrt{n!}}.$$

$\square$

Proof. We present a quicker method to compute A_n. Consider

$$|n+1\rangle = C_n\widehat{a}^\dagger\,|n\rangle$$

for some $C_n > 0$. We can determine C_n by imposing the normalisation condition.

$$\langle n+1|n+1\rangle = |C_n|^2\,\langle n|\,\widehat{a}\widehat{a}^\dagger\,|n\rangle$$
$$= C_n^2\,\langle n|\,\widehat{a}^\dagger\widehat{a} + [\widehat{a},\widehat{a}^\dagger]\,|n\rangle$$
$$= C_n^2\,\langle n|\,\widehat{N}+1\,|n\rangle$$
$$= C_n^2(n+1)\,\langle n|n\rangle$$
$$= C_n^2(n+1).$$

Therefore, $C_n = \frac{1}{\sqrt{n+1}}$ and

$$|n+1\rangle = \frac{\widehat{a}^\dagger}{\sqrt{n+1}}\,|n\rangle.$$

We can then start from $|0\rangle$,

$$|1\rangle = \widehat{a}^\dagger |0\rangle, \quad |2\rangle = \frac{\widehat{a}^\dagger}{\sqrt{2}} |1\rangle = \frac{(\widehat{a}^\dagger)^2}{\sqrt{2}} |0\rangle, \cdots$$

to conclude

$$|n\rangle = \frac{1}{\sqrt{n}} (\widehat{a}^\dagger)^n |0\rangle.$$

$\square$

Corollary 9.25. The state $|n\rangle$ is an eigenstate of $\widehat{N}$ with the eigenvalue n, i.e.

$$\widehat{N} |n\rangle = n |n\rangle.$$

Proof. Use Lemma 9.4. $\square$

Theorem 9.26

Let $|\Psi\rangle$ be an eigenstate of $\widehat{N}$ with the eigenvalue λ. Then, the following statements hold

1. $\lambda = n \in \mathbb{Z}_{\geq 0}$,

2. $|\Psi\rangle = C |n\rangle$ for some $C \in \mathbb{C}$ where $C \neq 0$.

Theorem 9.27

The eigenstates of $\widehat{H}$ are the states $|n\rangle$, and the corresponding eigenvalues are

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right).$$

The spectrum of $\widehat{H}$ is therefore,

$$\left\{ \frac{\hbar\omega}{2}, \frac{3\hbar\omega}{2}, \frac{5\hbar\omega}{2}, \cdots \right\}$$

Appendix

A Eigenvalue and eigenvectors

Definition A.1. Let V be a vector space. If $T : V \to V$ is a linear operator then, we call a vector $\boldsymbol{v} \in V$ an **eigenvector** for T if

$$\boldsymbol{v} \neq \boldsymbol{0} \quad \text{and} \quad T(\boldsymbol{v}) = \lambda \boldsymbol{v} \text{ for some } \lambda \in \mathbb{C}.$$

The scalar λ is called an **eigenvalue** of T.

Lemma A.2. Let T be a linear operator on a vector space V and $\boldsymbol{v} \in V$ and $\lambda \in \mathbb{C}$. Then $\boldsymbol{v}$ is an eigenvector of T with eigenvalue λ if and only if

$$\boldsymbol{v} \neq \boldsymbol{0} \quad \text{and} \quad \boldsymbol{v} \in \ker(T - \lambda I).$$

Theorem A.3

The scalar $\lambda \in \mathbb{C}$ is an eigenvalue of T is and only if $\det(T - \lambda I) = 0$.

Example A.4. Determine all eigenvalues and eigenvectors of the matrix

$$A = \begin{pmatrix} 1 & -10 \\ -5 & -4 \end{pmatrix}.$$

Solution. To find the eigenvalues we begin by solving the characteristic polynomial given by $p_A(\lambda) = \det(A - \lambda I) = 0$. We have that

$$A - \lambda I = \begin{pmatrix} 1 - \lambda & -10 \\ -5 & -4 - \lambda \end{pmatrix},$$

and so

$$\det(A - I\lambda) = (\lambda - 6)(\lambda + 9).$$

Thus, the eigenvalues of A are $\lambda_1 = 6$ and $\lambda_2 = -9$.
To find the eigenvectors of A, we have to solve the system $(A - \lambda I)\boldsymbol{v} = \boldsymbol{0}$ for $\lambda = 6, -9$.
Solving the system:

$$(A - 6I)\boldsymbol{v} = \begin{pmatrix} -5 & -10 \\ -5 & -10 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and so we have the system

$$\begin{cases} -5x - 10y &= 0 \\ -5x - 10y &= 0. \end{cases}$$

We conclude the eigenvector is of the form $x = -2y$ thus, by choosing $x = 1$ the eigenvector

$$\boldsymbol{v}_1 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

By a similar process the eigenvector

$$\boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Example A.5

Here we present an alternative method to finding the eigenvectors of the matrix A. From out previous work we know A has eigenvalues $\lambda = 6, -9$. We can choose an ansatz for the eigenvectors to be of the form

$$\boldsymbol{v} = \begin{pmatrix} 1 \\ \alpha \end{pmatrix},$$

and we know this eigenvector must satisfy the following equation

$$\begin{pmatrix} 1 & -10 \\ -5 & -4 \end{pmatrix} \begin{pmatrix} 1 \\ \alpha \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ \alpha \end{pmatrix}.$$

Therefore, we have the following set of equations:

$$\begin{cases} 1 - 10\alpha &= \lambda \\ -5 - 4\alpha &= \lambda\alpha; \end{cases}$$

from the first equation it follows that

$$\alpha = \frac{1}{10}(1 - \lambda).$$

We conclude the eigenvectors of A are

$$\boldsymbol{v}_{\lambda=6} = \begin{pmatrix} 1 \\ -0.5 \end{pmatrix} \quad \text{and} \quad \boldsymbol{v}_{\lambda=-9} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

A.1   Diagonalisation

Diagonal matrices have nice properties which makes working with them easy.

Proposition A.6. Properties of diagonal matrices:

- The determinant of $\text{diag}(a_1, \ldots, a_n)$ is the product $a_1 \cdots a_n$.

- Matrix multiplication: $\text{diag}(a_1, \ldots, a_n)\text{diag}(b_1, \ldots, b_n) = \text{diag}(a_1 b_1, \ldots, a_n b_n)$.

- The diagonal matrix $\text{diag}(a_1, \ldots, a_n)$ is invertible if and only if the entries $a_1, \ldots, a_n$ are all non-zero. In this case, we have

$$\text{diag}(a_1, \ldots, a_n) = \text{diag}(a_1^{-1}, \ldots, a_n^{-1}).$$

> **Theorem A.7**
>
> To diagonalise a matrix A we need a matrix P such that
>
> $$A = PDP^{-1}$$
>
> where D is a diagonal matrix. It follows that
>
> $$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$
>
> where λ_i are the eigenvalues of A and,
>
> $$P = (\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n)$$
>
> where $\boldsymbol{v}_i$ are the corresponding eigenvectors.

B   Multiple eigenvalue multiplicity

example of this in a second. For now, let us consider what happens when the eigenvalues have multiplicities, i.e. for each λ_i we have several eigenstates $|\lambda_i, 1\rangle, |\lambda_i, 2\rangle, \cdots, |\lambda_i, n_i\rangle$, which we assume to be orthonormal,

$$\langle \lambda_i, a | \lambda_j, b \rangle = \delta_{a,b} \delta_{i,j}. \tag{393}$$

In this case the state can be expanded as

$$|\Psi\rangle = \sum_{i=1}^{\infty} \sum_{a=1}^{n_i} c_{i,a} |\lambda_i, a\rangle. \tag{394}$$

In this case, the probability of obtaining the result λ_i is $\sum_{a=1}^{n_i} |c_{i,a}|^2$. However, it is harder to determine the final state – it is not uniquely determined by the eigenvalue λ_i and in general can be any linear combination $\sum_{a=1}^{n_i} b_a |\lambda_i, a\rangle$. The exact form of the final state depends on the precise details of the measurement of $\widehat{\mathcal{O}}$ and will not be discussed here.

C   Links

- Dirac delta function