Principles of Quantum Mechanics

Willoughby Seago

September 22, 2020

These are my notes for the *principles of quantum mechanics* course from the University of Edinburgh as part of the third year of the theoretical physics degree. When I took this course in the 2020/21 academic year it was taught by Professor Luigi Del Debbio[1] and Professor Arjun Berera[2]. These notes are based on the lectures delivered as part of this course, and the notes provided as part of this course, which are in the process of being published as a textbook. The content within is correct to the best of my knowledge but if you find a mistake or just disagree with something or think it could be improved please let me know.

These notes were produced using LaTeX[3]. Graphs where plotted using Python[4], Matplotlib[5], NumPy[6], and SciPy[7]. Diagrams were drawn with tikz[8].

This is version 1.1 of these notes, which is up to date as of 07/03/2021.

<div align="right">

Willoughby Seago

s1824487@ed.ac.uk

</div>

---

[1]https://www.ph.ed.ac.uk/people/luigi-del-debbio
[2]https://www.ph.ed.ac.uk/people/arjun-berera
[3]https://www.latex-project.org/
[4]https://www.python.org/
[5]https://matplotlib.org/
[6]https://numpy.org/
[7]https://scipy.org/scipylib/
[8]https://www.ctan.org/pkg/pgf

Contents

IV   Angular Momentum and Spin   60

List of Figures

List of Tables

Acronyms

CSCO   complete set of commuting/compatible observables.
EPR   Einstein–Podolsky–Rosen.
HUP   Heisenberg's uncertainty principle.
PDF   probability density function.
QED   quantum electrodynamics.
QM   quantum mechanics.
STM   scanning tunnelling microscope.
TDSE   time dependent Schrödinger equation.
TISE   time independent Schrödinger equation.

Part I

States

1   Quantum States

1.1   Classical States

In classical mechanics a point like system's state is entirely described by two 3-dimensional vectors, namely position, $\boldsymbol{x}$, and momentum, $\boldsymbol{p}$. These are 3-dimensional real vectors living in $\mathbb{R}^3$. We can instead view the object $(\boldsymbol{x}, \boldsymbol{p})$ as a 6-dimensional real vector living in $\mathbb{R}^6$. In this way we can completely describe the state of the system with one 6-dimensional vector. By this we mean that using $(\boldsymbol{x}, \boldsymbol{p})$ and Newton's laws we can compute the time evolution of the system, that is what the state will be at some time in the future, or what the state was at some time in the past. We can also compute other values such as energy and angular momentum, these are observables that we can measure along with position and momentum.

1.2   Quantum States

In quantum mechanics (QM) the state of a system can also be encoded in a single vector, although the way that we do this is more complex (literally). We start with the following postulate:

> **Postulate 1**
>
> The state of a quantum system is described by a complex vector.

In classical mechanics the state vector is real and therefore composed of values that can be measured. In QM we cannot directly measure the components of this complex vector. The way we measure observables will be covered in detail later. For now we focus on the mathematical side of this statement.

In this course we will use Dirac notation in which a vector is represented by $|v\rangle$, called a **ket**, where v is some label. For example we may use $|v\rangle$, $|\uparrow\rangle$, or $|n, m\rangle$. Here v, $\uparrow$, and n, m are simply whatever label is best suited to the problem at hand.

1.2.1   Vector Spaces

Our intuition for what is a vector is normally that it is some object with magnitude and direction. This definition fails when it comes to complex vectors and is not very useful for vectors of more than 3-dimensions. To be mathematically precise a **vector** is any element of a vector space, $\mathcal{H}$. This simply moves the problem on to the next definition, what is a vector space?

A **vector space**, $(\mathcal{H}, \mathbb{F}, +, \cdot)$, over a field, $\mathbb{F}$, is a set, $\mathcal{H}$ and two operations, $+$ and $\cdot$. We will restrict ourselves to $\mathbb{F} = \mathbb{R}, \mathbb{C}$. In the former case we get a real vector space and in the later a complex vector space. These operations are

- Vector addition, $+\colon \mathcal{H} \times \mathcal{H} \to \mathcal{H}$, which takes two vectors, $|u\rangle, |v\rangle \in \mathcal{H}$ and maps them to a third vector $|u\rangle + |v\rangle \in \mathcal{H}$.

- Scalar multiplication, $\cdot\colon \mathbb{F} \times \mathcal{H} \to \mathcal{H}$, which takes a vector, $|v\rangle \in \mathcal{H}$, and a scalar, $\lambda \in \mathbb{F}$, and maps them to a second vector $\lambda|v\rangle$.

These operations must satisfy the following properties:

- Associativity of vector addition – For all $|u\rangle, |v\rangle, |w\rangle \in \mathcal{H}$

$$|u\rangle + (|v\rangle + |w\rangle) = (|u\rangle + |v\rangle) + |w\rangle.$$

In other words the bracketing around vector addition doesn't matter, the result is the same however it is bracketed. This is true for more than three terms being added as well.

- Commutativity of vector addition – For all $|u\rangle, |v\rangle \in \mathcal{H}$

$$|u\rangle + |v\rangle = |v\rangle + |u\rangle.$$

In other words the order of terms in a sum doesn't matter, the result is the same either way.

- Existence of an additive identity – There exists $|0\rangle \in \mathcal{H}$ such that

$$|0\rangle + |v\rangle = |v\rangle$$

for all $|v\rangle \in \mathcal{H}$. $|0\rangle$ is called the zero vector. Be careful as $|0\rangle$ may not always denote the zero vector, for example, if we use $|p\rangle$ to denote a state with momentum p then $|0\rangle$ denotes a state with no momentum. This doesn't mean that this is the zero vector.

- Existence of additive inverses – For all $|v\rangle$ there exists $-|v\rangle \in \mathcal{H}$ such that

$$|v\rangle + (-|v\rangle) = |0\rangle$$

where $|0\rangle$ is the zero vector. We normally shorten this to

$$|v\rangle - |v\rangle = |0\rangle.$$

- Compatibility of scalar and field multiplication – for all $\lambda, \mu \in \mathbb{F}$ and $|v\rangle \in \mathcal{H}$ we have

$$(\lambda\mu)|v\rangle = \lambda(\mu|v\rangle)$$

here the product $\lambda\mu$ is computed using multiplication as defined in $\mathbb{F}$ and $\mu|v\rangle$ and $\lambda(\mu|v\rangle)$ are computed using scalar multiplication as defined for our vector space.

- Identity of scalar multiplication – If $1 \in \mathbb{F}$ is the multiplicative identity of $\mathbb{F}$ then

$$1|v\rangle = |v\rangle$$

for all $|v\rangle \in \mathcal{H}$.

- Distributivity of scalar multiplication over vector addition – For all $|u\rangle, |v\rangle \in \mathcal{H}$ and $\lambda \in \mathbb{F}$

$$\lambda(|u\rangle + |v\rangle) = \lambda|u\rangle + \lambda|v\rangle.$$

- Distributivity of scalar multiplication over field addition – For all $\lambda, \mu \in \mathbb{F}$ and $|v\rangle \in \mathcal{H}$

$$(\lambda + \mu)|v\rangle = \lambda|v\rangle + \mu|v\rangle.$$

The important take away here is that for a complex vector space, $\mathcal{H}$ if $|u\rangle, |v\rangle \in \mathcal{H}$ and $\alpha, \beta \in \mathbb{C}$ then

$$\alpha|u\rangle + \beta|v\rangle \in \mathcal{H}.$$

That is $\mathcal{H}$ is closed under vector addition and scalar multiplication. This easily generalises to $|v_i\rangle \in \mathcal{H}$ and $\alpha_i \in \mathcal{H}$ so

$$\sum_i \alpha_i|v_i\rangle \in \mathcal{H}$$

Any linear combination of vectors is again a vector.

Example 1.1. Let $|u\rangle$ and $|v\rangle$ be vectors with two complex components:

$$|u\rangle = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \qquad \text{and} \qquad |v\rangle = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

where $u_1, u_2, v_1, v_2 \in \mathbb{C}$. If we define the sum of these two vectors to be

$$|u\rangle + |v\rangle = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix}.$$

The addition on the left hand side is vector addition, the addition on the right hand side is field addition in $\mathbb{C}$. We define scalar multiplication by $\lambda \in \mathbb{C}$ as

$$\lambda|v\rangle = \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \end{pmatrix}.$$

Multiplication on the left hand side is scalar multiplication, multiplication on the right hand side is field multiplication in $\mathbb{C}$. The set, $\mathbb{C}^2$, of all such objects with this definition of vector addition and scalar multiplications is a complex vector space. We will use this throughout this section to give examples.

The fact that a linear combination of vectors is again a vector means that a linear combination of quantum states is again a quantum state. This is called the **superposition principle**. It is a direct consequence of postulate 1 and the closure of vector spaces under addition and multiplication. In fact it is possible to state the superposition principle as the first postulate and then derive that the states need to be elements of a vector space.

1.2.2 Scalar Products

The scalar product, or inner product, is a function:

$$\cdot \colon \mathcal{H} \times \mathcal{H} \to \mathbb{C}$$
$$|u\rangle, |v\rangle \mapsto \langle u|v\rangle = z \in \mathbb{C}.$$

The inner product also satisfies the following two properties:

- Conjugate symmetry – For all $|u\rangle, |v\rangle \in \mathcal{H}$

$$\langle u|v\rangle = \langle v|u\rangle^*$$

  where * denotes complex conjugation.

- Linearity in the second argument – For all $|u\rangle, |v\rangle, |w\rangle \in \mathcal{H}$ and $\alpha, \beta \in \mathbb{C}$

$$\langle u| \left[\alpha|v\rangle + \beta|w\rangle\right] = \alpha\langle u|v\rangle + \beta\langle u|w\rangle.$$

Combining these properties we get that for all $|u\rangle, |v\rangle, |w\rangle \in \mathcal{H}$ and $\alpha, \beta \in \mathbb{C}$

$$\left[\alpha\langle u| + \beta\langle v|\right]|w\rangle = \alpha^*\langle u|w\rangle + \beta^*\langle v|w\rangle.$$

A vector space equipped with an inner product is called a **Hilbert space**.[9]

Example 1.2. Given two vectors $|u\rangle, |v\rangle \in \mathbb{C}^2$ defined as

$$|u\rangle = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \qquad \text{and} \qquad |v\rangle = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

the following is a valid inner product:

$$\langle u|v\rangle = u_1^* v_1 + u_2^* v_2.$$

1.2.3 Linear Functionals

A **linear functional** is a function, f, defined on $\mathcal{H}$ that associates each vector with a complex number,

$$f \colon \mathcal{H} \to \mathbb{C}$$
$$|v\rangle \mapsto f(|v\rangle) = z \in \mathbb{C},$$

---

[9]Technically it is just an inner product space, a Hilbert space is then an inner product space that is Cauchy complete. That is all series of vectors that converge absolutely (the sum of the norms of the vectors converges) in $\mathcal{H}$ converge (the sum of the vectors converges) in $\mathcal{H}$. However all inner product spaces can be extended to be Hilbert spaces so the distinction isn't important for our purposes.

in such a way that

$$f(\alpha|u\rangle + \beta|v\rangle) = \alpha f(|u\rangle) + \beta f(|v\rangle)$$

for all $|u\rangle, |v\rangle \in \mathcal{H}$, and $\alpha, \beta \in \mathbb{C}$.

For a given vector, $|u\rangle \in \mathcal{H}$, we can define a linear functional, $f_{|u\rangle}$ that takes a vector $|v\rangle \in \mathcal{H}$ and returns a complex number through the correspondence

$$|v\rangle \mapsto z = f_{|u\rangle}(|v\rangle) = \langle u|v\rangle.$$

It can be shown that all bounded linear functionals that act on $\mathcal{H}$ can actually be represented like this. That is that for any linear functional defined there exists $|u\rangle \in \mathcal{H}$ such that the linear functional is equivalent to an inner product with this vector. This is called the **Riesz theorem**. This means that given any linear functional we can identify it with a vector and vice versa. This one-to-one correspondence led Dirac to introduce the **bra** notation where a linear functional is denoted $\langle u|$. Again u is just a label and this represents the linear functional $f_{|u\rangle}$ that has the result of $f_{|u\rangle}(|v\rangle) = \langle u|v\rangle$. The space of all linear functionals is called the **dual** of the vector space $\mathcal{H}$, and is denoted $\mathcal{H}^*$.

1.2.4   Norms

The **norm** of a vector is a function:

$$\|\cdot\| \colon \mathcal{H} \to \mathbb{R}$$
$$|v\rangle \mapsto \|v\|$$

where $\|v\|$ has some restraints:

- The triangle inequality – For all $|u\rangle, |v\rangle \in \mathcal{H}$

$$\||u\rangle + |v\rangle\| \leq \|u\| + \|v\|.$$

- Absolutely scaleable – For all $|v\rangle \in \mathcal{H}$, $\lambda \in \mathbb{C}$:

$$\|a|v\rangle\| = |\lambda|\|v\|.$$

- Positive definite – If $\|v\| = 0$ then $|v\rangle = |0\rangle$ is the zero vector.

Note that we use $\|\cdot\|$ for the norm and $|\cdot|$ for the absolute value of a complex number defined as $|z| = \sqrt{z^*z}$. This actually hints at the most common norm that we will use throughout this course. The norm induced by the inner product is

$$\|v\| = \sqrt{\langle v|v\rangle}.$$

Using the fact that the inner product is linear we have $\langle v|v\rangle = \langle v|v\rangle^*$ so we know that $\langle v|v\rangle \in \mathbb{R}$.

In QM it is conventional to have all state vectors have unital norm. The reason for this will become clear when we start to discuss how the physics is extracted from the state vector. Fortunately all states can be normalised such that their norm is 1, for all $|v\rangle \in \mathcal{H}$ if $\|v\| \neq 0$ then

$$\frac{1}{\sqrt{\langle v|v\rangle}}|v\rangle = \frac{1}{\|v\|}|v\rangle$$

is a vector with norm 1. This follows from the absolute scaleability and positive definiteness of the norm:

$$\left\|\frac{1}{\|v\|}|v\rangle\right\| = \left|\frac{1}{\|v\|}\right|\|v\| = \frac{1}{\|v\|}\|v\| = 1.$$

Note that if $\|v\| = 1$ then

$$\left\|e^{i\vartheta}|v\rangle\right\| = \left|e^{i\vartheta}\right|\|v\| = 1 \cdot 1 = 1.$$

Thus quantum states are actually only unique up to a factor of $e^{i\vartheta}$, called the phase.

Example 1.3. Let $|v\rangle \in \mathbb{C}^2$ be defined by

$$|v\rangle = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

then the norm of $|v\rangle$ is

$$\|v\| = \sqrt{\langle v|v\rangle} = \sqrt{v_1^* v_1 + v_2^* v_2} = \sqrt{|v_1|^2 + |v_2|^2}.$$

We can see how similar this is to the Euclidean norm that we are all used to. Also it is clearly only 0 if $v_1, v_2 = 0$, i.e. $|v\rangle = 0$.

1.2.5   Orthogonality

Two vectors, $|u\rangle, |v\rangle \in \mathcal{H}$ are **orthogonal** if and only if

$$\langle u|v\rangle = 0.$$

1.2.6   Operators

Operators are functions:

$$\hat{O} \colon \mathcal{H} \to \mathcal{H}$$
$$|v\rangle \mapsto \hat{O}|v\rangle = |w\rangle \in \mathcal{H}.$$

An operator, $\hat{O}$, is a linear operator if for all $|u\rangle, |v\rangle \in \mathcal{H}$ and $\alpha, \beta \in \mathcal{H}$

$$\hat{O}\left[\alpha|u\rangle + \beta|v\rangle\right] = \alpha\hat{O}|u\rangle + \beta\hat{O}|v\rangle.$$

If $\hat{O}_1$ and $\hat{O}_2$ are operators on $\mathcal{H}$ then their product, $\hat{O}_1\hat{O}_2$, is also an operator on $\mathcal{H}$ defined as

$$\hat{O}_1\hat{O}_2|v\rangle = \hat{O}_1\left[\hat{O}_2|v\rangle\right]$$

for $|v\rangle \in \mathcal{H}$. That is given a product of operators we just apply them one at a time from the right most operator to the left most.

This distinction in the order the operators apply is important because in general operators do not commute. That is

$$\hat{O}_1\hat{O}_2 \neq \hat{O}_2\hat{O}_1.$$

The commutator, defined as

$$[\hat{O}_1, \hat{O}_2] = \hat{O}_1\hat{O}_2 - \hat{O}_2\hat{O}_1,$$

gives a measure of to what degree $\hat{O}_1$ and $\hat{O}_2$ fail to commute. The commutator is itself an operator and turns out to be very important in QM.

We can extend the definition of a product of operators to any number of operators. In particular we look at the product of an operator with itself $n \in \mathbb{N}$ times. If $n > 0$ then

$$\hat{O}^n = \underbrace{\hat{O}(\hat{O}(\cdots(\hat{O}|v\rangle)))}_{n \text{ factors}}$$

for $|v\rangle \in \mathcal{H}$. Also we define

$$\hat{O}^0 = \mathbb{1}$$

where $\mathbb{1}$ is the identity operator defined by

$$\mathbb{1}|v\rangle = |v\rangle$$

for all $|v\rangle \in \mathcal{H}$. With these definitions we can now define functions of operators. In general if f is a function such that

$$f(x) = \sum_n c_n x^n$$

for $c_n \in \mathbb{C}$ then we can extend this function to act on operators by

$$f(\hat{O}) = \sum_n c_n \hat{O}^n,$$

where the coefficients c_n here are the same as they were for $f(x)$.

Example 1.4. The exponential function has a Taylor series given by

$$e^x = \sum_n \frac{1}{n!} x^n = 1 + x + \frac{x^2}{2!} + \cdots .$$

We can define the exponential of an operator as

$$\exp\left(\hat{O}\right) = \sum_n \frac{1}{n!} \hat{O}^n = \mathbb{1} + \hat{O} + \frac{\hat{O}}{2!} + \cdots .$$

Note that by linearity each term of the sum is an operator and therefore the whole sum is an operator. In general a function of an operator defined in this way will return another operator.

1.2.7   Bases

A **basis** of a vector space, $\mathcal{H}$, is a a set of linearly independent vectors, $\mathcal{B} = \{|e_k\rangle \mid k = 1, \ldots, n\} \subset \mathcal{H}$ such that all $|v\rangle \in \mathcal{H}$ can be expressed as a linear combination of basis vectors:

$$|v\rangle = \sum_k v_k |e_k\rangle.$$

Here $v_k \in \mathbb{C}$ are the components of $|v\rangle$ in the basis $\mathcal{B}$. The dimension of the vector space is given by the number of basis vectors, $|\mathcal{B}| = \dim(\mathcal{H})$. For example in $\mathbb{R}^3$

$$\boldsymbol{v} = v_1 \boldsymbol{e_1} + v_2 \boldsymbol{e_2} + v_3 \boldsymbol{e_3} = \sum_{i=1}^{3} v_i \boldsymbol{e_i}.$$

Here $v_i \in \mathbb{R}$ are the components of $\boldsymbol{v}$.

A basis $\mathcal{B} = \{|e_k\rangle\}$ is **orthonormal** if

$$\langle e_k | e_l \rangle = \delta_{kl}$$

for all k and l where δ_{kl} is the Kronecker delta defined as

$$\delta_{kl} = \begin{cases} 1, & k = l, \\ 0, & k \neq l. \end{cases}$$

We almost always work in an orthonormal basis simply because it makes computation so much easier.

We can use this property of an orthonormal basis to find the components of a vector $|v\rangle \in \mathcal{H}$ in the basis $\mathcal{B} = \{|e_k\rangle\}$:

$$\begin{aligned} \langle e_k | v \rangle &= \langle e_k | \left[\sum_l v_l |e_l\rangle \right] \\ &= \sum_l v_l \langle e_k | e_l \rangle \\ &= \sum_l v_l \delta_{kl} \\ &= v_k. \end{aligned}$$

We can also compute the norm of $|v\rangle \in \mathcal{H}$ using this. First note that $|v\rangle$ can be written as

$$|v\rangle = \sum_k v_k |e_k\rangle = \sum \langle e_k | v \rangle |e_k\rangle = \sum_k |e_k\rangle \langle e_k | v \rangle \tag{1.1}$$

using this we get

$$\langle v| = \sum_k v_k^* \langle e_k| = \sum \langle e_k | v \rangle^* \langle e_k| = \sum_k \langle v | e_k \rangle \langle e_k|$$

and then a direct computation gives us

$$\langle v | v \rangle = \left[\sum_k \langle v | e_k \rangle \langle e_k| \right] \left[\sum_l |e_l\rangle \langle e_l | v \rangle \right]$$

$$= \sum_{k,l} \langle v|e_k\rangle\langle e_k|e_l\rangle\langle e_l|v\rangle$$

$$= \sum_{k,l} \langle v|e_k\rangle\delta_{kl}\langle e_l|v\rangle$$

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

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

$$= \sum_{k} v_k^* v_k$$

$$= \sum_{k} |v_k|^2.$$

Notice that in equation 1.1 we had

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

We can think of the object $|e_k\rangle\langle e_k|$ as a **projection operator**, $\hat{\mathcal{P}}_k = |e_k\rangle\langle e_k|$, which projects $|v\rangle$ onto the basis vector $|e_k\rangle$. Projection operators have the following property:

$$\hat{\mathcal{P}}_k\hat{\mathcal{P}}_l = \delta_{kl}\mathcal{P}_k.$$

That is the same projection operator applied twice is the same as it applied once, it is idempotent, $\hat{\mathcal{P}}_k^2 = \hat{\mathcal{P}}_k$, and projection onto two different basis vectors is zero, $\hat{\mathcal{P}}_k\hat{\mathcal{P}}_l = 0$ if $k \neq l$. Since equation 1.1 is true for all $|v\rangle \in \mathcal{H}$ we have

$$\sum_{k} \hat{\mathcal{P}}_k = \sum_{k} |e_k\rangle\langle e_k| = \mathbb{1}$$

where 1 is the identity operator acting on $\mathcal{H}$, that is for all $|v\rangle \in \mathcal{H}$ we have $1|v\rangle = |v\rangle$. This is called the **completeness relation** and the vectors of the basis are said to be a **complete set** in $\mathcal{H}$. In a vector space of finite dimension any set of n linearly independent vectors is complete.

Note that in two different bases, $\mathcal{B} = \{|e_k\rangle\}$, and $\mathcal{B}' = \{|e_k'\rangle\}$, in general $v_k = \langle e_k|v\rangle \neq v_k' = \langle e_k'|v\rangle$ however we still have

$$|v\rangle = \sum_{k} v_k|e_k\rangle = \sum_{k} v_k'|e_k'\rangle.$$

So the vector is independent of the basis but its components aren't.

2   Two-State Systems

2.1   Calculations Within a Basis

2.1.1   Vectors in Two Bases

Let $\mathcal{H}$ be a vector space, and let $\mathcal{B}_1 = \{|e_k\rangle\}$ and $\mathcal{B}_2 = \{|e_k'\rangle\}$ be orthonormal bases of $\mathcal{H}$. We know that $|v\rangle \in \mathcal{H}$ can be written as

$$|v\rangle = \sum_{k} v_k|e_k\rangle = \sum_{k} v_k'|e_k'\rangle.$$

We can work out a relationship between v_k and v_k':

$$v_k' = \langle e_k'|v\rangle$$

$$= \langle e_k'|\mathbb{1}|v\rangle$$

$$= \langle e_k'|\sum_{l} |e_l\rangle\langle e_l|v\rangle$$

$$= \sum_{l} \langle e_k'|e_l\rangle\langle e_l|v\rangle$$

$$= \sum_{l} \langle e_k'|e_l\rangle v_l$$

$$= \sum_l U_{kl} v_l$$

where $U_{kl} = \langle e'_k | e_l \rangle$. U is a unitary matrix. We can show this quite easily:

$$
\begin{aligned}
UU^\dagger &= \sum_l U_{kl}(U^\dagger)_{lm} \\
&= \sum_l \langle e'_k | e_l \rangle (\langle e'_k | e_l \rangle)^\dagger \\
&= \sum_l \langle e'_k | e_l \rangle \langle e_l | e'_k \rangle \\
&= \langle e'_k | \mathbb{1} | e'_k \rangle \\
&= \langle e'_k | e'_k \rangle \\
&= \mathbb{1}
\end{aligned}
$$

This is only true if the bases are orthonormal.

2.1.2   Scalar Products in a Basis

Let $\mathcal{H}$ be a vector space and let $\mathcal{B} = \{|e_k\rangle\}$ be an orthonormal basis of $\mathcal{H}$. Let $|v\rangle, |w\rangle \in \mathcal{H}$. We know that

$$|v\rangle = \sum_k v_k |e_k\rangle, \qquad \text{and} \qquad |w\rangle = \sum_k w_k |e_k\rangle.$$

The scalar product of these two vectors is

$$
\begin{aligned}
\langle w | v \rangle &= \left[\sum_k w_k |e_k\rangle \right]^\dagger \left[\sum_l v_l |e_l\rangle \right] \\
&= \left[\sum_k w_k^* \langle e_k | \right] \left[\sum_l v_l |e_l\rangle \right] \\
&= \sum_{k,l} w_k^* v_l \langle e_k | e_l \rangle \\
&= \sum_{k,l} w_k^* v_l \delta_{kl} \\
&= \sum_k w_k^* v_k
\end{aligned}
$$

2.2   Operators in a Basis

Let $\mathcal{H}$ be a vector space and let $\mathcal{B} = \{|e_k\rangle\}$ be an orthonormal basis. Let $\hat{O}$ be a linear operator on $\mathcal{H}$. For $|v\rangle \in \mathcal{H}$ we have

$$|v\rangle = \sum_k v_k |e_k\rangle.$$

Define

$$|v'\rangle = \hat{O}|v\rangle$$

which we can expand in this basis as

$$|v'\rangle = \sum_k v'_k |e_k\rangle.$$

We can compute v'_k:

$$
\begin{aligned}
v'_k &= \langle e_k | v' \rangle \\
&= \langle e_k | \hat{O} | v \rangle \\
&= \langle e_k | \hat{O} \sum_l v_l | e_l \rangle
\end{aligned}
$$

$$= \sum_l v_l \langle e_k | \hat{O} | e_l \rangle$$

$$= \sum_l O_{kl} v_l$$

where $O_{kl} = \langle e_k | \hat{O} | e_l \rangle$ are the **matrix elements** of $\hat{O}$ between basis vectors.

2.3 Tensor Product

2.3.1 What is a Tensor Product?

The **tensor product**, $\otimes$, is an operation:

$$\otimes \colon \mathcal{H}_1 \times \mathcal{H}_2 \to \mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$$
$$|v^{(1)}\rangle, |v^{(2)}\rangle \mapsto |v^{(1)}\rangle \otimes |v^{(2)}\rangle = |v^{(1)}\rangle |v^{(2)}\rangle = |v^{(1)} v^{(2)}\rangle \in \mathcal{H}.$$

Don't worry yet about what $\mathcal{H}_1 \otimes \mathcal{H}_2$ is. The tensor product also satisfies the following:

- The tensor product is **symmetric** – For all $|v^{(1)}\rangle \in \mathcal{H}_1$, $|v^{(2)}\rangle \in \mathcal{H}_2$ we have

$$|v^{(1)}\rangle |v^{(2)}\rangle = |v^{(2)}\rangle |v^{(1)}\rangle.$$

- The tensor product is **linear** – If $|v^{(1)}\rangle \in \mathcal{H}_1$ is given by

$$|v^{(1)}\rangle = \alpha |w_1^{(1)}\rangle + \beta |w_2^{(1)}\rangle$$

for $\alpha, \beta \in \mathbb{C}$ and $|w_1^{(1)}\rangle, |w_2^{(1)}\rangle \in \mathcal{H}_1$ then the tensor product of $|v^{(1)}\rangle$ with $|v^{(2)}\rangle \in \mathcal{H}_2$ is

$$|v^{(1)}\rangle |v^{(2)}\rangle = \left[\alpha |w_1^{(1)}\rangle + \beta |w_2^{(1)}\rangle \right] |v^{(2)}\rangle$$
$$= \alpha |w_1^{(1)}\rangle |v^{(2)}\rangle + \beta |w_2^{(1)}\rangle |v^{(2)}\rangle.$$

Between these two properties we also have that the tensor product is linear in the second argument as well. Note that upper indices in brackets denote the vector space from where our vectors come and other indices are just part of the label that we use for convenience.

The set of vectors

$$\{ |v^{(1)}\rangle |v^{(2)}\rangle \mid |v^{(1)}\rangle \in \mathcal{H}_1, |v^{(2)}\rangle \in \mathcal{H}_2 \}$$

spans a vector space, $\mathcal{H}$, called the tensor product of $\mathcal{H}_1$ and $\mathcal{H}_2$. This vector space is denoted $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$.

If $\mathcal{B}_1 = \{ |e_k^{(1)}\rangle \}$ is a basis of $\{ |e_k^{(2)}\rangle \}$ are bases of $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively then

$$\mathcal{B} = \{ |e_k^{(1)} e_l^{(2)}\rangle \mid |e_k^{(1)}\rangle \in \mathcal{H}_1, |e_l^{(2)}\rangle \in \mathcal{H}_2 \}$$

is a basis of $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. This means that for all $|\psi\rangle \in \mathcal{H}$ there exists $c_{kl} \in \mathbb{C}$ such that

$$|\psi\rangle = \sum_{k,l} c_{kl} |e_k^{(1)} e_l^{(2)}\rangle$$

A simple counting argument shows that if $\mathcal{H}_1$ and $\mathcal{H}_2$ are finite dimensional then

$$\dim \mathcal{H} = |\mathcal{B}| = |\mathcal{B}_1| |\mathcal{B}_2| = \dim(\mathcal{H}_1) \dim(\mathcal{H}_2).$$

2.3.2 Why are Tensor Products Useful?

Suppose that a system is made of two composite systems, 1 and 2, which have $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively as their space of states. Then the composite system composed of both systems has $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ as its space of states.

2.3.3   Operators acting on Tensor Product Spaces

Say we have an operator, $\hat{O}^{(1)} \colon \mathcal{H}_1 \to \mathcal{H}_1$, defined such that

$$\hat{O}^{(1)}|v^{(1)}\rangle = |w^{(1)}\rangle$$

for $|v^{(1)}\rangle, |w^{(1)}\rangle \in \mathcal{H}_1$. We can define a new operator, $\hat{O}^{(1)} \colon \mathcal{H} \to \mathcal{H}$, such that

$$\hat{O}^{(1)}|v^{(1)}v^{(2)}\rangle = |w^{(1)}v^{(2)}\rangle.$$

Strictly speaking this is a new operator as its (co)domain is different but we use the same symbol as long as there is no chance of confusion. We can similarly define an operator, $\hat{O}^{(2)} \colon \mathcal{H}_2 \to \mathcal{H}_2$, defined by

$$\hat{O}^{(2)}|v^{(2)}\rangle = |w^{(2)}\rangle$$

for $|v^{(2)}\rangle, |w^{(2)}\rangle \in \mathcal{H}_2$. Again this can be extended to $\hat{O}^{(2)} \colon \mathcal{H} \to \mathcal{H}$ as

$$\hat{O}^{(2)}|v^{(1)}v^{(2)}\rangle = |v^{(1)}w^{(2)}\rangle.$$

Consider an operator, $\hat{O}^{(1)} \colon \mathcal{H}_1 \to \mathcal{H}_1$ defined by specifying its action on orthonormal basis vectors, $\mathcal{B}_1 = \{|e_k\rangle\}$:

$$\hat{O}^{(1)}|e_k^{(1)}\rangle = \sum_l |e_l^{(1)}\rangle\langle e_l^{(1)}|\hat{O}^{(1)}|e_k^{(1)}\rangle = \sum_l |e_l^{(1)}\rangle O_{mk}.$$

Here the first equality is achieved by the completeness relation and the second by the definition of the matrix elements of an operator. In the basis $\mathcal{B} = \{|e_k^{(1)}e_l^{(2)}\rangle\}$ we can write $|v\rangle \in \mathcal{H}$ as

$$|v\rangle = \sum_{k,l} v_{kl}|e_k^{(1)}e_l^{(2)}\rangle.$$

Using the linearity properties of the operator, $\hat{O}^{(1)}$, extended to act on $\mathcal{H}$, we get

$$\begin{aligned}
|v'\rangle &= \hat{O}^{(1)}|v\rangle \\
&= \hat{O}^{(1)}\sum_{k,l} v_{kl}|e_k^{(1)}e_l^{(2)}\rangle \\
&= \sum_{k,l} v_{kl}\hat{O}^{(1)}|e_k^{(1)}e_l^{(2)}\rangle \\
&= \sum_{k,l} v_{kl}\sum_m |e_m^{(1)}e_l^{(2)}\rangle O_{mk} \\
&= \sum_{m,l}\left[\sum_k O_{mk}v_{kl}\right]|e_m^{(1)}e_l^{(2)}\rangle.
\end{aligned}$$

Expanding $|v'\rangle$ in $\mathcal{B}$ we get

$$|v'\rangle = \sum_{m,l} v'_{ml}|e_m^{(1)}e_l^{(2)}\rangle,$$

which gives us

$$v'_{ml} = \sum_k O_{mk}v_{kl}.$$

We can do something similar for an operator initially defined on $\mathcal{H}_2$.

2.4   Two-State Systems

Two-state systems are quantum systems whose physical states live in a 2-dimensional complex vector space, $\mathcal{H}$. As a 2-dimensional vector space any basis has exactly two elements. For example the following are all two-state systems:

- A spin 1/2 particle can have either spin up, $|\uparrow\rangle$, or spin down, $|\downarrow\rangle$, or a superposition of both:

$$|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle.$$

- Light can be polarised in two orthogonal directions, vertical, $|\,|\rangle$, or horizontal, $|-\rangle$, or a superposition of both:

$$|\psi\rangle = \alpha|\,|\rangle + \beta|-\rangle.$$

- An atom with two very closely spaced energy levels, E_1 and E_2, can be viewed as comprising of *only* these two energy levels if the other energy levels are sufficiently far away. Thus an electron in this atom can be in the first energy level, $|E_1\rangle$, or the second, $|E_2\rangle$, or a superposition of both:

$$|\psi\rangle = \alpha|E_1\rangle + \beta|E_2\rangle.$$

- A qubit can be in either an on, $|1\rangle$, or an off, $|0\rangle$, state, or a superposition of both:

$$|\psi\rangle = \alpha|1\rangle + \beta|0\rangle.$$

Let $\mathcal{H}$ be a 2-dimensional vector space with an orthonormal basis, $\mathcal{B} = \{|1\rangle, |2\rangle\}$. A generic state, $|v\rangle \in \mathcal{H}$, can be written as

$$|v\rangle = v_1|1\rangle + v_2|2\rangle$$

where $v_1, v_2 \in \mathbb{C}$. We also require that all states are normalised so

$$|v_1|^2 + |v_2|^2 = 1.$$

Quantum states are only defined up to a complex phase so we impose an additional requirement that $v_1 \in \mathbb{R}$. This means that the four degrees of freedom that we start with (the real and complex components of v_1 and v_2) have been reduced by these two requirements to 2 degrees of freedom. As such we should be able to characterise the state by two real numbers. And indeed we can, one way to do this is

$$v_1 = \cos\frac{\vartheta}{2}, \qquad \text{and} \qquad v_2 = e^{i\varphi}\sin\frac{\vartheta}{2}.$$

The fact that we use $\vartheta/2$ instead of just ϑ is just a convention. Since ϑ and φ are angles it makes sense to view them in spherical coordinates as S^2, the sphere with radius 1. When we do this we call it the **Bloch sphere**, it is shown in figure 2.1. We see that the pure states, $|1\rangle$ and $|2\rangle$, are at the poles which can easily be shown by noting that $|1\rangle$ corresponds to $\vartheta = 0$ and $|2\rangle$ corresponds to $\vartheta = \pi$.
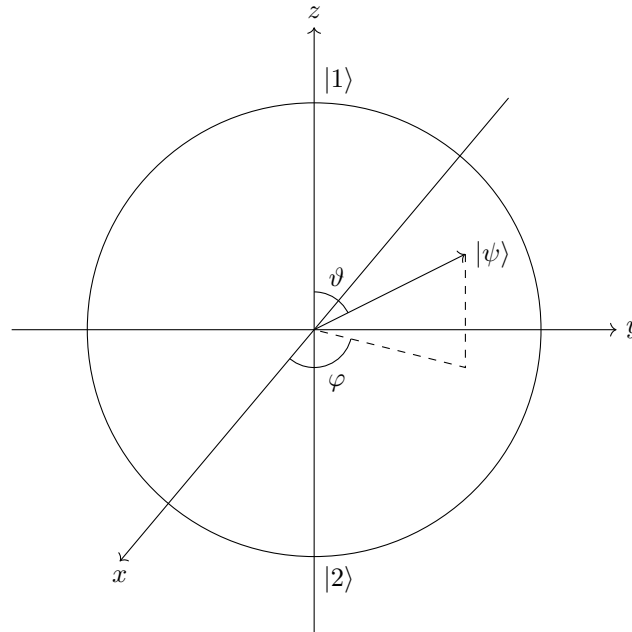
Figure 2.1: The Bloch Sphere, the two pure states, $|1\rangle$ and $|2\rangle$, are at the poles.

2.4.1 Two Two-State Systems

Suppose we have two spin $1/2$ particles. What is the space of states of this two particle system's spins? Each particle can be described by a two-state system $\mathcal{H}_i$ where i represents a quantity relating to particle i. $\mathcal{H}_i$ has an orthonormal basis of

$$\mathcal{B}_i = \{|1^{(i)}\rangle, |2^{(i)}\rangle\}$$

The state space of the whole two particle system is then $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. We can easily find a basis for this space:

$$\mathcal{B} = \{|1^{(1)}1^{(2)}\rangle, |1^{(1)}2^{(2)}\rangle, |2^{(1)}1^{(2)}\rangle, |2^{(1)}2^{(2)}\rangle\}.$$

From now on we suppress the upper index and take the first number to refer to particle 1 and the second number to refer to particle 2. For example $|1^{(1)}2^{(2)}\rangle$ becomes $|12\rangle$. Thus our basis is

$$\mathcal{B} = \{|11\rangle, |12\rangle, |21\rangle, |22\rangle\}.$$

A generic state, $|\psi\rangle \in \mathcal{H}$ can be written as

$$|\psi\rangle = c_{11}|11\rangle + c_{12}|12\rangle + c_{21}|21\rangle + c_{22}|22\rangle = \sum_{k,l} c_{kl}|kl\rangle,$$

where $c_{kl} \in \mathbb{C}$. We require this state to be normalised so that

$$\|\psi\| = |c_{11}|^2 + |c_{12}|^2 + |c_{21}|^2 + |c_{22}|^2 = \sum_{k,l} |c_{kl}|^2 = 1.$$

Suppose that $|\varphi^{(1)}\rangle \in \mathcal{H}_1$ and $|\varphi^{(2)}\rangle \in \mathcal{H}_2$ are normalised states. This means that we can write them as

$$|\varphi^{(1)}\rangle = c_1^{(1)}|1^{(1)}\rangle + c_2^{(1)}|2^{(1)}\rangle$$

and

$$|\varphi^{(2)}\rangle = c_1^{(2)}|1^{(2)}\rangle + c_2^{(2)}|2^{(2)}\rangle$$

where $c_k^{(l)} \in \mathbb{C}$ and

$$\left\|\varphi^{(1)}\right\|^2 = \left\|\varphi^{(2)}\right\|^2 = \left|c_1^{(1)}\right|^2 + \left|c_2^{(1)}\right|^2 = \left|c_1^{(2)}\right|^2 + \left|c_2^{(2)}\right|^2 = 1.$$

We can now compute the tensor product of these two objects:

$$|\varphi^{(1)}\varphi^{(2)}\rangle = \left[c_1^{(1)}|1^{(1)}\rangle + c_2^{(1)}|2^{(1)}\rangle\right] \left[c_1^{(2)}|1^{(2)}\rangle + c_2^{(2)}|2^{(2)}\rangle\right]$$
$$= c_1^{(1)}c_1^{(2)}|11\rangle + c_1^{(1)}c_2^{(2)}|12\rangle + c_2^{(1)}c_1^{(2)}|21\rangle + c_2^{(1)}c_2^{(2)}|22\rangle$$

The norm of this vector is

$$\left\|\varphi^{(1)}\varphi^{(2)}\right\| = \left|c_1^{(1)}c_1^{(2)}\right|^2 + \left|c_1^{(1)}c_2^{(2)}\right|^2 + \left|c_2^{(1)}c_1^{(2)}\right|^2 + \left|c_2^{(1)}c_2^{(2)}\right|^2$$
$$= \left|c_1^{(1)}\right|^2\left|c_1^{(2)}\right|^2 + \left|c_1^{(1)}\right|^2\left|c_2^{(2)}\right|^2 + \left|c_2^{(1)}\right|^2\left|c_1^{(2)}\right|^2 + \left|c_2^{(1)}\right|^2\left|c_2^{(2)}\right|^2$$
$$= \left|c_1^{(1)}\right| \left(\left|c_1^{(2)}\right|^2 + \left|c_2^{(2)}\right|^2\right) + \left|c_2^{(1)}\right| \left(\left|c_1^{(2)}\right|^2 + \left|c_2^{(2)}\right|^2\right)$$
$$= \left|c_1^{(1)}\right|^2\left\|\varphi^{(2)}\right\|^2 + \left|c_2^{(1)}\right|^2\left\|\varphi^{(2)}\right\|^2$$
$$= \left|c_1^{(1)}\right|^2 + \left|c_2^{(1)}\right|^2$$
$$= \left\|\varphi^{(1)}\right\|^2$$
$$= 1$$

So, given that $|\varphi^{(1)}\rangle$ and $|\varphi^{(2)}\rangle$ are normalised, $|\varphi^{(1)}\varphi^{(2)}\rangle$ is normalised.

It is important to note that not all $|\psi\rangle \in \mathcal{H}$ can be written as a tensor product of one vector from $\mathcal{H}_1$ and one vector from $\mathcal{H}_2$. For example this is not possible if

$$|\psi\rangle = \frac{1}{\sqrt{2}}|12\rangle + \frac{1}{\sqrt{2}}|21\rangle.$$

States like this, that cannot be written as a tensor product, are called **entangled states** and we will look at them in more detail later.

3   One-Dimensional System

Consider a one-dimensional system consisting of a line. The dimension here refers to the spatial dimension, not the dimension of the space of states. Any point on this line can be uniquely given by a coordinate, x. The state of the system is described by a **wave function**, $\psi\colon \mathbb{R} \to \mathbb{C}$. All information about the physical system is encoded in this function.

3.1   Discretised System

To understand how this fits with our discussion so far we consider a discretised system. We restrict our system to allow only an (arbitrary) number, N, of points, x_i. We choose $N = 6$. We also assume that they are evenly spaced a distance ε apart. See figure 3.1. We can now limit the domain of the
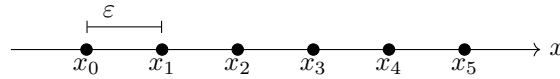
Figure 3.1: A one dimensional discrete system of $N = 6$ points, x_i, spaced evenly ε apart.

wave function to $\{x_i \mid i = 0, \ldots, 5\}$. We define

$$\psi_i = \sqrt{\varepsilon}\psi(x_i) \in \mathbb{C}.$$

We then collect this set of evaluations of $\psi(x)$ into one object:

$$|\psi\rangle = \begin{pmatrix} \psi_0 \\ \vdots \\ \psi_5 \end{pmatrix}$$

$|\psi\rangle$ is a member of a six-dimensional vector space, $\mathcal{H}$. Note that it is wrong to write $|\psi(x)\rangle$. $|\psi\rangle$ and $\psi(x)$ are different objects writing $|\psi(x)\rangle$ is similar to putting a coordinate label on a vector like $\boldsymbol{v}_x$, its just wrong.

3.1.1   Basis

By specifying a vector by its components we assume a basis. One such basis is

$$|\xi_0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad |\xi_1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, |\xi_5\rangle = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Any vector $\psi \in \mathcal{H}$ can be written as a linear combination of these basis vectors:

$$|\psi\rangle = \sum_{k=0}^{5} \psi_k |\xi_k\rangle$$

where as usual

$$\psi_k = \langle \xi_k | \psi \rangle.$$

For now we will make some assumptions about extracting physical information from state vectors and wave functions. We assume that $|\psi(x)|^2$ gives the probability density of finding the system at x. We also assume that $|\psi_i|^2 = \varepsilon|\psi(x_i)|^2$ gives the probability of finding the system at x_i. Note that the first of these is a probability density and the second is a probability. This is because the first is referring to a continuous variable, x, and the second refers to a discrete variable, x_i. This is the same as how a probability density function, p, gives a probability, $p(x)\,\mathrm{d}x$, here ε plays the role of $\mathrm{d}x$.

With these assumptions we have a physical description of the basis states. The vector $|\xi_i\rangle$ represents a system that is at x_i with probability 1.

3.1.2 Scalar Product and Norm

We can define a scalar product between $|\varphi\rangle, |\psi\rangle \in \mathcal{H}$

$$\langle\varphi|\psi\rangle = \sum_{k=0}^{6} \varphi_k^* \psi_k.$$

This allows us to induce a norm:

$$\|\psi\|^2 = \sum_{k=0}^{6} |\psi_k|^2.$$

As usual we require state vectors be normalised so

$$\|\psi\|^2 = \sum_{k=0}^{6} |\psi_k|^2 = 1.$$

3.2 Continuous System

Moving from this discretised system to a continuous one is as simple as taking the limiting case as $N \to \infty$ and $\varepsilon \to 0$. After doing this we have that $|\psi\rangle \in \mathcal{H}_\infty$ where $\mathcal{H}_\infty$ is an infinite dimensional vector space.

3.2.1 Basis

Define the following basis vectors for our discretised system:

$$|x_k\rangle = \frac{1}{\sqrt{\varepsilon}}|\xi_k\rangle.$$

Which gives us $|\xi_k\rangle = \sqrt{\varepsilon}|x_k\rangle$ A vector in the finite dimensional vector space, $\mathcal{H}$, can be written as

$$|\psi\rangle = \sum_{k=0}^{N} \psi_k |\xi_k\rangle = \sum_{k=0}^{N} \sqrt{\varepsilon}\psi(x_k)|\xi_k\rangle = \sum_{k=0}^{N} \varepsilon\psi(x_k)|x_k\rangle.$$

After a limiting process taking $N \to \infty$ and $\varepsilon \to 0$ we can write $|\psi\rangle \in \mathcal{H}_\infty$ as

$$|\psi\rangle = \int \mathrm{d}x\, \psi(x)|x\rangle$$

where

$$|x\rangle = \lim_{\varepsilon \to 0} \frac{1}{\sqrt{\varepsilon}}|\xi_k\rangle.$$

As normal we can find the component, here $\psi(x)$, with a scalar product with a basis vector:

$$\psi(x) = \langle x|\psi\rangle.$$

Using this we have

$$|\psi\rangle = \int \mathrm{d}x\, \langle x|\psi\rangle|x\rangle = \int \mathrm{d}x\, |x\rangle\langle x|\psi\rangle.$$

We see that $|x\rangle\langle x|$ is a projection operator and again the completeness relation holds, this time in the form

$$\int \mathrm{d}x\, |x\rangle\langle x| = \mathbb{1}.$$

3.2.2   Scalar Product and Norm

Let $|\psi\rangle, |\varphi\rangle \in \mathcal{H}$. Then their scalar product is

$$\langle\varphi|\psi\rangle = \sum_{k=0}^{N} \varphi_k^* \psi_k = \sum_{k=0}^{N} [\sqrt{\varepsilon}\varphi^*(x_k)][\sqrt{\varepsilon}\psi(x_k)] = \sum_{k=0}^{N} \varepsilon\varphi^*(x_k)\psi(x_k).$$

After a limiting process taking $N \to \infty$ and $\varepsilon \to 0$ we see that for $|\varphi\rangle, |\psi\rangle \in \mathcal{H}_\infty$ we have

$$\langle\varphi|\psi\rangle = \int \mathrm{d}x\, \varphi^*(x)\psi(x).$$

The norm induced by this for $|\psi\rangle \in \mathcal{H}_\infty$ is given by

$$\|\psi\|^2 = \langle\psi|\psi\rangle = \int \mathrm{d}x\, \psi^*(x)\psi(x) = \int \mathrm{d}x\, |\psi(x)|^2.$$

For states to be normalisable we require that $\|\psi\| < \infty$. The space of functions, $f\colon A \to \mathbb{C}$, that have this property is called the space of **square integrable functions** and is denoted $\mathrm{L}^2(A)$. For our example of a continuous system we are interested in $\mathrm{L}^2(\mathbb{R})$. If we limit our system to the interval $[a, b]$ where $a, b \in \mathbb{R}$ and $b > a$ then we are interested in the space $\mathrm{L}^2([a, b])$.

3.2.3   Scalar Product of Basis Vectors

In the discrete system $|\xi_k\rangle$ describes a state that is a x_k with probability 1. We want to know what $|x\rangle$ describes in our continuous case. Consider the inner product of two basis states:

$$\langle x|x'\rangle = \lim_{\varepsilon\to 0}\langle x_k|x_{k'}\rangle = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\langle\xi_k|\xi_{k'}\rangle.$$

Using the orthonormality of $\{\xi_k\}$ this gives us

$$\langle x|x'\rangle = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\delta_{kk'} = \begin{cases} 0, & x \neq x', \\ \infty, & x = x', \end{cases}$$

where we have used the fact that if $x = x'$ then $|\xi_k\rangle = |\xi_{k'}\rangle$ so $\delta_{kk'} = \delta_{kk} = 1$ so

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon} = \infty.$$

On the other hand if $x \neq x'$ then $|\xi_k\rangle \neq |\xi_{k'}\rangle$ so $\delta_{kk'} = 0$. Then in the limit we have

$$\lim_{\varepsilon\to 0}\frac{0}{\varepsilon} = \lim_{\varepsilon\to 0}\frac{\frac{\mathrm{d}}{\mathrm{d}\varepsilon}0}{\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\varepsilon} = \lim_{\varepsilon\to 0}\frac{0}{1} = 0.$$

We now compute the following integral for reasons that will become clear afterwards:

$$\begin{aligned}
\int \mathrm{d}x'\, f(x')\langle x|x'\rangle &= \lim_{\substack{N\to\infty \\ \varepsilon\to 0}} \sum_{k=0}^{N} \varepsilon f(x_{k'})\langle x_k|x_k\rangle \\
&= \lim_{\substack{N\to\infty \\ \varepsilon\to 0}} \sum_{k=0}^{N} \varepsilon f(x_{k'})\frac{1}{\varepsilon}\langle\xi_k|\xi_{k'}\rangle \\
&= \lim_{N\to\infty} \sum_{k=0}^{N} f(x_k)\langle\xi_k|\xi_{k'}\rangle \\
&= \lim_{N\to\infty} f(x_{k'})\delta_{kk'} \\
&= \lim_{\varepsilon\to 0} f(x_k) \\
&= f(x)
\end{aligned}$$

So we see that

$$\int \mathrm{d}x'\, f(x')\langle x|x'\rangle = f(x)$$

this, as well as being zero if $x \neq x'$ defines the **Dirac delta distribution**:

$$\delta(x - x') = \begin{cases} 0, & x \neq x', \\ \infty, & x = x' \end{cases}$$

and

$$\int \mathrm{d}x' \, f(x')\delta(x - x') = f(x).$$

So we can identify

$$\langle x | x' \rangle = \delta(x - x').$$

Using this we get around the fact that $\{|x\rangle\}$ are not normalisable, that is that $\langle x | x \rangle = \infty$, and we can still extract meaningful information as long as we only have $\langle x | x \rangle$ inside of integrals.

Part II

Observables

4   Observables

In QM the outcome of a measurement is the outcome of a stochastic variable. The theory can only predict the probability density function (PDF) of the outcome of a given experiment.

4.1   Some Statistics

A continuous random, or stochastic, variable is one that can take on a different value every time we measure it. The values that it can take can be characterised by a **probability density function (PDF)**. If x is a continuous random variable and can take any value in some set S then the PDF of x is a function

$$p \colon S \to [0, 1]$$

such that the probability that x is between x and $x + \mathrm{d}x$ is given by $p(x) \, \mathrm{d}x$. Using this we have that the probability that x is in the interval $\mathcal{I} = [a, b]$ is

$$P(x \in \mathcal{I}) = \int_a^b \mathrm{d}x \, p(x).$$

Since probabilities should add up to one we require p to be properly normalised so that

$$P(x \in S) = \int_S \mathrm{d}x \, p(x) = 1.$$

As a way of characterising a PDF we define its n**th moment** as

$$\mu_n = \int \mathrm{d}x \, p(x) x^n.$$

Clearly if p is properly normalised then $\mu_0 = 1$. The **mean** value of the random variable, x, is

$$\mu_1 = \langle x \rangle = \int \mathrm{d}x \, p(x) x.$$

The **variance** of the variable, x, is

$$\mathrm{Var}[x] = \mu_2 - (\mu_1)^2 = \langle x^2 \rangle - \langle x \rangle^2.$$

4.2   Position Operator

The state of the system is described by $\psi(x) = \langle x|\psi\rangle$ where $|\psi\rangle \in \mathcal{H}$, with $\mathcal{H}$ being some infinite dimensional vector space that has $\{|x\rangle\}$ as a basis. The PDF for finding a system in the state $|\psi\rangle \in \mathcal{H}$ at position x is

$$p(x) = |\psi(x)|^2.$$

The **expected value** of x is

$$\langle x\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\, p(x)x.$$

We introduce an operator, $\hat{X}\colon \mathcal{H} \to \mathcal{H}$, defined by

$$\langle x|\hat{X}|\psi\rangle = x\langle x|\psi\rangle,$$

or, equivalently,

$$\hat{X}\psi(x) = x\psi(x).$$

This allows us to write the mean value as

$$
\begin{aligned}
\langle x\rangle &= \int \mathrm{d}x\, |\psi(x)|^2 x \\
&= \int \mathrm{d}x\, \psi(x)^* x\psi(x) \\
&= \int \mathrm{d}x\, \langle\psi|x\rangle x\langle x|\psi\rangle \\
&= \int \mathrm{d}x\, \langle\psi|x\rangle\langle x|\hat{X}|\psi\rangle \\
&= \langle\psi|\hat{X}|\psi\rangle \\
&= \langle\hat{X}\rangle.
\end{aligned}
$$

Here we used the completeness relation to get rid of the integral and projection operators. Note that $\langle\hat{X}\rangle$ is just a short hand for $\langle\psi|\hat{X}|\psi\rangle$ when the state ψ is clear. It is not the expected value of the operator $\hat{X}$ per se but the expected value of x which is in some way related to the operator $\hat{X}$.

There is a theorem that states that if we know all the moments of a PDF then this uniquely determines the PDF. We can find the nth moment fairly easily:

$$
\begin{aligned}
\langle\psi|\hat{X}^n|\psi\rangle &= \int \mathrm{d}x\, \langle\psi|x\rangle\langle x|\hat{X}^n|\psi\rangle \\
&= \int \mathrm{d}x\, \langle\psi|x\rangle x^n \langle x|\psi\rangle \\
&= \int \mathrm{d}x\, \psi(x)^* x^n \psi(x) \\
&= \int \mathrm{d}x\, |\psi(x)|^2 x^n \\
&= \int \mathrm{d}x\, p(x)x^n \\
&= \mu_n.
\end{aligned}
$$

So by taking the matrix elements of $\hat{X}^n$ we can fully describe the PDF.

Here we have used the following theorem.

Theorem 1

If $\hat{O}\colon \mathcal{H} \to \mathcal{H}$ is an operator and acts on $|\psi\rangle \in \mathcal{H}$ as $\hat{O}|\psi\rangle = z|\psi\rangle$ where $z \in \mathbb{C}$ then $\hat{O}^n|\psi\rangle = z^n|\psi\rangle$.

Proof. We will prove this by induction. The basis case of $n = 0$ holds as by definition $\hat{O}^0 = \mathbb{1}$ so

$$\hat{O}^0|\psi\rangle = \mathbb{1}|\psi\rangle = |\psi\rangle = 1|\psi\rangle = z^0|\psi\rangle.$$

Now suppose that this holds for $n = k$, that is

$$\hat{O}^k|\psi\rangle = z^k|\psi\rangle.$$

Then

$$\begin{aligned}
\hat{O}^{k+1}|\psi\rangle &= \hat{O}(\hat{O}^k|\psi\rangle) \\
&= \hat{O}(z^k|\psi\rangle) \\
&= z^k(\hat{O}|\psi\rangle) \\
&= z^k(z|\psi\rangle) \\
&= z^{k+1}|\psi\rangle.
\end{aligned}$$

So it holds for $n = k + 1$. Thus by induction this holds for all $n \in \mathbb{N}$. □

4.3   Generic Observables

We define an **observable** as a quantity that can be measured in an experiment. Here an experiment is a black box that we provide somehow with a prepared state, $|\psi\rangle$, and after the experiment we are given a value for the observable that we were measuring. One important thing that we assume is that the experiment is reproducible, that is that we can prepare the system in the same way and perform the same measurement as many times as needed.

This gives rise to the second postulate of QM:

> **Postulate 2**
>
> In quantum mechanics each observable, O, is associated to a linear operator, $\hat{O} \colon \mathcal{H} \to \mathcal{H}$.

The naive approach to find the expected value of an observable, O, given a specific state, $|\psi\rangle$, might be to try $\hat{O}|\psi\rangle$. However this gives us another vector and we expect real numbers as the results of measurements. It turns out that the expected value of O for a system in the state $|\psi\rangle$ is given by the matrix elements of $\hat{O}$:

$$\langle O \rangle = \langle \psi | O | \psi \rangle.$$

4.3.1   Observing Observables

When we say that the value of an observable O is a stochastic variable is that if we set up the system in state $|\psi\rangle$ and measure O and get a result $O^{(1)}$, and then set up the system in state $|\psi\rangle$ and measure O and get a result $O^{(2)}$, and so on, then we end up with a set of values $\{O^{(1)}, O^{(2)}, \ldots, O^{(n)}\}$. The important thing is that it is possible that $O^{(i)} \neq O^{(j)}$, even accounting for experimental error. This is very different to classical mechanics where we would expect that a measurement gives the same result every time, up to some level of precision.

The expectation value of O is then

$$\langle O \rangle = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} O^{(k)}.$$

So far we have focused on the value of the observable but with QM we can predict more than this. Given an observable, O, we can predict

- the possible outcomes of a measurement of O,
- the probability of obtaining each of these possible outcomes.

All questions in QM need to be phrased in terms of possible outcomes and their probabilities, these are the only things that the theory predicts.

4.4   Eigenvalues and Eigenstates

The possible outcomes of measuring an observable, O, are given by the third postulate of QM:

> **Postulate 3**
>
> When measuring an observable, O, the possible outcomes, O_k, are the eigenvalues of the operator, $\hat{O}$.

If $\hat{O} \colon \mathcal{H} \to \mathcal{H}$ and $|\psi_k\rangle \in \mathcal{H}$ then $|\psi_k\rangle$ is an **eigenvector** or **eigenstate** with **eigenvalue** O_k if these are solutions to the following:

$$\hat{O}|\psi_k\rangle = O_k|\psi_k\rangle.$$

That is the only result of $\hat{O}$ when acting on $|\psi_k\rangle$ is to scale $|\psi_k\rangle$ by some amount. In general in a complex vector space eigenvalues are complex, we will see later that this does not give rise to the correct physics.

If O_k is an eigenvalue of the operator, $\hat{O}$, then we say that O_k is g-fold degenerate if there are exactly g linearly independent eigenvectors, $|u_k^{(n)}\rangle \in \mathcal{H}$ such that

$$\hat{O}|u_k^{(n)}\rangle = O_k|u_k^{(n)}\rangle$$

for $n = 1, \ldots, g$. That is that O_k is an eigenvalue for g fundamentally different eigenvectors (as opposed to scaled versions of the same eigenvector).

Any linear combination of the form

$$\sum_{n=1}^{g} c_n |u_k^{(n)}\rangle$$

where $c_k \in \mathbb{C}$ is also an eigenstate of $\hat{O}$, this is fairly easy to show using the linearity of $\hat{O}$:

$$\begin{aligned}
\hat{O} \sum_{n=1}^{g} c_n |u_k^{(n)}\rangle &= \sum_{n=1}^{g} c_n \hat{O}|u_k^{(n)}\rangle \\
&= \sum_{n=1}^{g} c_n O_k |u_k^{(n)}\rangle \\
&= O_k \sum_{n=1}^{g} c_n |u_k^{(n)}\rangle.
\end{aligned}$$

In fact any set of eigenvectors with degenerate eigenvalues forms a g-dimensional vector subspace of $\mathcal{H}$.

One logical question that we may ask now is what do the eigenstates represent? We can try to measure the expected value of O in a system that is the state $|\psi_k\rangle$ which is an eigenstate of $\hat{O}$ with eigenvalue O_k. We get

$$\begin{aligned}
\langle O \rangle &= \langle \psi_k | \hat{O} | \psi_k \rangle \\
&= \langle \psi_k | O_k | \psi_k \rangle \\
&= O_k \langle \psi_k | \psi_k \rangle \\
&= O_k.
\end{aligned}$$

So the expected value of O measured on its eigenstate, $|\psi_k\rangle$, is O_k. We can also compute the variance:

$$\begin{aligned}
\mathrm{Var}_{\psi_k}[O] &= \langle O^2 \rangle - \langle O \rangle^2 \\
&= \langle \psi_k | \hat{O}^2 | \psi_k \rangle - [\langle \psi_k | \hat{O} | \psi_k \rangle]^2 \\
&= \langle \psi_k | O_k^2 | \psi_k \rangle - [\langle \psi_k | O_k | \psi_k \rangle]^2 \\
&= O_k^2 \langle \psi_k | \psi_k \rangle - [O_k \langle \psi_k | \psi_k \rangle]^2 \\
&= O_k^2 - O_k^2 \\
&= 0.
\end{aligned}$$

where we have used the result of theorem 1, which states $\hat{O}^2|\psi_k\rangle = O_k^2|\psi_k\rangle$ so

$$\langle\psi_k|\hat{O}^2|\psi_k\rangle = \langle\psi_k|O_k^2|\psi_k\rangle = O_k^2\langle\psi_k|\psi_k\rangle.$$

The important thing is that the expected value is O_k and the variance is zero. This means that there is no spread in the values measured and so $O^{(i)} = O_k$ for all i. This means that in the state $|\psi_k\rangle$ when we measure O we get O_k with probability 1.

4.5 Hermitian Operators

4.5.1 Hermitian Conjugate

Let $\hat{O}\colon \mathcal{H} \to \mathcal{H}$. Then we define the **hermitian conjugate** of $\hat{O}$, denoted $\hat{O}^\dagger$, to be another operator, $\hat{O}^\dagger\colon \mathcal{H} \to \mathcal{H}$, defined such that for all $|\varphi\rangle, |\psi\rangle \in \mathcal{H}$

$$\langle\varphi|\hat{O}^\dagger|\psi\rangle = [\langle\psi|\hat{O}|\varphi\rangle]^*.$$

Compare this to the usual linear algebra definition:

$$O_{ij}^\dagger = O_{ji}^*.$$

We see that since $\langle\varphi|\hat{O}^\dagger|\psi\rangle$ gives the matrix elements of $\hat{O}^\dagger$ these two definitions are equivalent.

Example 4.1. Consider the operator $\hat{O} = \frac{\mathrm{d}}{\mathrm{d}x}$ defined by

$$\begin{aligned}
\langle x|\hat{O}|\psi\rangle &= \langle x|\frac{\mathrm{d}}{\mathrm{d}x}|\psi\rangle \\
&= \frac{\mathrm{d}}{\mathrm{d}x}\langle x|\psi\rangle \\
&= \frac{\mathrm{d}}{\mathrm{d}x}\psi(x).
\end{aligned}$$

Compute $\hat{O}^\dagger$.

$$\langle\varphi|\hat{O}^\dagger|\psi\rangle = [\langle\psi|\hat{O}|\varphi\rangle]^*$$

inserting completeness gives

$$\begin{aligned}
&= \left[\int_{-\infty}^{\infty} \mathrm{d}x\,\langle\psi|x\rangle\langle x|\hat{O}|\varphi\rangle\right]^* \\
&= \left[\int_{-\infty}^{\infty} \mathrm{d}x\,\psi(x)^*\frac{\mathrm{d}}{\mathrm{d}x}\varphi(x)\right]^*.
\end{aligned} \tag{4.1}$$

We can now integrate by parts:

$$\begin{aligned}
u &= \psi(x)^*, & v &= \varphi(x), \\
u' &= \tfrac{\mathrm{d}}{\mathrm{d}x}\psi(x)^*, & v' &= \tfrac{\mathrm{d}}{\mathrm{d}x}\varphi(x).
\end{aligned}$$

$$\langle\varphi|\hat{O}^\dagger|\psi\rangle = \left[[\psi(x)^*\varphi(x)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \mathrm{d}x\,\varphi(x)\frac{\mathrm{d}}{\mathrm{d}x}\psi(x)^*\right]^*$$

this first term is zero as ψ and φ are wave functions so are square integrable so they vanish at infinity leaving us with

$$\begin{aligned}
&= \left[-\int_{-\infty}^{\infty} \mathrm{d}x\,\varphi(x)\frac{\mathrm{d}}{\mathrm{d}x}\psi(x)^*\right]^* \\
&= -\int_{-\infty}^{\infty} \mathrm{d}x\,\varphi(x)^*\frac{\mathrm{d}}{\mathrm{d}x}\psi(x) \\
&= -\int_{-\infty}^{\infty} \mathrm{d}x\,\langle\varphi|x\rangle\langle x|\hat{O}|\psi\rangle \\
&= -\langle\varphi|\hat{O}|\psi\rangle.
\end{aligned}$$

So we can identify

$$\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^\dagger = -\frac{\mathrm{d}}{\mathrm{d}x}.$$

4.5.2   Hermitian Operator

An operator, $\hat{O}\colon \mathcal{H} \to \mathcal{H}$, is **hermitian** if $\hat{O}^\dagger = \hat{O}$. We saw in example 4.1 that $\frac{\mathrm{d}}{\mathrm{d}x}$ is *not* hermitian as $\hat{O}^\dagger = -\hat{O}$, this property is actually called being **anti-hermitian**.

Example 4.2. Show that $\hat{O} = -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}$ is hermitian. This operator will be important later.

$$\langle\varphi|\hat{O}^\dagger|\psi\rangle = [\langle\psi|\hat{O}|\varphi\rangle]^*$$

$$= \left[\int_{-\infty}^{\infty} \mathrm{d}x\, \langle\psi|x\rangle\langle x|\hat{O}|\varphi\rangle\right]^*$$

$$= \left[\int_{-\infty}^{\infty} \mathrm{d}x\, \psi(x)\left(-i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\right)|\varphi\rangle\right]^*$$

$$= \left[-i\hbar\int_{-\infty}^{\infty} \mathrm{d}x\, \psi(x)\frac{\mathrm{d}}{\mathrm{d}x}\varphi(x)\right]^*$$

$$= i\hbar\int_{-\infty}^{\infty} \mathrm{d}x\, \psi(x)^*\frac{\mathrm{d}}{\mathrm{d}x}\varphi(x)$$

see equation 4.1 for how to do this integral

$$= -i\hbar\int_{-\infty}^{\infty} \mathrm{d}x\, \varphi(x)^*\frac{\mathrm{d}}{\mathrm{d}x}\psi(x)$$

$$= \int_{-\infty}^{\infty} \mathrm{d}x\, \varphi(x)^*\left(-i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\right)\psi(x)$$

$$= \int_{-\infty}^{\infty} \mathrm{d}x\, \langle\varphi|x\rangle\langle x|\hat{O}|\psi\rangle$$

$$= \langle\varphi|\hat{O}|\psi\rangle.$$

So we can identify that

$$\left(-i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\right)^\dagger = -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}$$

so this is a hermitian operator.

4.5.3   Properties of Hermitian Operators

The following important properties of hermitian operators make incredibly important to QM:

1. Hermitian operators have real eigenvalues. That is if $\hat{O}\colon \mathcal{H} \to \mathcal{H}$ and

$$\hat{O}|\psi_k\rangle = O_k|\psi_k\rangle$$

   then

$$\hat{O} = \hat{O}^\dagger \implies O_k \in \mathbb{R}.$$

2. The eigenstates of a hermitian operator that belong to different eigenvalues are orthogonal. That is if $\hat{O}\colon \mathcal{H} \to \mathcal{H}$,

$$\hat{O}|\psi_k\rangle = O_k|\psi_k\rangle, \qquad \text{and} \qquad \hat{O}|\psi_l\rangle = O_l|\psi_l\rangle$$

   where $O_k \neq O_l$ then

$$\langle\psi_k|\psi_l\rangle = 0.$$

3. If $\hat{O}\colon \mathcal{H} \to \mathcal{H}$ then there exists an orthogonal basis of $\mathcal{H}$ made of eigenvectors of $\hat{O}$. In other words $|\psi\rangle \in \mathcal{H}$ can be written as

$$|\psi\rangle = \sum_k c_k|\psi_k\rangle$$

   where $c_k \in \mathbb{C}$ are given by

$$c_k = \langle\psi_k|\psi\rangle.$$

We can prove these fairly easily:

> ### Theorem 2
>
> Hermitian operators have real eigenvalues.

Proof. Suppose $\hat{O}$ is a hermitian operator on $\mathcal{H}$. Let $|\psi_k\rangle$ be a normalised eigenvector of $\hat{O}$ with eigenvalue O_k. Then

$$\hat{O}|\psi_k\rangle = O_k|\psi_k\rangle.$$

Taking an inner product with $\langle\psi_k|$ we get

$$\langle\psi_k|\hat{O}|\psi_k\rangle = \langle\psi_k|O_k|\psi_k\rangle = O_k\langle\psi_k|\psi_k\rangle = O_k \tag{4.2}$$

where we have used the fact that $|\psi_k\rangle$ is normalised so $\langle\psi_k|\psi_k\rangle = 1$. Taking the complex conjugate of this equation gives us

$$[\langle\psi_k|\hat{O}|\psi_k\rangle]^* = O_k^*.$$

Using the definition of the hermitian conjugate the left hand side is equivalent to

$$\langle\psi_k|\hat{O}^\dagger|\psi_k\rangle.$$

Since $\hat{O}$ is hermitian this in turn is equivalent to

$$\langle\psi_k|\hat{O}|\psi_k\rangle,$$

which, as we saw in equation 4.2, is just O_k. This means that we have

$$O_k = O_k^*$$

which can only be true if $O_k \in \mathbb{R}$. $\qquad\square$

> ### Theorem 3
>
> The eigenfunctions of a hermitian operator which belong to different eigenvalues are orthogonal.

Proof. Let $\hat{O}$ be an operator on $\mathcal{H}$. Suppose that

$$\hat{O}|\psi_1\rangle = O_1|\psi_1\rangle, \qquad \text{and} \qquad \hat{O}|\psi_2\rangle = O_2|\psi_2\rangle$$

where $O_1 \neq O_2$. Taking the inner product of the first of these equations with $\langle\psi_2|$ we have

$$\langle\psi_2|\hat{O}|\psi_1\rangle = O_1\langle\psi_2|\psi_1\rangle. \tag{4.3}$$

Instead taking the inner product of the second equation with $\langle\psi_1|$ we have

$$\langle\psi_1|\hat{O}|\psi_2\rangle = O_2\langle\psi_1|\psi_2\rangle.$$

Taking the complex conjugate, using the definition of the hermitian conjugate, and then using the fact that $\hat{O}^\dagger = \hat{O}$ the left hand side of the above equation gives us

$$[\langle\psi_1|\hat{O}|\psi_2\rangle]^* = \langle\psi_2|\hat{O}^\dagger|\psi_1\rangle = \langle\psi_2|\hat{O}|\psi_1\rangle.$$

The right hand side gives

$$O_2^*\langle\psi_1|\psi_2\rangle = O_2\langle\psi_1|\psi_2\rangle,$$

where we have used the fact that $\hat{O}$ is hermitian so its eigenvalues are real and therefore $O_2^* = O_2$. Comparing this with equation 4.3 we see that

$$O_2\langle\psi_2|\psi_1\rangle = O_1\langle\psi_2|\psi_1\rangle$$

which gives us

$$(O_2 - O_1)\langle\psi_2|\psi_1\rangle = 0.$$

Given that $\mathbb{C}$ is a field and so has no zero divisors either $O_2 - O_1 = 0$ or $\langle\psi_2|\psi_1\rangle = 0$. By our assumption $O_2 - O_1 \neq 0$ so we must have $\langle\psi_2|\psi_1\rangle = 0$. This means that $|\psi_1\rangle$ and $|\psi_2\rangle$ are orthogonal. $\qquad\square$

The reason that the first of these properties, that the eigenvalues are real, is important is that we expect real results from an experiment and therefore it only makes sense to have the eigenvalues of an observable be real. The other two properties become useful when doing theory as working in the basis given by the eigenvectors, called the eigenbasis, can simplify a lot of equations. The process of moving into this eigenbasis is called spectral decomposition.

4.5.4   Spectral Decomposition

Postulate 4

An observable, O, has an associated hermitian linear operator, $\hat{O} \colon \mathcal{H} \to \mathcal{H}$, which has a nonde-generate eigenvalue O_k and associated eigenvector $|\psi_k\rangle$. Given a normalised state, $|\psi\rangle \in \mathcal{H}$, the probability of measuring O to be O_k is

$$P_k = |c_k|^2 = |\langle \psi_k | \psi \rangle|^2$$

where $c_k \in \mathbb{C}$ are the coordinates of $|\psi\rangle$ in the basis made of the eigenvectors of $\hat{O}$:

$$|\psi\rangle = \sum_k c_k |\psi_k\rangle.$$

In the case that O_k is g-fold degenerate and has the vectors $\{|u_k^{(n)}\rangle\}$ as its eigenvectors then the probability of measuring O to be O_k is instead the sum of the contributions from the whole subspace spanned by $\{|u_k^{(n)}\rangle\}$:

$$P_k = \sum_{n=1}^{g} \left| \langle u_k^{(n)} | \psi \rangle \right|^2.$$

This postulate is the reason that we require states be normalised because this means that

$$\sum_k P_k = \sum_k \sum_{n=1}^{g_k} \left| \langle u_k^{(n)} | \psi \rangle \right|^2 = 1,$$

where g_k is the degeneracy of the eigenvalue O_k. In the case of non-degenerate eigenvalues the second sum has only one term and we recover the more familiar requirement that

$$\sum_k |\langle \psi_k | \psi \rangle|^2 = 1.$$

5   Collapse of the State Vector

Suppose we have a state $|\psi\rangle$ and we measure the observable O and get a value of O_k which has the associated eigenvector $|\psi_k\rangle$. Immediately after performing this measurement we know that the value of O is O_k with probability 1. This must mean that the system is in state $|\psi_k\rangle$. The act of measuring changes the state of the system. This is true for all measurements we could make.

Postulate 5

Immediately after a measurement that gives the result O_k, where O_k is a non-degenerate eigen-value of $\hat{O}$, the state of the system is $|\psi_k\rangle$, which is the eigenvector associated with O_k. The state vector has been projected onto the eigenstate by the process of performing the measurement.

Mathematically what this looks like is we start with a state, $|\psi\rangle$, and perform some sort of measurement. The outcome of the measurement is O_k with probability $|\langle \psi_k | \psi \rangle|^2$. After the experiment the state is

$$\frac{1}{\sqrt{\langle \psi | \hat{\mathcal{P}}_k | \psi \rangle}} \hat{\mathcal{P}}_k |\psi\rangle.$$

Here $\hat{\mathcal{P}}_k = |\psi_k\rangle\langle\psi_k|$ is the projection operator onto the eigenspace of $|\psi_k\rangle$.

If instead O_k has degeneracy g_k then

$$\hat{\mathcal{P}}_k = \sum_{n=1}^{g_k} |u_k^n\rangle\langle u_k^{(n)}|$$

where $|u_k^{(n)}\rangle$ are the eigenvectors with eigenvalue O_k for $n = 1, \ldots, g_k$. The state after measurement is then

$$\frac{1}{\sqrt{\sum_{n=1}^{g_k} \left| c_k^{(n)} \right|^2}} \sum_{n=1}^{g_k} c_k^{(n)} |u_k^{(n)}\rangle.$$

5.1 Compatible Observables

Suppose A and B are observables. If we take three measurements, first measuring A, then B, then A again, A and B are said to be **compatible** if and only if the result of the third measurement is equal to the result of the first measurement with probability one. That is measuring B doesn't change the outcome of measuring A.

Suppose that A and B have associated operators $\hat{A}$ and $\hat{B}$ respectively, which have no degenerate eigenvalues. Suppose also that

$$\hat{A}|u_i\rangle = A_i|u_i\rangle, \qquad \text{and} \qquad \hat{B}|v_i\rangle = B_i|v_i\rangle.$$

After we measure A if we get a value of A_j then the state of the system is $|u_j\rangle$. Next we measure B and we get B_k meaning that the system is in the state $|v_k\rangle$. The only way that we can guarantee another measurement of A will yield the same value as the first is if $|v_k\rangle = |u_j\rangle$. For this to hold for all possible values of the measurement we must have that $|v_k\rangle = |v_j\rangle$ for all possible values of (i, j). If there is no degeneracy then we have a one to one correspondence between eigenvectors of $\hat{A}$ and eigenvectors of $\hat{B}$. We say that $\hat{A}$ and $\hat{B}$ have a common eigenbasis.

The conditions under which A and B are compatible are summarised in the compatibility theorem:

Theorem 4: The Compatibility Theorem

Given two observables, A and B, with associated operators $\hat{A}$ and $\hat{B}$, then the following are equivalent:
  1. A and B are compatible observables,
  2. $\hat{A}$ and $\hat{B}$ have a common eigenbasis,
  3. The operators $\hat{A}$ and $\hat{B}$ commute: $[\hat{A}, \hat{B}] = 0$.

Due to the third condition compatible observables are sometimes called **commuting observables**.

5.2 Complete Sets of Compatible observables

Consider an observable A with associated operator $\hat{A}$. Let the eigenbasis of $\hat{A}$ be $\{u_i\}$. In the case of no degeneracy we have

$$\hat{A}|u_k\rangle = a_k|u_k\rangle.$$

We can identify the eigenvector $|u_k\rangle$ by its eigenvalue, which we do by writing it $|a_k\rangle$. A is then trivially a complete set of commuting/compatible observables (CSCO) by itself.

In the case where there is degeneracy we can introduce another observable, B, which is compatible with A. It is possible to find a common eigenbasis of $\hat{A}$ and $\hat{B}$. Call this eigenbasis $\{|v_k^{(n)}\rangle\}$ where $n = 1, \ldots, g_k$. Suppose that

$$\hat{A}|v_k^{(n)}\rangle = a_i|v_k^{(n)}\rangle, \qquad \text{and} \qquad \hat{B}|v_k^{(n)}\rangle = b_j|v_k^{(n)}\rangle.$$

If (a_i, b_j) uniquely identifies this eigenstate then A and B are a CSCO. We then write $|v_k^{(n)}\rangle = |a_i, b_j\rangle$.

If (a_i, b_j) does not uniquely identify $|v_k^{(n)}\rangle$ then we simply keep introducing observables that are compatible with all previously introduced observables until we have a tuple of eigenvalues that does uniquely identify each eigenvector.

That is $\{A, B, C, \ldots\}$ is a **complete set of commuting/compatible observables (CSCO)** if and only if:

  - All observables commute in pairs,

- Specifying all the eigenvalues of the operators identifies a unique eigenvector in the common eigenbasis.

Given a CSCO, for example, $\{A, B, C\}$, we can expand a state in the common eigenbasis as

$$|\psi\rangle = \sum_{n,p,q} c_{npq}|a_n, b_p, c_q\rangle.$$

The probability of measuring a_n, b_p, and c_q simultaneously is then $|c_{npq}|^2$.

6   Continuous Spectra

Up to now we have only considered eigenvalues and eigenvectors that can be identified by a discrete variable, k. We can also have eigenvalues that are represented by a continuous variable, f. For example

$$\hat{f}|f\rangle = f|f\rangle$$

where we follow the convention of identifying an eigenvector by the associated eigenvalue. Assuming that $\hat{f}$ is hermitian then $f \in \mathbb{R}$. The completeness relation for a continuous basis set is

$$\int \mathrm{d}f\, |f\rangle\langle f| = \mathbb{1}.$$

We can expand a generic state, $|\psi\rangle$, in this basis:

$$|\psi\rangle = \int \mathrm{d}f\, c(f)|f\rangle.$$

Here $c(f)$ plays the role of the coordinates of $|\psi\rangle$ in the same way that c_k did in the discrete case. The method for computing $c(f)$ is the same as in the discrete case:

$$
\begin{aligned}
c(f) &= \langle f|\psi\rangle \\
&= \int \mathrm{d}x\, \langle f|x\rangle\langle x|\psi\rangle \\
&= \int \mathrm{d}x\, f^*(x)\psi(x).
\end{aligned}
$$

Since we now have a continuous variable the probability of any one particular value has to be zero. However $|c(f)|^2$ does give the probability density such that

$$P(f \in [f_-, f_+]) = \int_{f_-}^{f_+} \mathrm{d}f\, |c(f)|^2.$$

To be properly normalised we demand that

$$\int_{-\infty}^{\infty} \mathrm{d}f\, |c(f)|^2 = 1.$$

Perhaps the most common continuous observable to consider is the position. The operator for position is $\hat{X}$ and its action on $|\psi\rangle$ is

$$\hat{X}|\psi\rangle = x|\psi\rangle.$$

Note that this is *not* an eigenvalue equation, x is a variable not a constant and this holds for all $|\psi\rangle$. This is simply a statement that the action of $\hat{X}$ is to multiply the state by x. The effect that this has on the wave function is as we would expect:

$$
\begin{aligned}
\hat{X}|\psi\rangle &= \hat{X}\left[\int \mathrm{d}x\, \psi(x)|x\rangle\right] \\
&= \int \mathrm{d}x\, \psi(x)\hat{X}|x\rangle
\end{aligned}
$$

$$= \int \mathrm{d}x\, \psi(x) x |x\rangle$$

$$= \int \mathrm{d}x\, \varphi(x) |x\rangle$$

$$= |\varphi\rangle$$

where $|\varphi\rangle$ is defined to be such that $\langle x|\varphi\rangle = \varphi(x) = x\psi(x)$. Being slightly lazy with notation we will sometimes write things like

$$\hat{X}\psi(x) = x\psi(x)$$

which is not, strictly speaking, defined as operators can only act on elements of $\mathcal{H}$, not on the coordinates of these vectors. However the meaning should be clear enough that it doesn't matter.

6.1   Orthonormality

Suppose that f and f' distinct eigenvalues of $\hat{f}$. Then by definition

$$\hat{f}|f\rangle = f|f\rangle, \qquad \text{and} \qquad \hat{f}|f'\rangle = f'|f'\rangle.$$

We can compute matrix elements of $\hat{f}$:

$$\langle f'|\hat{f}|f\rangle = f\langle f'|f\rangle,$$

or using the hermitian property of $\hat{f}$:

$$\begin{aligned}
\langle f'|\hat{f}|f\rangle &= (\langle f|\hat{f}^\dagger|f'\rangle)^* \\
&= (\langle f|\hat{f}|f'\rangle)^* \\
&= (f'\langle f|f'\rangle)^* \\
&= f'^*(\langle f|f'\rangle)^* \\
&= f'\langle f'|f\rangle.
\end{aligned}$$

Hence

$$(f - f')\langle f'|f\rangle = 0.$$

Assuming that $f \neq f'$ then $f - f' \neq 0$ so $\langle f'|f\rangle = 0$.

Consider the state $|\psi\rangle$ expanded in this basis:

$$|\psi\rangle = \int \mathrm{d}f\, c(f)|f\rangle.$$

We can compute $\|\psi\|$:

$$\begin{aligned}
\|\psi\|^2 &= \langle \psi|\psi\rangle \\
&= \int \mathrm{d}f'\, f^*(f')\langle f'| \int \mathrm{d}f\, c(f)|f\rangle \\
&= \int \mathrm{d}f'\, \mathrm{d}f\, c^*(f)c(f)\langle f'|f\rangle \\
&= \int \mathrm{d}f\, c^*(f)c(f)\langle f|f\rangle \\
&= \int \mathrm{d}f\, |c(f)|^2 \langle f|f\rangle
\end{aligned}$$

Here we have used that $\langle f'|f\rangle$ is zero apart from when $f = f'$ to replace f' with f while removing one of the integrals. We also require that $\|\psi\| = 1$ and we already have that

$$\int |c(f)|^2 = 1$$

so we must have

$$\int \mathrm{d}f\, \langle f|f\rangle = 1.$$

The properties that we have described so far for $\langle f'|f\rangle$ are exactly the definition of the Dirac delta function:

$$\langle f'|f\rangle = \delta(f - f').$$

Example 6.1. Consider a system on a line. The position operator, $\hat{X}$, is defined as

$$\hat{X}|x\rangle = x|x\rangle.$$

A generic state is

$$|\psi\rangle = \int \mathrm{d}x\, \psi(x)|x\rangle$$

where

$$\psi(x) = \langle x|\psi\rangle.$$

By definition

$$\langle x|x'\rangle = \delta(x - x').$$

Thus

$$\begin{aligned}
\langle \psi|\psi\rangle &= \int \mathrm{d}x\, \mathrm{d}x'\, \psi^*(x)\psi(x')\langle x|x'\rangle \\
&= \int \mathrm{d}x\, \mathrm{d}x'\, \psi^*(x)\psi(x')\delta(x - x') \\
&= \int \mathrm{d}x\, \psi^*(x)\psi(x) \\
&= \int \mathrm{d}x\, |\psi(x)|^2.
\end{aligned}$$

Consider now the state that has a wave function given by

$$\psi(x) = c \exp\left[\frac{i}{\hbar}p_0 x - \frac{(x - x_0)^2}{2\xi^2}\right]$$

where x_0, p_0, and ξ are constants and $c \in \mathbb{R}$ is a normalisation factor.

$$\begin{aligned}
\langle \psi|\psi\rangle &= \int_{-\infty}^{\infty} |\psi(x)|^2 \\
&= \int_{-\infty}^{\infty} c^2 \exp\left[-\frac{(x - x_0)^2}{2\xi^2}\right] \\
&= c^2 \xi\sqrt{\pi}.
\end{aligned}$$

Requiring a normalised state we then have

$$c = \frac{1}{\pi^{1/4}\sqrt{\xi}}.$$

The expected position is given by

$$\begin{aligned}
\langle x\rangle &= \int \mathrm{d}x\, x|\psi(x)|^2 \\
&= \frac{1}{\xi\sqrt{\pi}} \int \mathrm{d}x\, x \exp\left[-\frac{(x - x_0)^2}{2\xi^2}\right] \\
&= x_0.
\end{aligned}$$

Similarly it can be shown that

$$\langle \Delta x^2\rangle = \langle (x - x_0)^2\rangle = \xi^2.$$

This can be shown by computing the integrals or recognising the integrands as Gaussians and hence we are simply computing the moments of this distribution.

6.2   Momentum Operator

The momentum operator, $\hat{P}\colon \mathcal{H} \to \mathcal{H}$, is defined by

$$\langle x|\hat{P}|\psi\rangle = \hat{P}\psi(x) = -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\psi(x).$$

We saw in example 4.2 that this is a hermitian operator. One justification for calling this a momentum operator comes from a plane wave with momentum p:

$$\psi_p(x) = \exp\left[\frac{ipx}{\hbar}\right].$$

When we act on this with the momentum operator we get

$$\begin{aligned}\hat{P}\psi_p(x) &= -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\exp\left[\frac{ipx}{\hbar}\right] \\ &= -i\hbar\frac{ip}{\hbar}\exp\left[\frac{ipx}{\hbar}\right] \\ &= p\psi_p(x)\end{aligned}$$

So, at least in this case, the momentum is an eigenvalue of $\hat{P}$.

Momentum is a conserved variable. By Noether's every conserved variable has a related transformation. The transformation related to momentum is translations. We can describe a small translation about the coordinate x by an amount a as

$$\psi(x+a) = \psi(x) + a\frac{\mathrm{d}}{\mathrm{d}x}\psi(x) + \mathcal{O}a^2.$$

However $\mathrm{d}/\mathrm{d}x$ is an anti-hermitian operator. Fortunately there is an easy fix for this, if A is anti-hermitian then iA is hermitian. This means we can write

$$\psi(x+a) = \psi(x) + \frac{i}{\hbar}\left(-i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\right)\psi(x) + \mathcal{O}a^2.$$

The factor of $\hbar$ is convention and ensures that we have the correct units. In this way the momentum operator is the generator of translations.

6.3   The Canonical Commutation Relation

Consider the two operators that we have met so far that are related to actual observables, the position, $\hat{X}$, and momentum, $\hat{P}$. The commutator of these two operators is fairly easy to compute:

$$\begin{aligned}\hat{X}\hat{P}\psi(x) &= \hat{X}\left[-i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\psi(x)\right] \\ &= -i\hbar x\frac{\mathrm{d}}{\mathrm{d}x}\psi(x) \\ \hat{P}\hat{X}\psi &= \hat{P}x\psi(x) \\ &= -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}[x\psi(x)] \\ &= -i\hbar x\frac{\mathrm{d}}{\mathrm{d}x}\psi(x) - i\hbar\psi(x) \\ [\hat{X},\hat{P}]\psi &= \hat{X}\hat{P} - \hat{P}\hat{X} \\ &= -i\hbar x\frac{\mathrm{d}}{\mathrm{d}x}\psi(x) + i\hbar x\frac{\mathrm{d}}{\mathrm{d}x}\psi(x) + i\hbar\psi(x) \\ &= i\hbar\psi(x)\end{aligned}$$

This holds for all ψ therefore we conclude that

$$\boxed{[\hat{X},\hat{P}] = i\hbar.}$$

This is called the **canonical commutation relation**. From this we can derive the **Heisenberg's uncertainty principle (HUP)**:

$$\Delta x \Delta p \geq \frac{\hbar}{2}.$$

Part III

Dynamics

7 Dynamics

7.1 Dynamics in Classical Mechanics

In classical mechanics the dynamics of a system can be entirely determined from Newton's second law:

$$\boldsymbol{F} = m\boldsymbol{a} = m\frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{d}t^2}.$$

The force term encodes information about the external forces applied to the system, the mass is a characteristic of the system and the acceleration provides a time dependence. Solving this differential equation can give us the equations of motion. From these we can predict the future of the system but also the past.

7.2 Dynamics in Quantum Mechanics

> **Postulate 6**
>
> The dynamics of a system described by the state $|\Psi(t)\rangle$, at time t, is determined by the **Schrödinger equation**:
>
> $$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle.$$
>
> Where $\hat{H}$ is the **Hamiltonian**, a hermitian operator corresponding to the energy of the system.

The state vectors, $|\Psi(t)\rangle$, now include an explicit time dependence.

The Hamiltonian, for a single particle in one dimension, is given in an analogous way to classical mechanics by

$$\hat{H} = \hat{T} + \hat{V}.$$

Here $\hat{T}$ is the operator associated with the kinetic energy and $\hat{V}$ is the operator associated with the potential energy. In general it is often possible to take a classical formula and change x to $\hat{X}$, and p to $\hat{P}$ and have it still be valid, although we do have to be careful about whether operators commute which we don't have to consider with classical mechanics. Fortunately this is valid for this case and we have

$$T = \frac{p^2}{2m} \to \hat{T} = \frac{\hat{P}^2}{2m},$$
$$V = V(x) \to \hat{V} = V(\hat{X}).$$

Here the function V on the left, which acts on real numbers, and the function V on the right, which acts on operators, are technically different as their domains are different but they have the same form. For example in a harmonic potential,

$$V(x) = \frac{1}{2}kx^2$$

and the operator equivalent is

$$V(\hat{X}) = \frac{1}{2}k\hat{X}^2.$$

Consider a state $|\Psi(t)\rangle$, after a small amount of time, ε, this becomes

$$|\Psi(t+\varepsilon)\rangle = |\Psi(t)\rangle + \varepsilon\frac{\mathrm{d}}{\mathrm{d}t}|\Psi(t)\rangle$$
$$= |\Psi(t)\rangle - \frac{i\varepsilon}{\hbar}\hat{H}|\Psi(t)\rangle$$

In this way we can view the change in time as the result of the action of $\hat{H}$.

7.3 Time Evolution of a Wave Function

The wave function related to the state $|\Psi(t)\rangle$ is, as normal, found by the inner product with the position eigenbasis vectors:

$$\Psi(x,t) = \langle x|\Psi(t)\rangle.$$

If we take the inner product of the left hand side of the Schrödinger equation with one of these basis vectors we get

$$\langle x|i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\Psi(t)\rangle = i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\langle x|\Psi(t)\rangle = i\hbar\frac{\partial}{\partial t}\Psi(x,t).$$

Here we have used the fact that $\langle x|$ is independent of time so commutes with a time derivative. Taking the inner product with the right hand side we have

$$\langle x|\hat{H}|\Psi(t)\rangle = \hat{H}\langle x|\Psi(t)\rangle = \hat{H}\Psi(x,t).$$

In the case that $\hat{H} = \hat{T} + \hat{V}$,

$$\hat{T} = \frac{\hat{P}^2}{2m}, \qquad \text{and} \qquad \hat{V} = V(\hat{X})$$

the Schrödinger equation becomes

$$\hat{H}\Psi(x,t) = \left[\frac{\hat{P}^2}{2m} + V(\hat{X})\right]\Psi(x,t) = i\hbar\frac{\partial}{\partial t}\Psi(x,t).$$

We can calculate the action of $\hat{P}^2\Psi(x,t)$ easily:

$$\hat{P}^2\Psi(x,t) = -i\hbar\frac{\partial}{\partial x}\left(-i\frac{\partial}{\partial x}\hbar\right) = -\hbar^2\frac{\partial^2}{\partial x^2}\Psi(x,t).$$

Since the action of $\hat{X}$ is simply multiplication by x we have that $V(\hat{X}) = V(x)$. Schrödinger's equation becomes

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right]\Psi(x,t) = i\hbar\frac{\partial}{\partial t}\Psi(x,t).$$

Notice that this has a second order derivative with respect to position but only a first order derivative with respect to time. This means that time and position are treated differently and therefore this formula fails to be relativistic.

7.4 Justification of the Schrödinger Equation

While the Schrödinger equation is a postulate so doesn't need justifying it can't hurt to take a brief look at the logic that one could apply in coming up with it in the first case. A plane wave has the wave function

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

This describes the propagation of monochromatic light. Applying $i\hbar\partial_t$ to this we get

$$i\hbar\frac{\partial}{\partial t}\Psi(x,t) = i\hbar\frac{\partial}{\partial t}Ae^{-i(\omega t - kx)}.$$
$$= i\hbar(-i\omega t)Ae^{-i(\omega t - kx)}.$$
$$= \hbar\omega\Psi(x,t)$$
$$= 2\pi\nu\hbar\Psi(x,t)$$

where $\nu = \omega/2\pi$ is the frequency of the wave. The eigenvalue, $2\pi\nu\hbar$, is precisely the energy of the photon which justifies the use of the Hamiltonian as the operator associated with the total energy.

As well as this we also require $\|\Psi(t)\| = 1$ for all times, t:

$$\|\Psi(t)\|^2 = \int \mathrm{d}x \, \Psi^*(x,t)\Psi(x,t) = 1.$$

Since this integral is constant with respect to time the time derivative of this integral must vanish:

$$
\begin{aligned}
0 &= \frac{\mathrm{d}}{\mathrm{d}t} \int \mathrm{d}x \, \Psi^*(x,t)\Psi(x,t) \\
&= \int \mathrm{d}x \, \frac{\partial}{\partial t}[\Psi^*(x,t)\Psi(x,t)] \\
&= \int \mathrm{d}x \left[\left(\frac{\partial}{\partial t}\Psi^*(x,t)\right)\Psi(x,t) + \Psi^*(x,t)\left(\frac{\partial}{\partial t}\Psi(x,t)\right) \right]
\end{aligned}
$$

This occurs if ∂_t is represented by an anti-hermitian operator. This is why there is a factor of i, since if A is anti-hermitian then iA is hermitian so $i\partial_t$ is represented by a hermitian operator. The factor of $\hbar$ is required for the dimensions to match correctly.

7.5   Eigenstates of the Hamiltonian

The Hamiltonian is the operator associated with the energy. Assuming that the spectrum of $\hat{H}$ is discrete the eigenvalue equation for $\hat{H}$ is

$$\boxed{\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle.}$$

This is called the **time independent Schrödinger equation (TISE)**. $E_n \in \mathbb{R}$ are the possible outcomes of measuring the energy of the system and $|\psi_n\rangle$ is a state in which, when measuring the energy, we get E_n with probability 1.

7.5.1   Time Evolution of the Energy Eigenstates

Suppose we have a system that starts in one of the energy eigenstates. That is

$$|\Psi(0)\rangle = |\psi_n\rangle.$$

We make the ansatz that the time evolution of the system, that is the state of the system after some time, t, is given by

$$|\Psi(t)\rangle = e^{-iE_n t/\hbar}|\Psi(0)\rangle = e^{-iE_n t/\hbar}|\psi_n\rangle.$$

We can easily check that this satisfies the Schrödinger equation:

$$
\begin{aligned}
i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\Psi(t)\rangle &= i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-iE_n t/\hbar}|\psi_n\rangle\right) \\
&= i\hbar|\psi_n\rangle\left(\frac{\mathrm{d}}{\mathrm{d}t}e^{-iE_n t/\hbar}\right) \\
&= i\hbar|\psi_n\rangle\left(-\frac{i}{\hbar}E_n e^{-iE_n t/\hbar}\right) \\
&= E_n e^{-iE_n t/\hbar}|\psi_n\rangle \\
&= e^{-iE_n t/\hbar}\hat{H}|\psi_n\rangle \\
&= \hat{H}|\Psi(t)\rangle.
\end{aligned}
$$

Here we have used the fact that the state, $|\psi_n\rangle$, is independent of time so commutes with $\mathrm{d}/\mathrm{d}t$. Also the phase factor, $\exp(-iE_n t/\hbar) \in \mathbb{C}$, commutes with the operator $\hat{H}$.

This also trivially satisfies the boundary conditions:

$$|\Psi(0)\rangle = e^{-iE_n 0/\hbar}|\psi_n\rangle = |\psi_n\rangle.$$

Hence we conclude that the ansatz is correct.

7.5.2   Time Evolution of Matrix Elements

Let $\hat{O}$ be a generic operator that is constant with respect to time. Define the matrix element

$$O_n(t) = \langle \Psi_n(t)|\hat{O}|\Psi_n(t)\rangle$$

where

$$|\Psi_n(t)\rangle = e^{-iE_nt/\hbar}|\psi_n\rangle.$$

This gives us

$$\langle \Psi_n(t)| = e^{iE_nt/\hbar}\langle \psi_n|.$$

Notice that $O_n(t)$ may have time dependence as the states have time dependence even if the operator doesn't. However it actually turns out that $O_n(t)$ *doesn't* have time dependence:

$$\begin{aligned}
O_n(t) &= \langle \Psi_n(t)|\hat{O}|\Psi_n(t)\rangle \\
&= \langle \psi_n|e^{iE_nt/\hbar}\hat{O}e^{-iE_nt/\hbar}|\psi_n\rangle \\
&= \langle \psi_n|\hat{O}|\psi_n\rangle \\
&= O_n(0).
\end{aligned}$$

Hence $O_n(t)$ is independent of t. Here we have used that the phase factors are simply complex numbers so commute with the operators.

We can conclude that when the system is prepared in an eigenstate of the Hamiltonian all expected values of operators are constant. For this reason we call $|\psi_n\rangle$ **stationary states**.

Example 7.1. Consider the operator $\hat{\mathcal{P}}_x = |x\rangle\langle x|$. The matrix element of this is

$$\langle \Psi(t)|\hat{\mathcal{P}}_x|\Psi(t)\rangle = \langle \Psi(t)|x\rangle\langle x|\Psi(t)\rangle = |\Psi(x,t)|^2.$$

This must be independent of time if $|\Psi(0)\rangle = |\psi_n\rangle$. As usual $|\Psi(x,t)|^2$ gives the probability of finding the system at x. Since this is independent of time the expected position of x is constant. This justifies the name 'stationary state'.

7.6   Time Evolution of a Generic State

We define a **time evolution operator**, $\hat{U}(t)$, that acts on a state at time $t = 0$ to give the value of that state at time t:

$$\hat{U}(t)|\Psi(0)\rangle = |\Psi(t)\rangle.$$

The form of this operator is

$$\hat{U}(t) = \exp\left(-\frac{i}{\hbar}\hat{H}t\right).$$

Notice the similarity to the time evolution of a single state. All that has happened is the eigenvalue has been replaced by the operator. As usual with operators the exponential is defined through its Taylor series:

$$\exp\left(-\frac{i}{\hbar}\hat{H}t\right) = \sum_{k=0}^{\infty}\frac{1}{k!}\left(-\frac{i}{\hbar}\right)^k\hat{H}^kt^k$$

We can show that

$$\hat{U}(t)|\Psi(0)\rangle$$

is a solution to the Schrödinger equation:

$$\begin{aligned}
i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\hat{U}|\Psi(0)\rangle &= i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\exp\left(-\frac{i}{\hbar}\hat{H}t\right)|\Psi(0)\rangle \\
&= i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\sum_{k=0}^{\infty}\frac{1}{k!}\left(-\frac{i}{\hbar}\right)^k\hat{H}^kt^k|\Psi(0)\rangle
\end{aligned}$$

Note that only the terms t^k have time dependnece:

$$= i\hbar \sum_{k=0}^{\infty} \frac{1}{k!} \left(-\frac{i}{\hbar}\right)^k \hat{H}^k |\Psi(0)\rangle \frac{\mathrm{d}}{\mathrm{d}t} t^k$$

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

The first term in the sum is zero so we can disregard it

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

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

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

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

Let $k' = k - 1$:

$$= \hat{H} \sum_{k'=0}^{\infty} \frac{1}{k'!} \left(-\frac{i}{\hbar}\right)^{k'} \hat{H}^{k'} |\Psi(0)\rangle t^{k'}$$

$$= \hat{H} \exp\left(-\frac{i}{\hbar}\hat{H}t\right) |\Psi(0)\rangle.$$

$$= \hat{H}\hat{U}(t)|\Psi(0)\rangle.$$

Hence $\hat{U}(t)|\Psi(0)\rangle$ is a solution to the Schrödinger equation with the boundary condition that $U(0)|\Psi(0)\rangle = |\Psi(0)\rangle$. Note that in showing this we assumed nothing about the derivative of $e^{\hat{A}t}$. We have shown that derivative is the same as if $\hat{A}$ was a constant real number, that is

$$\frac{\mathrm{d}}{\mathrm{d}t} e^{\hat{A}t} = \hat{A}e^{\hat{A}t}.$$

From now on we will assume this.

We can use this to recover the time evolution of the eigenbasis of the Hamiltonian:

$$\hat{U}(t)|\psi_n\rangle = \exp\left(-\frac{i}{\hbar}\hat{H}t\right)|\psi_n\rangle$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} \left(-\frac{i}{\hbar}\right)^k t^k \hat{H}^k |\psi_n\rangle$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} \left(-\frac{i}{\hbar}\right)^k t^k E_n^k |\psi_n\rangle$$

$$= \exp\left(-\frac{i}{\hbar}E_n t\right)|\psi_n\rangle$$

7.7   The Eigenbasis of the Hamiltonian

For any mildly complex system the form of the Hamiltonian can make actually calculating the time evolution operator infeasible. For this reason it is common to work in the eigenbasis of the Hamiltonian as the time evolution of these states is known. A generic state, $|\Psi(t)\rangle$, can be expanded as

$$|\Psi(t)\rangle = \sum_n c_n(t)|\psi_n\rangle.$$

Where
$$c_n(t) = \langle \psi_n | \Psi(t) \rangle$$

Only the coordinates have time dependence in this basis, not the basis vectors. In terms of the time evolution operator this is

$$\begin{aligned}
|\Psi(t)\rangle &= \hat{U}(t)|\Psi(0)\rangle \\
&= \hat{U}(t)\sum_n c_n(0)|\psi_n\rangle \\
&= \sum_n c_n(0)\hat{U}(t)|\psi_n\rangle \\
&= \sum_n c_n(0)e^{-iE_n t/\hbar}|\psi_n\rangle.
\end{aligned}$$

Be careful here, the time evolution of each basis state is different as E_n is potentially different for each n. Since only the coordinates have time dependence we can see that the time evolution of the coordinates is given by

$$e^{iE_n t/\hbar}c_n(0).$$

7.8   Time Evolution of Expected Values

Suppose that $\hat{O}$ is a hermitian operator that is time independent. The expected value upon measuring O at time t for a state, $|\Psi(t)\rangle$, is

$$O(t) = \langle \Psi(t)|\hat{O}|\Psi(t)\rangle.$$

To see how this varies with time we take the derivative:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}O(t) &= \frac{\mathrm{d}}{\mathrm{d}t}[\langle\Psi(t)|\hat{O}|\Psi(t)\rangle] \\
&= \left(\frac{\mathrm{d}}{\mathrm{d}t}\langle\Psi(t)|\right)\hat{O}|\Psi(t)\rangle + \langle\Psi(t)|\hat{O}\left(\frac{\mathrm{d}}{\mathrm{d}t}|\Psi(t)\rangle\right)
\end{aligned}$$

The action of $\mathrm{d}/\mathrm{d}t$ on $|\Psi(t)\rangle$ is given by Schrödinger's equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}|\Psi(t)\rangle = -\frac{i}{\hbar}\hat{H}|\Psi(t)\rangle.$$

We need to find the action of $\mathrm{d}/\mathrm{d}t$ on $\langle\Psi(t)|$. First we define

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\Psi(t)| = \langle\Phi(t)|.$$

Then using the definition of the hermitian conjugate we have

$$\begin{aligned}
\left(\frac{\mathrm{d}}{\mathrm{d}t}\langle\Psi(t)|\right)\hat{O}|\Psi(t)\rangle &= \langle\Phi(t)|\hat{O}|\Psi(t)\rangle \\
&= (\langle\Psi(t)|\hat{O}^\dagger|\Phi(t)\rangle)^*
\end{aligned}$$

noting that $\hat{O}$ is hermitian we have

$$\begin{aligned}
&= (\langle\Psi(t)|\hat{O}|\Phi(t)\rangle)^* \\
&= \left(\langle\Psi(t)|\hat{O}\frac{\mathrm{d}}{\mathrm{d}t}|\Psi(t)\rangle\right)^* \\
&= \left(\langle\Psi(t)|\hat{O}\left[-\frac{i}{\hbar}\right]\hat{H}|\Psi(t)\rangle\right)^* \\
&= \frac{i}{\hbar}(\langle\Psi(t)|\hat{O}\hat{H}|\Psi(t)\rangle)^* \\
&= \frac{i}{\hbar}\langle\Psi(t)|(\hat{O}\hat{H})^\dagger|\Psi(t)\rangle \\
&= \frac{i}{\hbar}\langle\Psi(t)|\hat{H}^\dagger\hat{O}^\dagger|\Psi(t)\rangle
\end{aligned}$$

$$= \frac{i}{\hbar} \langle \Psi(t) | \hat{H} \hat{O} | \Psi(t) \rangle$$

Hence

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} O(t) &= \frac{i}{\hbar} \langle \Psi(t) | \hat{H} \hat{O} | \Psi(t) \rangle - \frac{i}{\hbar} \langle \Psi(t) | \hat{O} \hat{H} | \Psi(t) \rangle \\
&= \frac{i}{\hbar} \langle \Psi(t) | (\hat{H} \hat{O} - \hat{O} \hat{H}) | \Psi(t) \rangle \\
&= \frac{i}{\hbar} \langle \Psi(t) | [\hat{H}, \hat{O}] | \Psi(t) \rangle \\
&= \frac{i}{\hbar} \langle [\hat{H}, \hat{O}] \rangle.
\end{aligned}
\tag{7.1}
$$

We see that if $\hat{O}$ and $\hat{H}$ commute then the expected value of O is constant with respect to time.

Note that if we allowed the operator $\hat{O}$ to have time dependence then after applying the product rule we would have had an extra term. Dealing with this extra term is easier if we work with wave functions:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} O(t) &= \frac{\mathrm{d}}{\mathrm{d}t} [\langle \Psi(t) | \hat{O} | \Psi(t) \rangle] \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \int \mathrm{d}x \, \langle \Psi(t) | x \rangle \langle x | \hat{O} | \Psi(t) \rangle \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \int \mathrm{d}x \, \langle \Psi(t) | x \rangle \hat{O} \langle x | \Psi(t) \rangle \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \int \mathrm{d}x \, \Psi^*(x, t) \hat{O} \Psi(x, t) \\
&= \int \mathrm{d}x \, \frac{\partial}{\partial t} [\Psi^*(x, t) \hat{O} \Psi(x, t)] \\
&= \int \mathrm{d}x \left(\frac{\partial}{\partial t} \Psi^*(x, t) \right) \hat{O} \Psi(x, t) \\
&\quad + \int \mathrm{d}x \, \Psi^*(x, t) \left(\frac{\partial}{\partial t} \hat{O} \right) \Psi(x, t) \\
&\quad + \int \mathrm{d}x \, \Psi^*(x, t) \hat{O} \left(\frac{\partial}{\partial t} \Psi(x, t) \right)
\end{aligned}
$$

using the work above for the derivatives of the wave functions this becomes

$$= \frac{i}{\hbar} \langle [\hat{H}, \hat{O}] \rangle + \int \mathrm{d}x \, \Psi^*(x, t) \frac{\partial \hat{O}}{\partial t} \Psi(x, t)$$

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle O \rangle = \frac{i}{\hbar} \langle [\hat{H}, \hat{O}] \rangle + \left\langle \frac{\partial \hat{O}}{\partial t} \right\rangle.$$

In the case that $\hat{O}$ is time independent then the last term is zero and this reduces to the case above.

8   One-Dimensional Systems

8.1   Recap

The state of the system is described by a vector, $|\psi\rangle \in \mathcal{H}$. We work in the position basis, $\{|x\rangle\}$, which are the eigenvectors of the position operator, $\hat{X}$:

$$\hat{X} |x\rangle = x |x\rangle.$$

This is a complete set so

$$\int \mathrm{d}x \, |x\rangle \langle x| = \mathbb{1}.$$

We can use this to expand $|\psi\rangle$ in this basis:

$$|\psi\rangle = \int \mathrm{d}x \, \langle x | \langle x | \psi \rangle = \int \mathrm{d}x \, \psi(x) |x\rangle.$$

Where

$$\psi(x) = \langle x|\psi \rangle$$

is a complex valued function called the wave function.

8.2   Time Independent Schrödinger Equation

The TISE is the eigenvalue equation for the Hamiltonian, $\hat{H}$:

$$\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle.$$

We assume here that the spectrum of $\hat{H}$ is discrete, and hence labelled by an integer, n, but this needn't be the case. We can take the inner product of the TISE with the basis vectors and on the right hand side we get

$$\langle x|E_n|\psi\rangle = E_n\langle x|\psi\rangle = E_n\psi(x).$$

On the left hand side we have

$$\langle x|\hat{H}|\psi\rangle = \hat{H}\psi(x).$$

We want to know what the form of $\hat{H}$ is when it acts on a wave function. We take

$$\hat{H} = \hat{T} + \hat{V} = \frac{\hat{P}^2}{2m} + V(\hat{X})$$

where

$$\hat{P} = -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}, \quad\quad \text{and} \quad\quad \hat{X} = x.$$

Hence, expressed as a differential operator on a wave function, we have

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x).$$

Thus the TISE can be written as

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)\right]\psi_n(x) = E_n\psi(x).$$

This is the most general form of the TISE written as a differential equation. Generally our goal is to solve this for the eigenvalues, E_n, and eigenfunctions, ψ_n. Again, this still holds for a continuous spectrum.

8.3   Free Particle

For a free particle there is no potential, that is $V(x) = 0$. The TISE reduces to

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

Here we have dispensed with n subscripts as it turns out that the energy will have a continuous spectrum in this case. Rearranging this equation we need to solve

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi(x) = -\frac{2mE}{\hbar^2}\psi(x).$$

We define

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

We only need to consider the positive square root as the minimum potential is zero and therefore the minimum energy is zero so we don't need to consider $-k$. The TISE, in terms of k, becomes

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi(x) = -k^2\psi(x).$$

The solutions to this are known to be

$$\psi_{k,+}(x) = e^{ikx}$$

$$\psi_{k,-}(x) = e^{-ikx}$$

We can perform a brief dimensions check here.

$$[k]^2 = \frac{[\text{mass}][\text{energy}]}{[\text{energy}]^2[\text{time}]^2} = \frac{[\text{mass}]}{[\text{energy}][\text{time}]^2} = \frac{[\text{mass}]}{[\text{mass}][\text{length}]^2[\text{time}]^{-2}[\text{time}]^2} = [\text{length}]^{-2}$$

$$\implies [k] = [\text{length}]^{-1} \implies [kx] = [\text{length}]^{-1}[\text{length}] = 1.$$

This means that the argument of the exponential is dimensionless, as it must be. The energy eigenvalue is given by

$$E = \frac{\hbar^2 k^2}{2m}.$$

For each value of this eigenvalue there are two eigenfunctions, $\psi_{k,+}$ and $\psi_{k,-}$. This means that $\hat{H}$ is not, in this case, a CSCO.

We can introduce a different observable, which is compatible with $\hat{H}$, in a way such that we get a CSCO. It turns out that the momentum has these required properties. First we shall show that it is compatible with the energy by showing that it commutes with $\hat{H}$:

$$[\hat{P}, \hat{H}] = [\hat{P}, C\hat{P}^2] = 0$$

where C is a constant scalar and we have used that all operators commute with themselves. Hence $\hat{P}$ and $\hat{H}$ are compatible. Next we compute the eigenvalues of $\hat{P}$ with the eigenfunctions $\psi_{k,\pm}(x)$:

$$\hat{P}\psi_{k,\pm}(x) = -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}e^{\pm ikx} = -i\hbar(\pm ik)e^{\pm ikx} = \pm\hbar k\psi_{k,\pm}(x).$$

This means that we can distinguish between $\psi_{k,+}$ and $\psi_{k,-}$ by the momentum eigenvalue, $\hbar k$ and $-\hbar k$ respectively. Another dimension check here is a good idea:

$$[\hbar k] = [\text{energy}][\text{time}][\text{length}]^{-1} = [\text{momentum}][\text{length}][\text{length}]^{-1} = [\text{momentum}]$$

so the eigenvalues of the momentum have units of momentum as we would expect.

We can identify $\psi_{k,\pm}$ with

$$\psi_{k,\pm}(x) = \langle x|\psi_{k,\pm}\rangle = \langle x|E, \pm\hbar k\rangle.$$

Here $|E, \pm\hbar k\rangle \in \mathcal{H}$ is such that

$$\hat{H}|E, \pm\hbar k\rangle = E|E, \pm\hbar k\rangle, \qquad \text{and} \qquad \hat{P}|E, \pm\hbar k\rangle = \pm\hbar k|E, \pm\hbar k\rangle.$$

This uniquely identifies all eigenvectors and so we conclude that $\{\hat{H}, \hat{P}\}$ is a CSCO.

If we consider the two states $|E, \pm\hbar k\rangle$ we see that they have equal and opposite momentum. We can think of $|E, +\hbar k\rangle$ as propagating in the $+x$ direction with momentum $+\hbar k$, and $|E, -\hbar k\rangle$ as propagating in th $-x$ direction with momentum $-\hbar k$.

We could come up with an operator that swaps between $|E, +\hbar k\rangle$ and $|E, -\hbar k\rangle$. Since they only differ in direction we can change between these states by changing the parity of the system, we do this with the parity operator, $\hat{\mathscr{P}}$, defined as

$$\langle x|\hat{\mathscr{P}}|\psi\rangle = \hat{\mathscr{P}}\psi(x) = \psi(-x) = \langle -x|\psi\rangle.$$

We can check that this does indeed transform $|E, \pm\hbar k\rangle$ into $|E, \mp\hbar k\rangle$:

$$\hat{\mathscr{P}}\psi_{k,\pm}(x) = \psi_{k,\pm}(-x) = e^{\pm ik(-x)} = e^{\mp ikx} = \psi_{k,\mp}(x).$$

We can check that $\hat{\mathscr{P}}$ is hermitian:

$$\langle\varphi|\hat{\mathscr{P}}^\dagger|\psi\rangle = (\langle\psi|\hat{\mathscr{P}}|\varphi\rangle)^*$$

$$= \left(\int_{-\infty}^{\infty}\mathrm{d}x\,\langle\psi|x\rangle\langle x|\hat{\mathscr{P}}|\varphi\rangle\right)^*$$

$$= \left(\int_{-\infty}^{\infty} \mathrm{d}x \, \psi^*(x) \hat{\mathscr{P}} \varphi(x) \right)^*$$

$$= \left(\int_{-\infty}^{\infty} \mathrm{d}x \, \psi^*(x) \varphi(-x) \right)^*$$

Let $u = -x$ then $\mathrm{d}x \to -\mathrm{d}u$ and $(-\infty, \infty) \to (\infty, -\infty)$:

$$= \left(- \int_{\infty}^{-\infty} \mathrm{d}u \, \psi^*(-u) \varphi(u) \right)^*$$

$$= \left(\int_{-\infty}^{\infty} \mathrm{d}u \, \psi^*(-u) \varphi(u) \right)^*$$

$$= \int_{-\infty}^{\infty} \mathrm{d}u \, \varphi^*(u) \psi(-u)$$

$$= \int_{-\infty}^{\infty} \mathrm{d}u \, \varphi^*(u) \hat{\mathscr{P}} \psi(u)$$

$$= \int_{-\infty}^{\infty} \mathrm{d}u \, \langle \varphi | u \rangle \langle u | \hat{\mathscr{P}} | \psi \rangle$$

$$= \langle \varphi | \hat{\mathscr{P}} | \psi \rangle.$$

This holds for all states, $|\psi\rangle, |\varphi\rangle \in \mathcal{H}$, therefore we conclude that $\hat{\mathscr{P}}^\dagger = \hat{\mathscr{P}}$ so $\hat{\mathscr{P}}$ is hermitian.

8.4   Time Dependent Schrödinger Equation

The Schrödinger equation for a time dependent state, $|\Psi(t)\rangle \in \mathcal{H}$, is

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

Taking the inner product with the basis vectors on the left hand side we get

$$i\hbar \langle x | \frac{\mathrm{d}}{\mathrm{d}t} |\Psi(t)\rangle = i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \langle x |\Psi(t)\rangle = i\hbar \frac{\partial}{\partial t} \Psi(x,t).$$

Here we have used the fact that $|x\rangle$ is independent of time so commutes with $\mathrm{d}/\mathrm{d}t$. The right hand side gives

$$\langle x | \hat{H} |\Psi(t)\rangle = \hat{H} \Psi(x,t) = \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \Psi(x,t).$$

Hence we have the Schrödinger equation in the most general differential equation form:

$$\left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \Psi(x,t) = -i\hbar \frac{\partial}{\partial t} \Psi(x,t).$$

8.4.1   General Properties of a Wave Function

From this there are a few restriction that we can place on the form of a wave function:

1. $\Psi(x,t)$ must be a single valued function of both x and t.

2. For the first derivatives to exist $\Psi(x,t)$ must be continuous in x and t at all values of x and t.

3. For the second position derivative to exists $\partial_x \Psi(x,t)$ must be continuous in x at all values of x.

We can relax the last of these conditions if the potential, $V(x)$, has a singularity at x_0 then $\partial_x \Psi(x,t)$ can also have a singularity at x_0 if it occurs in such a way that the singularities cancel out in the entire equation.

8.5   Physical Properties

8.5.1   Minimum Energy

Suppose we have a state $|\psi\rangle \in \mathcal{H}$. Then we can calculate the expectation value of the energy as follows:

$$\langle \psi | \hat{H} | \psi \rangle = \int_{-\infty}^{\infty} \mathrm{d}x \, \langle \psi | x \rangle \langle x | \hat{H} | \psi \rangle$$

$$= \int_{-\infty}^{\infty} \mathrm{d}x \, \psi^*(x) \hat{H} |\psi\rangle$$

$$= \int_{-\infty}^{\infty} \mathrm{d}x \, \psi^*(x) \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \psi(x)$$

$$= -\int_{-\infty}^{\infty} \mathrm{d}x \, \frac{\hbar^2}{2m} \psi^*(x) \frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + \int_{-\infty}^{\infty} \mathrm{d}x \, \psi^*(x) V(x) \psi(x)$$

$$= -\int_{-\infty}^{\infty} \mathrm{d}x \, \frac{\hbar^2}{2m} \psi^*(x) \frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + \int \mathrm{d}x \, V(x) |\psi(x)|^2.$$

Focusing on the first integral we can integrate by parts:

$$u = \psi^*, \qquad\qquad\qquad v = \frac{\mathrm{d}\psi}{\mathrm{d}x},$$

$$u' = \frac{\mathrm{d}\psi^*}{\mathrm{d}x}, \qquad\qquad\qquad v' = \frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}.$$

Noting that $(\psi^*)' = (\psi')^*$ we have

$$\int \mathrm{d}x \, \frac{\hbar^2}{2m} \psi^*(x) \frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = \left[\psi^* \frac{\mathrm{d}\psi}{\mathrm{d}x} \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \mathrm{d}x \, \frac{\mathrm{d}\psi}{\mathrm{d}x}^* \frac{\mathrm{d}\psi}{\mathrm{d}x}$$

$$= -\int_{-\infty}^{\infty} \mathrm{d}x \left| \frac{\mathrm{d}\psi}{\mathrm{d}x} \right|^2$$

Hence the expectation value of the energy is

$$\langle \psi | \hat{H} | \psi \rangle = \int_{-\infty}^{\infty} \mathrm{d}x \left| \frac{\mathrm{d}\psi}{\mathrm{d}x} \right|^2 + \int_{-\infty}^{\infty} \mathrm{d}x \, V(x) |\psi(x)|^2$$

noting that the first integral is always non-negative we have

$$\geq \int_{-\infty}^{\infty} \mathrm{d}x \, V(x) |\psi(x)|^2$$

assuming that V is bounded below (which it must be for a physical potential) by $V_{\min}$ we get

$$\geq \int_{-\infty}^{\infty} \mathrm{d}x \, V_{\min} |\psi(x)|^2$$

$$= V_{\min} \int_{-\infty}^{\infty} |\psi(x)|^2$$

$$= V_{\min}$$

assuming that $|\psi\rangle$ is properly normalised. This means that the expected energy of the system is always at least the minimum of the potential. For the case of a free particle $V(x) = 0$ so $V_{\min} = 0$ which justifies only looking at positive values of E.

8.5.2   Ehrenfest's Theorems

Suppose we have a particle of mass m in a state $|\Psi(t)\rangle$. We can compute the expected value of $m\hat{X}$,

$$\langle m\hat{X} \rangle_t = \langle \Psi(t) | m\hat{X} | \Psi(t) \rangle.$$

Here we add a subscript t to the normal expected value notation to remind us that this value is time dependent. As a time dependent variable one question we may ask is how does it vary with time:

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle m\hat{X} \rangle_t = \frac{\mathrm{d}}{\mathrm{d}t} \langle \Psi(t) | m\hat{X} | \Psi(t) \rangle$$

$$= m \frac{\mathrm{d}}{\mathrm{d}t} \langle \Psi(t) | \hat{X} | \Psi(t) \rangle$$

$$= \frac{mi}{\hbar} \langle \Psi(t) | [\hat{H}, \hat{X}] | \Psi(t) \rangle.$$

Here we have used the result derived in equation 7.1. We can fairly easily compute this commutator:

$$
\begin{aligned}
[\hat{H}, \hat{X}] &= \left[\frac{\hat{P}^2}{2m} + V(\hat{X}), \hat{X} \right] \\
&= \left[\frac{\hat{P}^2}{2m}, \hat{X} \right] + \underbrace{[V(\hat{X}), \hat{X}]}_{=0} \\
&= \frac{1}{2m}[\hat{P}^2, \hat{X}] \\
&= \frac{1}{2m}(\hat{P}^2 \hat{X} - \hat{X} \hat{P}^2) \\
&= \frac{1}{2m}(\hat{P}^2 \hat{X} - \hat{P} \hat{X} \hat{P} + \hat{P} \hat{X} \hat{P} - \hat{X} \hat{P}^2) \\
&= \frac{1}{2m}(\hat{P}(\hat{P}\hat{X} - \hat{X}\hat{P}) + (\hat{P}\hat{X} - \hat{X}\hat{P})\hat{P}) \\
&= \frac{1}{2m}(\hat{P}[\hat{P}, \hat{X}] + [\hat{P}, \hat{X}]\hat{P}) \\
&= -\frac{1}{2m}(\hat{P}[\hat{X}, \hat{P}] + [\hat{X}, \hat{P}]\hat{P}) \\
&= -\frac{1}{2m}(\hat{P}i\hbar + i\hbar\hat{P}) \\
&= -\frac{i\hbar}{m}\hat{P}.
\end{aligned}
$$

Here we have used the fact that $V(\hat{X})$ is a function of $\hat{X}$ so $[V(\hat{X}), \hat{X}] = 0$, and the canonical commutation relation

$$[\hat{X}, \hat{P}] = i\hbar.$$

Substituting this back in we get

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\langle m\hat{X}\rangle_t &= -\frac{mi}{\hbar}\langle \Psi(t)|i\hbar\hat{P}|\Psi(t)\rangle \\
&= \langle \Psi(t)|\hat{P}|\Psi(t)\rangle \\
&= \langle \hat{P}\rangle_t
\end{aligned}
$$

This result is called Ehrenfest's first theorem. Compare it to the classical equivalent:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle m\hat{X}\rangle_t = \langle \hat{P}\rangle_t \longleftrightarrow \frac{\mathrm{d}}{\mathrm{d}t}(mx) = p.$$

Notice that $\langle \hat{P}\rangle_t$ also depends on time. If we compute how it changes with time we get

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\langle \hat{P}\rangle_t &= \frac{\mathrm{d}}{\mathrm{d}t}\langle \Psi(t)|\hat{P}|\Psi(t)\rangle \\
&= \frac{i}{\hbar}\langle \Psi(t)|[\hat{H}, \hat{P}]|\Psi(t)\rangle.
\end{aligned}
$$

Computing the commutator we get

$$
\begin{aligned}
[\hat{H}, \hat{P}] &= \left[\frac{\hat{P}^2}{2m} + V(\hat{X}), \hat{P} \right] \\
&= \left[\frac{\hat{P}}{2m}, \hat{P} \right]\psi(x) + [V(\hat{X}), \hat{P}]
\end{aligned}
$$

Computing this with a test function $\psi(x)$ we get

$$
\begin{aligned}
\left[V(x), -i\hbar\frac{\mathrm{d}}{\mathrm{d}x} \right] &= -i\hbar V(x)\frac{\mathrm{d}}{\mathrm{d}x}\psi(x) + i\hbar\frac{\mathrm{d}}{\mathrm{d}x}(V(x)\psi(x)) \\
&= -i\hbar V(x)\frac{\mathrm{d}\psi}{\mathrm{d}x} + i\hbar\frac{\mathrm{d}V}{\mathrm{d}x}\psi(x) + i\hbar V(x)\frac{\mathrm{d}\psi}{\mathrm{d}x}
\end{aligned}
$$

$$= i\hbar \frac{\mathrm{d}V}{\mathrm{d}x}\psi(x)$$

this holds for all ψ so

$$[\hat{H}, \hat{P}] = i\hbar \frac{\mathrm{d}}{\mathrm{d}x} V(\hat{X})$$

Hence

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \hat{P}\rangle_t = \frac{i}{\hbar}\langle \Psi(t)|i\hbar \frac{\mathrm{d}}{\mathrm{d}x}V(\hat{X})|\Psi(t)\rangle$$

$$= -\langle \Psi(t)|\frac{\mathrm{d}}{\mathrm{d}x}V(\hat{X})|\Psi(t)\rangle$$

$$= -\left\langle \frac{\mathrm{d}}{\mathrm{d}x}V(\hat{X})\right\rangle_t.$$

This is called Ehrenfest's second theorem. Compare it to its classical counterpart

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \hat{P}\rangle_t = -\left\langle \frac{\mathrm{d}}{\mathrm{d}x}V(\hat{X})\right\rangle \longleftrightarrow \frac{\mathrm{d}}{\mathrm{d}t}p = -\frac{\mathrm{d}}{\mathrm{d}x}V(x) = F$$

8.6   Degeneracy of the Hamiltonian

> **Theorem 5**
>
> If the eigenvalues of the Hamiltonian for a one-dimensional system form a discrete spectrum then there is no degeneracy.

Proof. Suppose there exists an eigenvalue, E, of the Hamiltonian that has at least two-fold degeneracy. Then there exists $|\psi_1\rangle, |\psi_2\rangle \in \mathcal{H}$ such that both $|\psi_1\rangle$ and $|\psi_2\rangle$ are eigenvectors of $\hat{H}$ with eigenvalue E:

$$\hat{H}|\psi_1\rangle = E|\psi_1\rangle$$

$$\hat{H}|\psi_2\rangle = E|\psi_2\rangle.$$

The Schrödinger equation gives us

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

for any wave function ψ. Rearranging we get

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = -\frac{2m}{\hbar^2}[E - V(x)]\psi(x).$$

Specifically in the case of $\psi_1(x) = \langle x|\psi_1\rangle$ and $\psi_2(x) = \langle x|\psi_2\rangle$ we have

$$\frac{\mathrm{d}^2\psi_1}{\mathrm{d}x^2} = -\frac{2m}{\hbar^2}[E - V(x)]\psi_1(x),$$

$$\frac{\mathrm{d}^2\psi_2}{\mathrm{d}x^2} = -\frac{2m}{\hbar^2}[E - V(x)]\psi_2(x).$$

Multiplying the first of these by $\psi_2(x)$ and the second by $\psi_1(x)$ we get

$$\frac{\mathrm{d}^2\psi_1}{\mathrm{d}x^2}\psi_2(x) = -\frac{2m}{\hbar^2}[E - V(x)]\psi_1(x)\psi_2(x),$$

and

$$\psi_1(x)\frac{\mathrm{d}^2\psi_2}{\mathrm{d}x^2} = -\psi_1(x)\frac{2m}{\hbar^2}[E - V(x)]\psi_2(x)$$

respectively. Subtracting the second of these from the first we get

$$\frac{\mathrm{d}^2\psi_1}{\mathrm{d}x^2}\psi_2(x) - \psi_1(x)\frac{\mathrm{d}^2\psi_2}{\mathrm{d}x^2} = 0 \tag{8.1}$$

Consider now

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\mathrm{d}\psi_1}{\mathrm{d}x}\psi_2(x) - \psi_1(x)\frac{\mathrm{d}\psi_2}{\mathrm{d}x}\right] = \frac{\mathrm{d}^2\psi_1}{\mathrm{d}x^2}\psi_2(x) + \frac{\mathrm{d}\psi_1}{\mathrm{d}x}\frac{\mathrm{d}\psi_2}{\mathrm{d}x} - \frac{\mathrm{d}\psi_1}{\mathrm{d}x}\frac{\mathrm{d}\psi_2}{\mathrm{d}x} - \psi_1(x)\frac{\mathrm{d}^2\psi_2}{\mathrm{d}x^2}$$
$$= \frac{\mathrm{d}^2\psi_1}{\mathrm{d}x^2}\psi_2(x) - \psi_1(x)\frac{\mathrm{d}^2\psi_2}{\mathrm{d}x^2}.$$

This is exactly the quantity that we arrived at in equation 8.1 and we showed it to be zero. This means that

$$\frac{\mathrm{d}\psi_1}{\mathrm{d}x}\psi_2(x) - \psi_1(x)\frac{\mathrm{d}\psi_2}{\mathrm{d}x} = C$$

where C is some constant. Since this is true for all x we can evaluate at any x to find C. We choose $x \to \infty$ as wave functions must vanish at infinity. Thus

$$C = \lim_{x\to\infty}\left[\frac{\mathrm{d}\psi_1}{\mathrm{d}x}\psi_2(x) - \psi_1(x)\frac{\mathrm{d}\psi_2}{\mathrm{d}x}\right] = 0.$$

Which means

$$\frac{\mathrm{d}\psi_1}{\mathrm{d}x}\psi_2(x) = \psi_1(x)\frac{\mathrm{d}\psi_2}{\mathrm{d}x}$$
$$\implies \frac{\psi_1'(x)}{\psi_1(x)} = \frac{\psi_2'(x)}{\psi_2(x)}$$
$$\implies \frac{\mathrm{d}}{\mathrm{d}x}\ln(\psi_1(x)) = \frac{\mathrm{d}}{\mathrm{d}x}\ln(\psi_2(x))$$
$$\implies \ln(\psi_1(x)) = \ln(\psi_2(x)) + \tilde{\kappa}$$

where $\tilde{\kappa}$ is a constant of integration. From this we get

$$\psi_1(x) = \kappa\psi_2(x)$$

where $\kappa = e^{\tilde{\kappa}}$. For $|\psi_1\rangle$ and $|\psi_2\rangle$ to be properly normalised we must have that κ is just some phase factor with unit modulus otherwise both couldn't be normalised at once. This means that $|\psi_1\rangle$ and $|\psi_2\rangle$ differ by at most a complex phase factor and therefore describe the same state meaning that E is not a degenerate eigenvalue.   □

Note that this only holds for discrete spectra as in the last step we assume that the states can be normalised to a finite value which is not the case for a continuous state which can only be normalised to have a finite integral. We saw in the case of the free particle that we do indeed get degenerate energy eigenvalues. Each E had two eigenfunctions, $\psi_{k,\pm}$, and we needed to introduce a second observable, the momentum, to lift the degeneracy.

We will also see later that this theorem doesn't hold in more dimensions.

9   Potentials

In this section we will look at some basic one-dimensional potentials, $V(x)$. The goal will be to find the wave function, ψ, that satisfies the TISE for these potentials:

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)\right]\psi(x) = E\psi(x).$$

9.1   Potential Step

9.1.1   The Potential

The potential step is the simplest non-trivial potential ($V(x) = V_0$ for all x is trivial as we are free to choose $V_0 = 0$). The **potential step** is

$$V(x) = V\Theta(x) = \begin{cases} 0, & x < 0, \\ V, & x > 0. \end{cases}$$

For some $V > 0$. Here Θ is the Heaviside step function. We showed in the previous section that $E \geq V_{\min}$ therefore in this case $E \geq 0$. There are two cases of interest, the first is $E > V$ and the second is $E < V$.
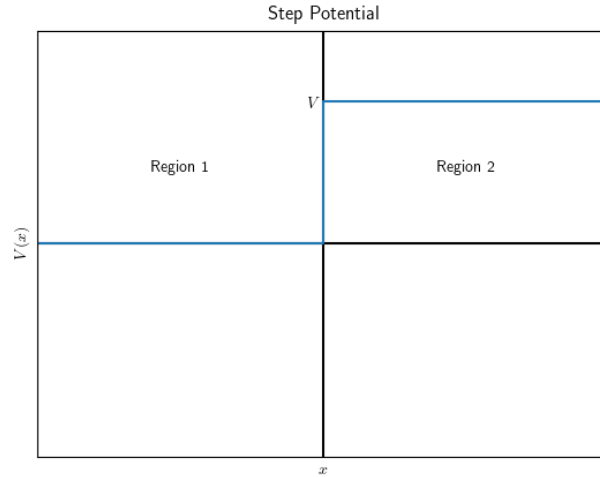
Figure 9.1: The step potential

9.1.2   Case 1: $E > V$

Classically we would expect a particle incident from the left with $E > V$ to enter region 2, $x > 0$. Since the total energy of the particle is constant and given by $H = T + V$ then we would expect that upon entering region 2 the kinetic energy, and hence momentum, of the classical particle will have to decrease to keep the total energy constant.

The TISE can be split into two section, in region 1, $x < 0$, we have

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi(x) = E\psi(x) \implies \psi''(x) = -\frac{p^2}{\hbar^2}\psi(x)$$

where

$$p = \sqrt{2mE},$$

this is exactly the momentum of a classical particle with energy E and a potential energy of 0. In region 2 we have

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V\right]\psi(x) = E\psi(x) \implies \psi''(x) = -\frac{\bar{p}}{\hbar^2}\psi(x)$$

where

$$\bar{p} = \sqrt{2m(E - V)},$$

this is exactly the momentum of a classical particle with energy E and a potential energy of V. As we would expect the momentum in region 2, $\bar{p}$, is less than the momentum in region 1, p. Fortunately we already know the solution to the wave equation in both regions:

$$\psi(x) = \begin{cases} Ae^{ipx/\hbar} + Be^{-ikp/\hbar}, & x < 0, \\ Ce^{i\bar{p}x/\hbar} + De^{-i\bar{p}x/\hbar} & x > 0. \end{cases}$$

The solutions are a superposition of waves with momentum $\pm p$ and $\pm\bar{p}$ in the respective region. By 'wave with momentum p' we mean a wave function such that p is an eigenvalue of $\hat{P}$.

Suppose that a particle comes in from the left. Then we expect a wave propagating to the right in both regions and a wave propagating to the left in region 1 corresponding to the wave reflecting off the boundary. However there is no reason to expect a wave propagating left in region 2 as once the wave enters that region there is nothing to cause a reflection. Therefore $D = 0$.

We are left with three degrees of freedom, A, B, and C. We can fix one by requiring a normalised wave function, say we choose A for this purpose. We then use the requirements of continuity of ψ and ψ' to

fix the other two degrees of freedom. Clearly ψ and ψ' are continuous at all $x \neq 0$ so we only need to consider continuity at $x = 0$. We require that

$$\lim_{x \to 0^+} \psi(x) = \lim_{x \to 0^-} \psi(x),$$

and

$$\lim_{x \to 0^+} \psi'(x) = \lim_{x \to 0^-} \psi'(x).$$

The first of these gives us

$$\lim_{x \to 0^+} \psi(x) = \lim_{x \to 0^+} \left[A e^{ipx/\hbar} + B e^{-ikx/\hbar} \right] = A + B$$

$$\lim_{x \to 0^-} \psi(x) = \lim_{x \to 0^-} C e^{i\bar{p}x/\hbar} = C.$$

Hence

$$A + B = C.$$

The second leads to

$$\lim_{x \to 0^+} \psi'(x) = \lim_{x \to 0^+} \left[\frac{ip}{\hbar} A e^{ipx/\hbar} - \frac{ip}{\hbar} B e^{-ikx/\hbar} \right] = \frac{ip}{\hbar} A + \frac{ip}{\hbar} B$$

$$\lim_{x \to 0^-} \psi'(x) = \lim_{x \to 0^-} \frac{i\bar{p}}{\hbar} C e^{i\bar{p}x/\hbar} = \frac{i\bar{p}}{\hbar} C.$$

Hence

$$p(A + B) = \bar{p}C.$$

We can then show that

$$B = \frac{p - \bar{p}}{p + \bar{p}} A, \qquad \text{and} \qquad C = \frac{2p}{p + \bar{p}} A.$$

This imposes no restrictions on the energy, E can take any value as long as it is greater than V. The Hamiltonian has a continuous spectrum. Notice that in general $B \neq 0$ so there is a reflected wave going left in region 1.

9.1.3   Case 2: $E < V$

The situation is very similar if $E < V$ but now region 2 is classically forbidden, meaning that a classical particle could not enter into this region. The equations for the wave function are the same but since $E < V$ $E - V$ is now negative meaning that $\bar{p}$ is imaginary:

$$\bar{p} = \sqrt{2m(E - V)} = i\sqrt{2m(V - E)} = i\tilde{p}.$$

In region 2 the solution to the wave function is now

$$\psi(x) = C e^{-\tilde{p}x/\hbar} + D e^{\tilde{p}x/\hbar}.$$

The first term is a decaying exponential and is finite in region 2. The second is an increasing exponential and therefore goes to infinity as $x \to \infty$. This is not allowed so we discard this term in order for ψ to be normalisable. Hence in region 2

$$\psi(x) = C e^{-\tilde{p}x/\hbar}.$$

Region 2 is classically forbidden but the wave function is non-zero in region 2 therefore there is a non-zero probability that the particle will be found in region 2. This time we can show that

$$B = \frac{p - i\tilde{p}}{P + i\tilde{p}} A.$$

The fraction here has unit modulus so we must have that $|B|^2 = |A|^2$, this means that they just differ by a phase, δ:

$$B = e^{-2i\delta(E)} A,$$

here we emphasise the fact that δ is a function of the energy (not the Dirac delta). It can also be shown that

$$\tilde{p} = p \cot \delta.$$

We define λ as the length that the particle penetrates the classically forbidden region and we note that it scales as $\hbar/p$.

9.2 Potential Barrier

Consider the potential

$$V(x) = \begin{cases} V, & x \in [0, a], \\ 0, & x \notin [0, a], \end{cases}$$

for some positive constant, V. Again there are two cases, $E > V$ and $E < V$. In the cases that $E > V$
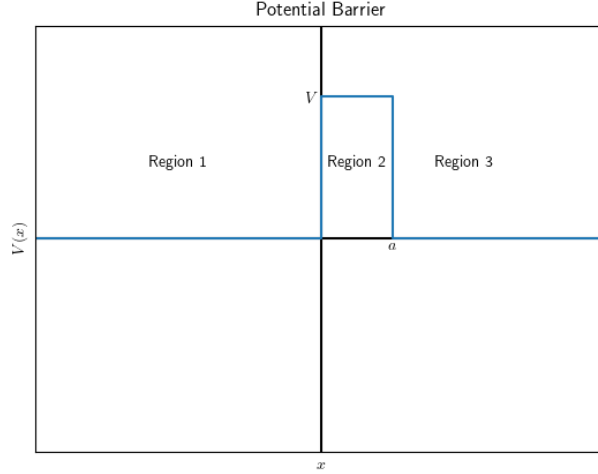
Figure 9.2: The potential barrier

this is very similar to the same case with the potential step. The wave function will be oscillatory in all three regions with an amplitude that depends on the value of the potential. The interesting case is when $E < V$.

Classically when $E < V$ region 2 is forbidden. A particle entering from the left has no way to go past the origin, it will reflect off the origin and go back to $-\infty$. The solution to the Schrödinger equation is

$$\psi(x) = \begin{cases} Ae^{ipx/\hbar} + Be^{ipx/\hbar}, & x \in (-\infty, 0), \\ Ce^{-\tilde{p}x/\hbar} + De^{\tilde{p}x/\hbar}, & x \in (0, a), \\ Fe^{ipx/\hbar} + Ge^{ipx/\hbar}, & x \in (a, \infty). \end{cases}$$

Here p and $\tilde{p}$ are defined as they were for the potential step. Notice that since x is finite in $(0, a)$ there is no reason to discard the term $De^{\tilde{p}x/\hbar}$ as it is finite for all $x \in (0, a)$. For the same reason as before if we have a particle coming in from the left we must have that $G = 0$. So in region 3

$$\psi(x) = Fe^{ipx/\hbar} = AS(E)e^{ip(x-a)/\hbar}.$$

The last part of this is the conventional way to write this term. Of particular interest is $S(E)$, this is the ratio of the probability amplitude of the wave entering the barrier at $x = 0$, to the probability amplitude of the wave leaving the barrier at $x = a$. $S(E)$ can be shown to be

$$S(E) = \frac{2ip\tilde{p}}{(p^2 - \tilde{p}^2)\sinh(\tilde{p}a/\hbar) + 2ip\tilde{p}\cosh(\tilde{p}a/\hbar)}.$$

The transmissivity, T, is defined as the probability that the particle tunnels through the barrier, that is

$$T = |S(E)|^2 = \left[1 + \frac{\sinh^2(\tilde{p}a/\hbar)}{4(E/V)(1 - E/V)}\right]^{-1}.$$

For $E < V$ this is a monotonically increasing function of E for a fixed value of a. This means that as the energy increases the probability that the particle tunnels through the barrier increases. Importantly

$$T \propto e^{-ka}$$

for some constant k. Therefore as a increases for a fixed value of E the probability that the particle tunnels into region 3 decreases.

9.2.1   Tunnelling

Tunnelling is a phenomenon without classical analogue. It can be experimentally verified and is one of many confirmations that quantum mechanics yields correct predictions. Tunnelling has many applications in physics.

The process of alpha decay can be seen as a process of the alpha particle tunnelling out of a potential well created by the nucleus. Figure 9.3 shows a phenomenological potential used to model alpha decay. The alpha particle is in a potential well and typically has less energy than the maximum potential. However it is still possible for the alpha particle to tunnel out of the well, and out of the nucleus.
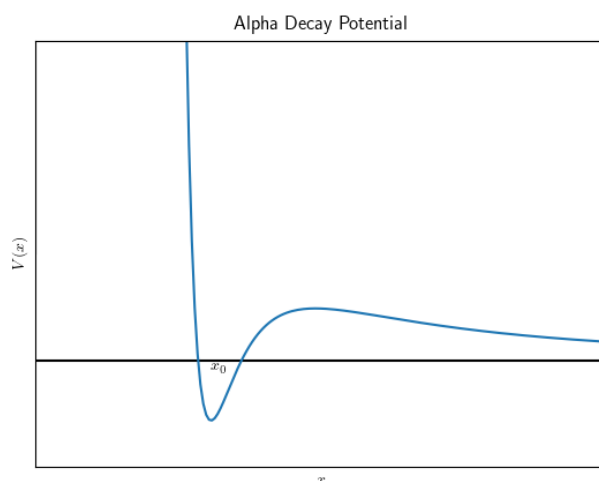
Figure 9.3: The potential modelling the alpha decay process. An alpha particle in the nucleus is in the potential well at x_0, it can then tunnel outside of this well to some positive x value even if it has less energy than the peak on the right of the potential well.

9.2.2   Scanning Tunnelling Microscope

A scanning tunnelling microscope (STM) is made of a needle that is placed near the surface of a material. A current of electrons is applied to the needle. We can model the air gap between the needle point and the surface as a potential barrier. The electrons tunnel across this gap and a current flows. The current is proportional to the probability of an electron tunnelling so

$$I \propto e^{-ka}.$$

This allows for very small changes in the size of the air gap to be measured accurately by measuring the current. Scanning the STM over the whole surface we can build up a picture of the 3-dimensional structure of the surface.

9.3   Infinite Potential Well

An infinite potential well of width a is a potential,

$$V(x) = \begin{cases} 0, & x \in (-a/2, a/2) \\ \infty, & x \notin (-a/2, a/2). \end{cases}$$

We are looking for the solutions to the TISE with positive energy (as the minimum potential is zero and the energy must be at least as large as this). We can view each boundary at $\pm a/2$ as a potential step and then let $V \to \infty$. The solution in these regions is

$$\psi \sim e^{-\bar{p}\delta/\hbar}.$$

For finite V

$$\tilde{p} = \sqrt{2m(V - E)}.$$

As $V \to \infty$ $\tilde{p} \to \infty$ and so $\psi \to 0$. We see that the regions outside of the well are not only classically forbidden but completely inaccessible, even to a quantum particle.

Therefore we only need to solve the TISE in the region $(-a/2, a/2)$ and apply boundary conditions at $x = \pm a/2$. The boundary conditions here are $\psi(\pm a/2) = 0$ as we require continuity of ψ. Note that since V is not finite we do not require continuity of ψ'. The general form of the solution in the well is that of a free particle:

$$\psi(x) = Ae^{ikx} + Be^{-ikx}$$

where $k = p/\hbar$ and $p = \sqrt{2mE}$ is the classical momentum. At the boundaries we have

$$\psi(a/2) = Ae^{ika/2} + Be^{-ika/2} = 0 \implies B = -Ae^{ika},$$

substituting this into the second boundary condition we have

$$\begin{aligned}
\psi(-a/2) &= Ae^{-ika/2} - Ae^{ika}e^{ika/2} \\
&= Ae^{ika/2}(e^{-ika} - e^{ika}) \\
&= Ae^{ika/2}[\cos(ka) - i\sin(ka) - \cos(ka) - i\sin(ka)] \\
&= 2iAe^{ika/2}\sin(ka).
\end{aligned}$$

We require this to be zero. Since $2iAe^{ika/2} \neq 0$ (assuming A is non-zero, as it must be for any wave function of interest) we must have that $\sin(ka) = 0$. Therefore

$$ka = n\pi \implies k = \frac{n\pi}{a}, \qquad n \in \mathbb{N}.$$

From this we have

$$\frac{n\pi}{a} = k = \frac{p}{\hbar} = \frac{\sqrt{2mE}}{\hbar} \implies E = \frac{n^2\pi^2\hbar^2}{2ma^2}.$$

Since $n \in \mathbb{N}$ the energy, E, is quantised. From now on we will denote the energy of the nth state as

$$E_n = \frac{1}{2m}\left(\frac{\hbar\pi n}{a}\right)^2.$$

If we normalise the wave functions we find that the properly normalised eigenstates of the Hamiltonian are

$$u_n = \begin{cases} \sqrt{\frac{2}{a}}\sin\left(\frac{n\pi x}{a}\right), & n \text{ even,} \\ \sqrt{\frac{2}{a}}\cos\left(\frac{n\pi x}{a}\right), & n \text{ odd.} \end{cases}$$

Since $u_n(x) = 0$ is boring we consider u_1 to be the ground state. In this state the particle is most likely to be found at $x = 0$. The first four wave functions, and the associated probability densities, are shown in figure 9.4. The generic state of a particle in an infinite square well potential is a linear combination of $\{u_n(x)\}$. That is a linear superposition of eigenstates of $\hat{P}$ with eigenvalues

$$\pm p_n = \pm\hbar k_n = \pm\frac{n\pi\hbar}{a}.$$

9.3.1 Zero Point Energy

Notice that in the ground state $p_1 = \pi\hbar/a \neq 0$. This can be seen as a direct consequence of HUP. Since the particle is constrained to have $x \in (-a/2, a/2)$ this gives us a maximum uncertainty in x of $\Delta x = a$. Hence

$$\Delta p \Delta x \geq \frac{\hbar}{2} \implies \Delta p \geq \frac{\hbar}{2a}.$$

Therefore the momentum cannot be zero, as then we would know the value with zero uncertainty, which is certainly less than $\hbar/2a$. The energy in this state is consequently also non-zero:

$$E \sim \frac{(\Delta p)^2}{2m} \geq \frac{\hbar^2}{8ma}.$$
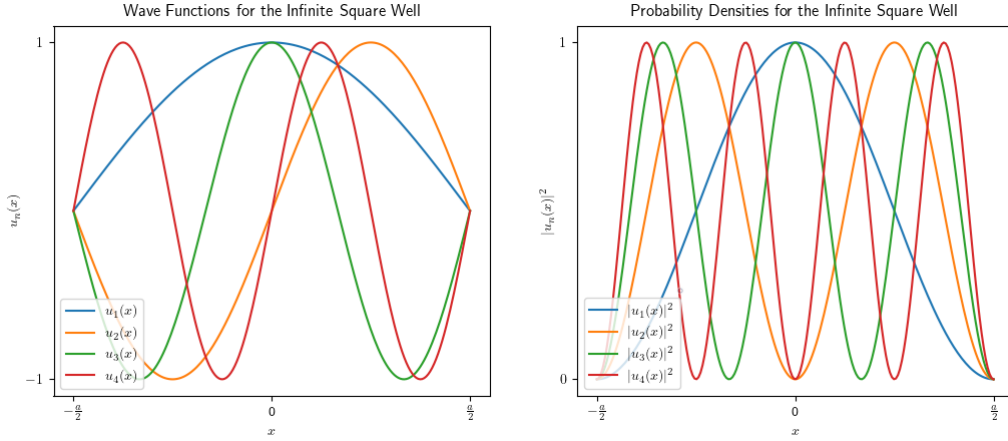
Figure 9.4: The first four wave functions, u_n, for the infinite square well and the the associated probability densities.

The kinetic energy due to the uncertainty principle is called the zero point energy. One consequence of this is that helium can be a liquid at $T = 0\,\mathrm{K}$. This is because in order to form a solid the atoms must be constrained to a lattice giving them some non-zero zero point energy, which scales inversely to the mass, for a particle as light as helium it is enough to overcome interatomic forces and the system remains in a liquid phase.

9.3.2   Symmetry of the Solutions

The energy eigenstates have definite parity, that is if we define the parity operator, $\hat{\mathscr{P}}$, to have the action

$$\hat{\mathscr{P}}\psi(x) = \psi(-x),$$

then the energy eigenstates are also eigenstates of the parity operator:

$$\hat{\mathscr{P}}u_n(x) = (-1)^{n+1}u_n(x).$$

This can be seen from the explicit form of u_n and using the fact that sine and cosine are odd and even functions respectively. This means that u_n are simultaneously eigenfunctions of $\hat{\mathscr{P}}$, $\hat{P}$, and $\hat{H}$.

We can show that in general if the potential is symmetric, so $V(-x) = V(x)$, then $[\hat{H}, \hat{\mathscr{P}}] = 0$, that is $\hat{H}$ is symmetric under $\hat{\mathscr{P}}$. This then implies the existence of a basis of simultaneous eigenstates of $\hat{H}$ and $\hat{\mathscr{P}}$.

It is fairly simple to show the above paragraph is true. Suppose that V is a symmetric potential. Then

$$\hat{\mathscr{P}}\hat{H}\psi(x) = \hat{\mathscr{P}}\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)\right]\psi(x)$$
$$= \hat{\mathscr{P}}\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2} + V(x)\psi(x)\right]$$
$$= -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi(-x)}{\mathrm{d}x^2} + V(-x)\psi(-x)$$
$$= -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi(-x)}{\mathrm{d}x^2} + V(x)\psi(-x).$$

Here we have used the fact that

$$\hat{\mathscr{P}}\frac{\mathrm{d}}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}-x} = -\frac{\mathrm{d}}{\mathrm{d}x}$$

and hence

$$\hat{\mathscr{P}}\frac{\mathrm{d}^2}{\mathrm{d}x^2} = -\frac{\mathrm{d}}{\mathrm{d}x}\left(-\frac{\mathrm{d}}{\mathrm{d}x}\right) = \frac{\mathrm{d}^2}{\mathrm{d}x^2}.$$

Also

$$\hat{H}\hat{\mathscr{P}}\psi(x) = \hat{H}\psi(-x)$$

$$= \left[-\frac{\hbar^2}{2m} \frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x) \right] \psi(-x)$$

$$= -\frac{\hbar^2}{2m} \frac{\mathrm{d}^2 \psi(-x)}{\mathrm{d}x^2} + V(x)\psi(-x).$$

Hence

$$[\hat{\mathscr{P}}, \hat{H}] = 0.$$

We know that this means there is a common basis of eigenstates but how do we find it? The states

$$v_n^+(x) = u_n(x) + u_n(-x), \qquad \text{and} \qquad v_n^-(x) = u_n(x) - u_n(-x)$$

are one such basis. These states can be compactly written as

$$v_n^\pm(x) = (1 \pm \hat{\mathscr{P}})u_n(x).$$

It is then easy to show that these are eigenstates of $\hat{H}$ and $\hat{\mathscr{P}}$:

$$\hat{H}v_n^\pm(x) = \hat{H}(1 \pm \hat{\mathscr{P}})u_n(x) = \hat{H}u_n(x) \pm \hat{H}\hat{\mathscr{P}}u_n(x) = \hat{H}u_n(x) \pm \hat{\mathscr{P}}\hat{H}u_n(x)$$
$$= E_n u_n(x) + E_n \hat{\mathscr{P}}u_n(x) = E_n(1 \pm \hat{\mathscr{P}})u_n = E_n v_n^\pm(x).$$

and

$$\hat{\mathscr{P}}v_n^\pm(x) = \hat{\mathscr{P}}(1 \pm \hat{\mathscr{P}})u_n(x) = \hat{\mathscr{P}}u_n(x) \pm \hat{\mathscr{P}}^2 u_n(x) = \hat{\mathscr{P}}u_n(x) \pm (-1)^{n+1}(-1)^{n+1}u_n(x)$$
$$= \hat{\mathscr{P}}u_n(x) \pm u_n(x) = \hat{\mathscr{P}}u_n(x) \pm u_n(x) = \pm(1 \pm \hat{\mathscr{P}})u_n(x)$$

9.4   Finite Potential Well

The finite potential well has the potential

$$V(x) = \begin{cases} -V_0, & x \int (-a/2, a/2), \\ 0, & x \notin (-a/2, a/2), \end{cases}$$

for some positive constant V_0. This potential is symmetric under parity change. This means that we can look for a set of energy eigenstates that are also parity eigenstates. The energy eigenvalues must be greater than $-V_0$. States with $-V_0 < E < 0$ are called **bound states**. The wave function for these states decays exponentially for $|x| > a/2$ so the probability of finding the particle outside of the well becomes small quickly. The states with $E > 0$ are simply plane waves that have a momentum change and chance of reflection at each boundary, this isn't that interesting and is just like the potential step so we won't consider these states further.

The Schrödinger equation is

$$\psi''(x) = \begin{cases} -\frac{2m}{\hbar^2} E \psi(x), & |x| > a/2, \\ -\frac{2m}{\hbar^2}(E + V_0)\psi(x), & |x| < a/2. \end{cases}$$

Unlike the infinite well the wave function is non-zero outside the well. For $-V_0 < E < 0$ the even parity solutions are

$$\psi(x) = \begin{cases} A\cos(px/\hbar), & |x| < a/2, \\ Ce^{-\bar{p}x/\hbar}, & x > a/2, \\ Ce^{\bar{p}x/\hbar}, & x < -a/2. \end{cases}$$

The odd parity solutions are

$$\psi(x) = \begin{cases} A\sin(px/\hbar), & |x| < a/2, \\ Ce^{-\bar{p}x/\hbar}, & x > a/2, \\ -Ce^{\bar{p}x/\hbar}, & x < -a/2. \end{cases}$$

We have introduced the momenta

$$p = \sqrt{2m(E + V_0)}, \qquad \text{and} \qquad \bar{p}\sqrt{-2mE}.$$

There are two things to notice here, first $-V_0 < E < 0$ so $p, \bar{p} \in \mathbb{R}$, and here the factor of $(E + V_0)$ appears in p whereas previously it appeared in $\bar{p}$. The reason for this is that previously we viewed the minimum of the potential as 0 and the maximum as V but here the minimum is $-V_0$ and the maximum is 0.

Imposing continuity at $a/2$ in both the wave function and its derivative we have for the even solutions that

$$A \cos \left(\frac{pa}{2\hbar} \right) = C e^{-\bar{p}a/2\hbar}$$

$$-\frac{p}{\hbar} A \sin \left(\frac{pa}{2\hbar} \right) = -\frac{\bar{p}a}{\hbar} C e^{-\bar{p}a/2\hbar}.$$

Dividing the first equation by the second we get the quantisation condition

$$p \tan \left(\frac{pa}{2\hbar} \right) = \bar{p}.$$

Similarly for the odd parity solutions we have

$$A \sin \left(\frac{pa}{2\hbar} \right) = C e^{-\bar{p}a/2\hbar}$$

$$\frac{p}{\hbar} A \cos \left(\frac{pa}{2\hbar} \right) = -\frac{\bar{p}a}{\hbar} C e^{-\bar{p}a/2\hbar}.$$

This leads to

$$p \cot \left(\frac{pa}{2\hbar} \right) = -\bar{p}.$$

There is always at least one even bound state. For small V_0, corresponding to a shallow well, it can be shown that

$$E = -\frac{m V_0^2 a^2}{2\hbar^2}$$

for the ground state.

10 Harmonic Oscillators

10.1 The Classical Harmonic Oscillators

The equation of motion for a classical particle of mass m undergoing one-dimensional harmonic oscillations, undergoing a restoring force proportional to the displacement, is

$$m\ddot{x} = -kx.$$

The constant k is often called the spring constant due to the common example of a mass on a spring where this is exactly what k is. The solution is oscillatory with an angular frequency

$$\omega = \sqrt{\frac{k}{m}}.$$

Using the chain rule we have

$$\ddot{x} = \frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = \frac{\mathrm{d}v}{\mathrm{d}t} = \frac{\mathrm{d}v}{\mathrm{d}x} \frac{\mathrm{d}x}{\mathrm{d}t} = v \frac{\mathrm{d}v}{\mathrm{d}x}.$$

Integrating the equation of motion we then have

$$m \int v \frac{\mathrm{d}v}{\mathrm{d}x} \, \mathrm{d}x = -k \int x \, \mathrm{d}x \implies \frac{1}{2} m v^2 = -\frac{1}{2} k x^2 + E$$

where E is the total energy of the system, which appears as a constant of integration. Thus

$$\frac{1}{2} m v^2 + \frac{1}{2} k x^2 = E.$$

Identifying the first term as the kinetic energy we see that the potential energy is

$$V(x) = \frac{1}{2} k x^2 = \frac{1}{2} m \omega^2 x^2.$$

The total energy is

$$E = \frac{1}{2}m\omega^2 L^2$$

where L is the amplitude of the oscillations. It is easy to show that this is the maximum energy by noting that at $x = L$ the particle is instantaneously stationary as it changes direction so the kinetic energy is zero.

10.2   The Quantum Harmonic Oscillator

The transformation to the quantum case is straight forward. The potential is simply

$$V(\hat{X}) = \frac{1}{2}m\omega^2 \hat{X}^2.$$

Assuming a discrete spectrum we have to solve the TISE:

$$\hat{H}|n\rangle = E_n|n\rangle,$$

where

$$\hat{H} = \frac{\hat{P}^2}{2m} + \frac{1}{2}m\omega^2 \hat{X}^2.$$

Taking the inner product of the TISE with an eigenstate of $\hat{X}$ we have

$$\langle x|\hat{H}|n\rangle = E_n\langle x|n\rangle = E_n u_n(x)$$

defining $u_n(x) = \langle x|n\rangle$ as the energy eigenfunction with energy E_n. The full TISE is then

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{1}{2}m\omega^2 x^2 \right] u_n(x) = E_n u_n(x).$$

10.3   The Solutions to the Harmonic Oscillator

We will start by discussing the solutions to this and look at deriving them in later sections. The first thing to note is that the energy is quantised as follows:

$$E_n = \left(n + \frac{1}{2} \right)\hbar\omega, \qquad \text{where} \qquad n \in \mathbb{N}.$$

The ground state corresponds to $n = 0$ and has the non-zero energy $E_0 = \hbar\omega/2$. The energy eigenfunctions are

$$u_n(x) = C_n e^{-\alpha^2 x^2/2} H_n(\alpha x).$$

Here C_n is a normalisation constant, $\alpha^2 = m\omega/\hbar$ and H_n are polynomials of degree n called the **Hermite polynomials**. The Hermite polynomials are orthogonal in such that

$$\int_{-\infty}^{\infty} \mathrm{d}s \, e^{-s^2} H_m(s) H_n(s) = 2^n \sqrt{\pi} n! \delta_{mn},$$

which is only non-zero if $n = m$. The first four Hermite polynomials are

$$H_0(s) = 1, \qquad\qquad\qquad H_2(s) = 4s^2 - 2,$$
$$H_1(x) = 2s, \qquad\qquad\qquad H_3(s) = 8s^3 - 12s.$$

Note that if n is even (odd) then $H_n(s)$ only contains even (odd) powers of s, this is true for all n. This means that the Hermite polynomials have even (odd) parity for even (odd) n. Since the Gaussian term of u_n has even parity this means that the parity of u_n is the same as H_n. That is

$$\hat{\mathscr{P}} u_n(x) = u_n(-x) = (-1)^n u_n(x).$$

We see that the energy eigenstates are eigenstates of the parity operator, $\hat{\mathscr{P}}$, which is a direct consequence of the potential being symmetric:

$$V(-x) = \frac{1}{2}m\omega^2(-x)^2 = \frac{1}{2}m\omega^2 x^2 = V(x).$$

Since u_n has definite parity $|u_n(x)|^2$ is even. Therefore

$$\langle x \rangle_n = \langle n|\hat{X}|n \rangle = \int_{-\infty}^{\infty} dx\, x|u_n(x)|^2 = 0,$$

as the integrand is always odd and an integral of an odd function over an even interval, $(-\infty, \infty)$, is zero. Strictly speaking this only holds if the integral converges but in this case it clearly does as the Gaussian part goes to zero much faster than the polynomial parts can go to infinity. A quick sanity check here would be to think about the potential and since it is symmetric the particle is equally likely to be at a positive x as it is at a negative x and therefore the expected value should be zero.

The next quantity that we might want to compute is the expected value of x^2 which is

$$\langle x^2 \rangle_n = \int_{-\infty}^{\infty} dx\, x^2|u_n(x)|^2 = \left(n + \frac{1}{2}\right)\frac{1}{\alpha^2}.$$

From this we can see that for larger energies it becomes more probable to find the particle at large x, that is the variance, $\Delta x_n = \langle x \rangle_n - \langle x \rangle_n^2 = \langle x^2 \rangle_n$, increases with energy:

$$E_n = m\omega^2 \langle x^2 \rangle_n.$$

Compare this to the classical energy,

$$E = \frac{1}{2}m\omega^2 L^2.$$

10.4   Solving the Harmonic Oscillator (The Fancy Way)

10.4.1   Factorising the Hamiltonian

The Hamiltonian is

$$\hat{H} = \frac{\hat{P}^2}{2m} + \frac{1}{2}m\omega^2\hat{X}^2.$$

We can factor out $\hbar\omega$ from this:

$$\hat{H} = \hbar\omega\left[\frac{1}{2m\hbar\omega}\hat{P}^2 + \frac{m\omega}{2\hbar}\hat{X}^2\right].$$

Since $\hbar\omega$ has units of energy the term in the brackets must be dimensionless. We define two operators

$$\hat{\xi} = \sqrt{\frac{m\omega}{2\hbar}}, \qquad \text{and} \qquad \hat{\eta} = \frac{1}{\sqrt{2m\hbar\omega}}\hat{P}.$$

These are simply rescaled, dimensionless, versions of the position and momentum operators. The Hamiltonian can now be written as

$$\hat{H} = \hbar\omega[\hat{\xi}^2 + \hat{\eta}^2].$$

Since $\hat{\xi}$ and $\hat{\eta}$ are proportional to the position and momentum operators we know that they don't commute, in fact their commutator is

$$\begin{aligned}
[\hat{\xi}, \hat{\eta}] &= \sqrt{\frac{m\omega}{2\hbar}}\frac{1}{\sqrt{2m\hbar\omega}}[\hat{X}, \hat{P}] \\
&= \frac{1}{2\hbar}[\hat{X}, \hat{P}] \\
&= \frac{i}{2}.
\end{aligned}$$

There are two checks that it is good to do here. First since $\hat{\xi}$ and $\hat{\eta}$ are dimensionless their commutator should be too, and it is. Second since $\hat{X}$ and $\hat{P}$ are Hermitian then so are $\hat{\xi}$ and $\hat{\eta}$ and therefore their commutator should be anti-Hermitian, and it is.

We now define a new operator:

$$\hat{a} = \hat{\xi} + i\hat{\eta}, \qquad \text{and} \qquad \hat{a}^\dagger = \hat{\xi} - i\hat{\eta}.$$

The motivation behind this is that if $\hat{\xi}$ and $\hat{\eta}$ commuted then we could factorise the Hamiltonian as

$$\hbar\omega(\hat{\xi} + i\hat{\eta})(\hat{\xi} - i\hat{\eta}).$$

However we cannot do this. We can write this new operator in terms of $\hat{X}$ and $\hat{P}$:

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\hat{X} + \frac{i}{\sqrt{2m\omega\hbar}}\hat{P} = \sqrt{\frac{m\omega}{2\hbar}}\left[\hat{X} + \frac{i}{m\omega}\hat{P}\right]$$

$$\hat{a}^{\dagger} = \sqrt{\frac{m\omega}{2\hbar}}\hat{X} - \frac{i}{\sqrt{2m\omega\hbar}}\hat{P} = \sqrt{\frac{m\omega}{2\hbar}}\left[\hat{X} - \frac{i}{m\omega}\hat{P}\right]$$

We can then compute

$$\hat{a}\hat{a}^{\dagger} = (\hat{\xi} + i\hat{\eta})(\hat{\xi} - i\hat{\eta})$$
$$= \hat{\xi}^2 + i\hat{\eta}\hat{\xi} - i\hat{\xi}\hat{\eta} + \hat{\eta}^2$$
$$= \hat{\xi} + i[\hat{\eta}, \hat{\xi}] + \hat{\eta}^2,$$

and

$$\hat{a}^{\dagger}\hat{a} = (\hat{\xi} - i\hat{\eta})(\hat{\xi} + i\hat{\eta})$$
$$= \hat{\xi}^2 + i\hat{\xi}\hat{\eta} - i\hat{\eta}\hat{\xi} + \hat{\eta}^2$$
$$= \hat{\xi}^2 - i[\hat{\eta}, \hat{\xi}] + \hat{\eta}^2.$$

Subtracting these two equations we have

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

and adding the equations we have

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

A more useful form for the Hamiltonian can be found by adding zero:

$$\hat{H} = \frac{1}{2}\hbar\omega(\hat{a}\hat{a}^{\dagger} + \hat{a}^{\dagger}\hat{a}) = \frac{1}{2}\hbar\omega(\hat{a}\hat{a}^{\dagger} - \hat{a}^{\dagger}\hat{a} + \hat{a}^{\dagger}\hat{a} + \hat{a}^{\dagger}\hat{a}) = \frac{1}{2}\hbar\omega([\hat{a}, \hat{a}^{\dagger}] + 2\hat{a}^{\dagger}\hat{a}) = \hbar\omega\left(\hat{a}^{\dagger}\hat{a} + \frac{1}{2}\right).$$

10.4.2   Creation and Annihilation

For this section we will need the commutators of $\hat{a}$ and $\hat{a}^{\dagger}$ with the Hamiltonian, fortunately these are fairly easy to compute now:

$$[\hat{H}, \hat{a}] = \left[\hbar\omega\left(\hat{a}^{\dagger}\hat{a} + \frac{1}{2}\right), \hat{a}\right] = \hbar\omega[\hat{a}^{\dagger}\hat{a}, \hat{a}]$$

using the fact that $[1/2, \hat{a}] = 0$. Then

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

Therefore

$$[\hat{H}, \hat{a}] = \hbar\omega[\hat{a}^{\dagger}\hat{a}, \hat{a}] = -\hbar\omega\hat{a}.$$

Similarly

$$[\hat{H}, \hat{a}^{\dagger}] = \left[\hbar\omega\left(\hat{a}^{\dagger}\hat{a} + \frac{1}{2}\right), \hat{a}^{\dagger}\right] = \hbar\omega[\hat{a}^{\dagger}\hat{a}, \hat{a}^{\dagger}] = \hbar\omega(\hat{a}^{\dagger}\hat{a}\hat{a}^{\dagger} - \hat{a}^{\dagger}\hat{a}^{\dagger}\hat{a}) = \hbar\omega\hat{a}^{\dagger}[\hat{a}, \hat{a}^{\dagger}] = \hbar\omega\hat{a}^{\dagger}.$$

Suppose $|n\rangle$ is an energy eigenstate. We would like to know how $\hat{a}$ acts on this since working in the energy eigenbasis we can then describe the action of $\hat{a}$ on any state. To do this we will use that

$$\hat{H}\hat{a} = \hat{a}\hat{H} + \hat{H}\hat{a} - \hat{a}\hat{H} = \hat{a}\hat{H} + [\hat{H}, \hat{a}].$$

Thus

$$\hat{H}(\hat{a}|n\rangle) = \hat{a}\hat{H}|n\rangle + [\hat{H}, \hat{a}]|n\rangle$$
$$= E_n\hat{a}|n\rangle - \hbar\omega\hat{a}|n\rangle$$
$$= (E_n - \hbar\omega)\hat{a}|n\rangle.$$

So we see that $\hat{a}|n\rangle$ is another eigenstate of $\hat{H}$ with energy $E_n - \hbar\omega$, unless $\hat{a}|n\rangle = 0$. For this reason we say that $\hat{a}$ is a **lowering operator** as its action is to return an eigenstate of lower energy. It is also called an **annihilation operator** as it annihilates one quantum of energy, $\hbar\omega$, from the system.[10] Similarly

$$\hat{H}(\hat{a} \dagger |n\rangle) = \hat{a}\hat{H}|n\rangle + [\hat{H}, \hat{a}^\dagger]$$
$$= E_n\hat{a}|n\rangle + \hbar\omega\hat{a}^\dagger$$
$$= (E_n + \hbar\omega)\hat{a}^\dagger|n\rangle.$$

So again $\hat{a}^\dagger|n\rangle$ is an energy eigenstate, this time with higher energy $E_n + \hbar\omega$. We call $\hat{a}^\dagger$ a **raising operator** as its action is to return an eigenstate of higher energy. It is also called a **creation operator** as it adds one quantum of energy, $\hbar\omega$, to the system.

If $|n\rangle$ is an energy eigenstate with energy E_n then define $|n \pm 1\rangle$ to be an energy eigenstate with energy $E_n \pm \hbar\omega$. Then

$$\hat{a}|n\rangle = c_n|n-1\rangle, \qquad \text{and} \qquad \hat{a}^\dagger|n\rangle = d_n|n+1\rangle.$$

Here c_n and d_n are constants of proportionality and not eigenvalues.

By successive application of the annihilation operator, $\hat{a}$, we can get states with lower and lower energy unless at some point $\hat{a}|0\rangle = 0$ we call the state $|0\rangle$ the ground state. Such a state must exist as the potential is bounded below and the energy is always at least as large as the minimum of the potential. We can find the ground state energy by applying the Hamiltonian:

$$\hat{H}|0\rangle = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)|0\rangle = \hbar\omega\hat{a}^\dagger\underbrace{\hat{a}|0\rangle}_{=0}\frac{\hbar\omega}{2}|0\rangle = \frac{\hbar\omega}{2}|0\rangle$$

so we see that the ground state of the Harmonic oscillator has energy $\hbar\omega/2$.

Going the other way we can start from the ground state and construct states with progressively higher and higher energies by applying the creation operator, $\hat{a}^\dagger$. Unlike with the annihilation operator there is no reason that this should ever yield zero as the energy is in theory unbounded above. The first application of the creation operator gives

$$\hat{a}^\dagger|0\rangle = d_0|1\rangle$$

where $|1\rangle$ is a state with energy

$$E_1 = \hbar\omega\left(1 + \frac{1}{2}\right) = \frac{3}{2}\hbar\omega.$$

A second application gives

$$\hat{a}^{\dagger 2}|0\rangle = d_0\hat{a}^\dagger|1\rangle = d_0 d_1|2\rangle.$$

At this point we see that we really need to know what c_n and d_n are. We find them by considering the expectation values of $\hat{a}^\dagger\hat{a}$ and $\hat{a}\hat{a}^\dagger$ and requiring all states be normalised:

$$\langle n|\hat{a}^\dagger\hat{a}|n\rangle = c_n\langle n|\hat{a}^\dagger|n-1\rangle$$
$$= c_n\langle n-1|\hat{a}|n\rangle^*$$
$$= c_n c_n^*\langle n-1|n-1\rangle$$
$$= |c_n|^2.$$

Also

$$\hat{a}^\dagger\hat{a} = \frac{1}{\hbar\omega}\hat{H} - \frac{1}{2}$$

---

[10]$\hat{a}$ and $\hat{a}^\dagger$ don't correspond to physical processes so we don't need to consider energy conservation.

so

$$\begin{aligned}
\langle n|\hat{a}^\dagger \hat{a}|n\rangle &= \frac{1}{\hbar\omega}\langle n|\hat{H}|n\rangle - \frac{1}{2}\langle n|n\rangle \\
&= \frac{1}{\hbar\omega}E_n\langle n|n\rangle - \frac{1}{2}\langle n|n\rangle \\
&= \frac{1}{\hbar\omega}\left(n+\frac{1}{2}\right)\hbar\omega - \frac{1}{2} \\
&= n
\end{aligned}$$

Since states are only defined up to a phase factor we are free to choose c_n to be a real, positive, number so

$$c_n = \sqrt{n}.$$

For d_n consider

$$\begin{aligned}
\langle n|\hat{a}\hat{a}^\dagger|n\rangle &= d_n\langle n|\hat{a}|n+1\rangle \\
&= d_n\langle n+1|\hat{a}^\dagger|n\rangle^* \\
&= d_n d_n^*\langle n+1|n+1\rangle
\end{aligned}$$

Also

$$\hat{a}\hat{a}^\dagger = \frac{1}{\hbar\omega}\hat{H} + \frac{1}{2}$$

so

$$\begin{aligned}
\langle n|\hat{a}\hat{a}^\dagger|n\rangle &= \frac{1}{\hbar\omega}\langle n|\hat{H}|n\rangle + \frac{1}{2}\langle n|n\rangle \\
&= \frac{1}{\hbar\omega}E_n\langle n|n\rangle + \frac{1}{2}\langle n|n\rangle \\
&= \frac{1}{\hbar\omega}\left(n+\frac{1}{2}\right)\hbar\omega + \frac{1}{2} \\
&= n+1.
\end{aligned}$$

So, again choosing d_n to be real and positive, we have

$$d_n = \sqrt{n+1}.$$

So the full action of the raising and lowering operators is

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle, \qquad \text{and} \qquad \hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle.$$

We see that the action of $\hat{a}$ on the ground state is

$$\langle x|\hat{a}|0\rangle = \langle x|\sqrt{0}|?\rangle = 0$$

for all $|x\rangle$, as expected.

10.4.3   The Wave Function

We have introduced these creation and annihilation operators but we have yet to show that they allow us to find the correct wave functions for the system. We define $\langle x|n\rangle = u_n(x)$ and also we know for the ground state that $\langle x|\hat{a}|0\rangle = \hat{a}u_0(x) = 0$. We can also expand this with the definition of $\hat{a}$ in terms of position and momentum operators:

$$\begin{aligned}
0 &= \hat{a}u_0(x) \\
&= \sqrt{\frac{m\omega}{2\hbar}}\left[\hat{X} + \frac{i}{m\omega}\hat{P}\right]u_0(x) \\
&= \sqrt{\frac{m\omega}{2\hbar}}\left[x - \frac{i}{m\omega}i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\right]u_0(x) \\
\implies 0 &= xu_0(x) + \frac{\hbar}{m\omega}\frac{\mathrm{d}u_0}{\mathrm{d}x}
\end{aligned}$$

$$\implies \frac{\mathrm{d}u_0}{\mathrm{d}x} = -\frac{m\omega}{\hbar}xu_0(x)$$
$$= \alpha^2 xu_0(x)$$
$$\implies \frac{1}{u_0}\frac{\mathrm{d}u_0}{\mathrm{d}x} = -\alpha x$$
$$\implies \int \frac{1}{u_0}\,\mathrm{d}u_0 = -\alpha^2 \int x\,\mathrm{d}x$$
$$\implies \ln(u_0) = -\frac{1}{2}\alpha^2 x^2 + \ln(N)$$

where $\ln C$ is a constant of integration

$$\implies u_0 = Ce^{-\alpha^2 x^2/2}.$$

We can find C by requiring the state to be normalised so

$$1 = \int_{-\infty}^{\infty} |u_0(x)|^2\,\mathrm{d}x$$
$$= C^2 \int_{-\infty}^{\infty} e^{-\alpha^2 x^2}$$
$$= C^2 \frac{\sqrt{\pi}}{\alpha}$$
$$\implies C = \frac{\sqrt{\alpha}}{\pi^{1/4}}.$$

where we have used the standard result

$$\int_{-\infty}^{\infty} e^{-p(x+q)^2}\,\mathrm{d}x = \sqrt{\frac{\pi}{p}}.$$

We can then construct any wave function by successive applications of $\hat{a}^\dagger$. However we must account for the constants of proportionality, for example

$$\hat{a}^{\dagger 4}|0\rangle = \sqrt{1}\hat{a}^{\dagger 3}|1\rangle = \sqrt{1}\sqrt{2}\hat{a}^{\dagger 2}|2\rangle = \sqrt{1}\sqrt{2}\sqrt{3}\hat{a}^\dagger|3\rangle = \sqrt{1}\sqrt{2}\sqrt{3}\sqrt{4}|4\rangle = \sqrt{4!}|4\rangle.$$

In general

$$|n\rangle = \frac{1}{\sqrt{n!}}\hat{a}^{\dagger n}|0\rangle.$$

Note that these states may still not be normalised as $\hat{a}^\dagger$ is not Hermitian so there is no guarantee that it preserves inner products. In terms of wave functions then we have

$$u_n(x) = \frac{1}{\sqrt{n!}}\hat{a}^{\dagger n}u_0(x)$$
$$= \frac{1}{\sqrt{n!}}\left[\sqrt{\frac{m\omega}{2\hbar}}\left(\hat{X} - \frac{i}{m\omega}\hat{P}\right)\right]^n u_0(x)$$
$$= \frac{1}{\sqrt{n!}}\left[\sqrt{\frac{m\omega}{2\hbar}}\left(x - \frac{\hbar}{m\omega}\frac{\mathrm{d}}{\mathrm{d}x}\right)\right]^n u_0(x)$$

Computing this we will recover the Hermite polynomials.

11   Systems in Three Spatial Dimensions

11.1   States

As with the one-dimensional case the state is described by a state vector, $|\psi\rangle \in \mathcal{H}$. The position eigenbasis is now represented by position vectors, $\boldsymbol{r}$, so the basis is $\{|\boldsymbol{r}\rangle\}$. A generic state vector is then

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

where
$$\psi(\boldsymbol{r}) = \langle \boldsymbol{r} | \psi \rangle.$$

The probability density is given by $|\psi(\boldsymbol{r})|^2$ and $|\psi(\boldsymbol{r})|^2 \, \mathrm{d}^3 r$ then gives the probability that the particle is found in the volume element $\mathrm{d}^3 r$ at position $\boldsymbol{r}$. Introducing some time dependence a generic state becomes
$$|\Psi(t)\rangle = \int \mathrm{d}^3 r \, |\boldsymbol{r}\rangle\langle \boldsymbol{r}|\Psi(t)\rangle = \int \mathrm{d}^3 r \, \Psi(\boldsymbol{r}, t)|\boldsymbol{r}\rangle,$$

where
$$\Psi(\boldsymbol{r}, t) = \langle \boldsymbol{r} | \Psi(t) \rangle.$$

An inner product between two states, $|\Psi(t)\rangle, |\Phi(t)\rangle \in \mathcal{H}$ is

$$\langle \Phi(t)|\Psi(t)\rangle = \int \mathrm{d}^3 r \, \Phi^*(\boldsymbol{r}, t)\Psi(\boldsymbol{r}, t).$$

As usual we require normalised states so

$$\langle \Psi(t)|\Psi(t)\rangle = \int_{\text{all space}} \mathrm{d}^3 r \, |\Psi(t)|^2 = 1.$$

The only change so far from the one-dimensional case is that x becomes $\boldsymbol{r}$ and $\mathrm{d}x$ becomes $\mathrm{d}^3 r$.

11.2   Observables

An observable, O, is still associated with a Hermitian operator, $\hat{O} \colon \mathcal{H} \to \mathcal{H}$ which we can think of as acting on the wave function with the usual identification that

$$\hat{O}|\psi\rangle = \langle \boldsymbol{r}|\hat{O}|\psi\rangle.$$

The eigenvalue equation is still
$$\hat{O}\psi_k(\boldsymbol{r}) = O_k \psi_k(\boldsymbol{r})$$

where $\{\psi_k\}$ are the eigenfunctions of $\hat{O}$. Again $\{|\psi_k\rangle\}$ form a basis for $\mathcal{H}$. The eigenvalues, $\{O_k\}$, are all possible outcomes of the measurement and the specific eigenvalue O_k, when measuring O in state $|\psi\rangle \in \mathcal{H}$, occurs with probability
$$|c_k|^2 = |\langle \psi_k|\psi\rangle|^2.$$

The Hermitian conjugate is still defined as

$$\langle \varphi|\hat{O}^\dagger|\psi\rangle = \langle \psi|\hat{O}|\varphi\rangle^*,$$

or in terms of wave functions

$$\int \mathrm{d}^3 r \, \varphi^*(\boldsymbol{r})\hat{O}^\dagger\psi(\boldsymbol{r}) = \left[\int \mathrm{d}^3 r \, \psi^*(\boldsymbol{r})\hat{o}\varphi(\boldsymbol{r}) \right]^*.$$

11.2.1   Position Operator

The three-dimensional position operator is a vector of one-dimensional position operators:

$$\hat{\boldsymbol{X}} = (\hat{X}, \hat{Y}, \hat{Z}) = (\hat{X}_1, \hat{X}_2, \hat{X}_3).$$

Which have the expected application of multiplication by their respective Cartesian coordinate:

$$\hat{X}_k \psi(\boldsymbol{r}) = x_k \psi(\boldsymbol{r})$$

where $\boldsymbol{r} = (x_1, x_2, x_3)$.

11.2.2   Momentum Operator

The three-dimensional momentum operator is a vector of one-dimensional momentum operators:

$$\hat{\boldsymbol{P}} = (\hat{P}_x, \hat{P}_y, \hat{P}_z) = (\hat{P}_1, \hat{P}_2, \hat{P}_3).$$

Each one-dimensional operator is

$$\hat{P}_k = -i\hbar\frac{\partial}{\partial x_k} = -i\hbar\partial_k,$$

which when applied to a wave function gives

$$\hat{P}_k\psi(\boldsymbol{r}) = -i\hbar\frac{\partial\psi}{\partial x_i}\bigg|_{\boldsymbol{r}}.$$

We can concisely write the momentum operator with some vector calculus notation as

$$\hat{\boldsymbol{P}} = -i\hbar\nabla$$

where $\nabla = (\partial_1, \partial_2, \partial_3)$ is the gradient operator.

11.2.3   The Canonical Commutation Relation

The partial derivatives commute so

$$\hat{P}_i\hat{P}_j\psi(\boldsymbol{r}) = -\hbar^2\frac{\partial^2\psi}{\partial x_i\partial x_j} = -\hbar^2\frac{\partial^2\psi}{\partial x_j\partial x_i} = \hat{P}_j\hat{P}_i\psi(\boldsymbol{r}).$$

This means that

$$[\hat{P}_i, \hat{P}_j] = 0.$$

Clearly $x_ix_j = x_jx_i$ so

$$[\hat{X}_i, \hat{X}_j] = 0.$$

More interestingly we can look at $[\hat{X}_i, \hat{P}_j]$. We know that in the case $i = j$ this is the same as the one-dimensional canonical commutation relation and therefore equal to $i\hbar$. In general

$$\hat{X}_i\hat{P}_j\psi = -i\hbar\hat{X}_i\frac{\partial\psi}{\partial x_j} = -i\hbar x_i\frac{\partial\psi}{\partial x_j}$$

and

$$\hat{P}_j\hat{X}_i\psi = \hat{P}_j[x_i\psi] = -i\hbar\frac{\partial x_i}{\partial x_j} - i\hbar x_i\frac{\partial\psi}{\partial x_j} = \hat{P}_j[x_i\psi] = -i\hbar\delta_{ij} - i\hbar x_i\frac{\partial\psi}{\partial x_j}.$$

Therefore

$$[\hat{X}_i, \hat{P}_j] = -i\hbar x_i\frac{\partial\psi}{\partial x_j} + i\hbar\delta_{ij} + i\hbar x_i\frac{\partial\psi}{\partial x_j} = i\hbar\delta_{ij}.$$

In the case that $i = j$ we have that $\delta_{ij} = 1$ and so this reduces to the expected relation in one-dimension.

Since orthogonal position and momentum operators commute there is no need for them to follow the HUP, however parallel position and momentum operators must still follow the HUP, thus

$$\Delta x_i\Delta p_j \geq \frac{\hbar}{2}\delta_{ij}.$$

11.3   Dynamics

The dynamics of the three-dimensional system are still given by the Schrödinger equation,

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

where

$$\hat{H} = \frac{\hat{\boldsymbol{P}}\cdot\hat{\boldsymbol{P}}}{2m} + V(\hat{\boldsymbol{X}}).$$

The action of this one a wave function is then

$$\hat{H} = \frac{-\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r}).$$

The TISE is still

$$\hat{H}|\psi\rangle = E|\psi\rangle$$

and the time evolution operator is still

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

11.3.1   Three-Dimensional Harmonic Oscillator

The three-dimensional isotropic harmonic oscillator has the same potential

$$V(\boldsymbol{r}) = \frac{1}{2}m\omega^2 r^2 = \frac{1}{2}m\omega^2(x^2 + y^2 + z^2).$$

We can separate the TISE into a product of three terms,

$$\psi(\boldsymbol{r}) = X(x)Y(y)Z(z),$$

which each depend on only one coordinate. Using this the TISE becomes

$$EX(x)Y(y)Z(z) = \left[-\frac{\hbar^2}{2m}\nabla^2 + \frac{1}{2}m\omega^2(x^2 + y^2 + z^2)\right]X(x)Y(y)Z(z)$$

$$= -\frac{\hbar^2}{2m}\left(\frac{\mathrm{d}^2 X}{\mathrm{d}x^2}Y(y)Z(z) + X(x)\frac{\mathrm{d}^2 Y}{\mathrm{d}y^2}Z(z) + X(x)Y(y)\frac{\mathrm{d}^2 Z}{\mathrm{d}z^2}\right)$$

$$+ \frac{1}{2}m\omega^2(x^2 + y^2 + z^2)X(x)Y(y)Z(z)$$

Since the three terms X, Y, and Z only depend on one coordinate and therefore their derivatives with respect to any other coordinate are zero. Dividing through by $X(x)Y(y)Z(z)$ and factorising we have

$$E = \left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2 X}{\mathrm{d}x^2}\frac{1}{X} + \frac{1}{2}m\omega^2 x^2\right] + \left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2 Y}{\mathrm{d}y^2}\frac{1}{Y} + \frac{1}{2}m\omega^2 y^2\right] + \left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2 Z}{\mathrm{d}z^2}\frac{1}{Z} + \frac{1}{2}m\omega^2 z^2\right].$$

The left hand side is a constant, each term in brackets depends on only one coordinate, and we can vary two coordinates without varying the third. This means that each term in brackets must be constant so

$$E_{n_x} + E_{n_y} + E_{n_z} = E$$

where

$$E_{n_x} = \left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2 X}{\mathrm{d}x^2}\frac{1}{X} + \frac{1}{2}m\omega^2 x^2\right],$$

and E_{n_y} and E_{n_z} are similarly defined. We can identify each term as a Harmonic oscillator in one dimension and we already know the solutions for these. The energy corresponding to the x direction is then

$$E_{n_x} = \left(n_x + \frac{1}{2}\right)\hbar\omega.$$

The solution in the x direction is

$$X(x) = u_{n_x}(x) = C_{n_{(x)}}e^{-\alpha^2 x^2/2}H_{n_x}(\alpha x).$$

The same can be done for the y and z directions. The energy eigenvalue for the whole system is then

$$E_n = \left(n_x + n_y + n_z + \frac{3}{2}\right)\hbar\omega = \left(n + \frac{3}{2}\right)\hbar\omega.$$

The solution to the whole system is

$$\psi_{n_x,n_y,n_z}(\boldsymbol{r}) = u_{n_x}(x)u_{n_y}(y)u_{n_z}(z).$$

n	n_x	n_y	n_z	g_n
0	0	0	0	1
	1	0	0	
1	0	1	0	3
	0	0	1	
	2	0	0	
	0	2	0	
2	0	0	2	6
	1	1	0	
	1	0	1	
	0	1	1	

Table 11.1: Demonstrating the degeneracy of the E_n energy eigenvalue, for $n = 0, 1, 2$, for the three-dimensional harmonic oscillator.

One new thing that we didn't have in the one-dimensional case is degeneracy. Suppose the total energy is $5\hbar\omega/2$. That is $n = 1$. Then there are three ways to have this happen, any one of n_x, n_y, and n_z could be one and the other two will be zero. We say that E_n is three-fold degenerate. Table 11.1 shows different ways that the same energy eigenvalue, E_n, can be degenerate for $n = 0, 1, 2$. Clearly these grow quite quickly. With some basic combinatorics we can find a general statement for the degeneracy of the nth energy eigenvalue. First choose n_x to be some integer such that $0 \leq n_x \leq n$. Then $n_y + n_z = n - n_x$ so pick n_y to be some integer such that $0 \leq n_y \leq n - n_x$. Finally choose n_z such that $n_x + n_y + n_z = n$. There are $n + 1$ ways to choose n_x and then for each of these choices there are $n - n_x + 1$ ways to choose n_y and then n_z is fixed by these choices. Thus

$$
\begin{aligned}
g_n &= \sum_{n_x=0}^{n+1} (n - n_x + 1) \\
&= \sum_{n_x=0}^{n+1} (n + 1) - \sum_{n_x=0}^{n+1} n_x \\
&= (n + 1)(n + 1) - \frac{1}{2}(n + 1)(n + 2) \\
&= (n + 1)\left[(n + 2) - \frac{1}{2}(n + 2)\right] \\
&= \frac{1}{2}(n + 1)(n + 2)
\end{aligned}
$$

Where we have used the result that

$$
\sum_{k=0}^{N} k = \frac{1}{2}N(N + 1)
$$

which can be readily proven by induction, as we have done in appendix B.

Part IV

Angular Momentum and Spin

12   Angular Momentum

12.1   Angular Momentum in Classical Mechanics

In classical mechanics the angular momentum of a particle at position $\boldsymbol{x}$ with momentum $\boldsymbol{p}$ is

$$
\boldsymbol{L} = \boldsymbol{x} \times \boldsymbol{p}.
$$

As a vector equation this can be written as three scalar equations:

$$L_x = yp_z - zp_y, \qquad L_y = zp_x - xp_z, \qquad \text{and} \qquad L_z = xp_y - yp_x.$$

These can be written more concisely as

$$L_i = \varepsilon_{ijk} x_j p_k$$

where $(x_1, x_2, x_3) = (x, y, z)$ and $(p_1, p_2, p_3) = (p_x, p_y, p_z)$ and ε_{ijk} is the Levi-Civita symbol defined as

$$\varepsilon_{ijk} = \begin{cases} 1, & (i, j, k) = (1, 2, 3), (2, 3, 1), (3, 1, 2), \\ -1, & (i, j, k) = (1, 3, 2), (2, 1, 3), (3, 2, 1), \\ 0, & \text{otherwise,} \end{cases}$$

and we have applied the Einstein summation convention of summing over repeated indices.

12.2   Angular Momentum in Quantum Mechanics

12.2.1   Angular Momentum Operators

Angular momentum is an observable and so has an associated Hermitian operator. The **angular momentum operator** for the angular momentum in the $\boldsymbol{e_i}$ direction is

$$\hat{L}_i = \varepsilon_{ijk} \hat{X}_j \hat{P}_k.$$

We can write this as a cross product of vectors of operators:

$$\hat{\boldsymbol{L}} = \hat{\boldsymbol{X}} \times \hat{\boldsymbol{P}}.$$

For a state $|\psi\rangle \in \mathcal{H}$ if we define $\hat{L}_k |\psi\rangle = |\varphi\rangle$ then

$$\langle \boldsymbol{r} | \hat{L}_k | \psi \rangle = \langle \boldsymbol{r} | \varphi \rangle = \varphi(\boldsymbol{r}) = \hat{L}_k \psi(\boldsymbol{r}) = -\hbar \varepsilon_{ijk} x_j \frac{\partial \psi}{\partial x_k}.$$

12.2.2   Angular Momentum Commutators

Since $\hat{L}_i$ are defined as a product of position and momentum operators we expect the commutation relations between different components of the angular momentum to be non-trivial, for example:

$$\begin{aligned} [\hat{L}_x, \hat{L}_y] &= [\hat{Y}\hat{P}_z - \hat{Z}\hat{P}_y, \hat{Z}\hat{P}_x - \hat{X}\hat{P}_z] \\ &= [\hat{Y}\hat{P}_z, \hat{Z}\hat{P}_x] - [\hat{Z}\hat{P}_y, \hat{Z}\hat{P}_x] - [\hat{Y}\hat{P}_z, \hat{X}\hat{P}_z] + [\hat{Z}\hat{P}_y, \hat{X}\hat{P}_z] \\ &= [\hat{Y}\hat{P}_z, \hat{Z}\hat{P}_x] - \hat{Z}[\hat{P}_y, \hat{P}_x] - [\hat{Y}, \hat{X}]\hat{P}_z + [\hat{Z}\hat{P}_y, \hat{X}\hat{P}_z] \\ &= \hat{Y}\hat{P}_z\hat{Z}\hat{P}_x - \hat{Z}\hat{P}_x\hat{Y}\hat{P}_z + \hat{Z}\hat{P}_y\hat{X}\hat{P}_z - \hat{X}\hat{P}_z\hat{Z}\hat{P}_y \\ &= \hat{Y}\hat{P}_x(\hat{P}_z\hat{Z} - \hat{Z}\hat{P}_z) + \hat{X}\hat{P}_y(\hat{Z}\hat{P}_z - \hat{P}_z\hat{Z}) \\ &= \hat{Y}\hat{P}_x[\hat{P}_z, \hat{Z}] + \hat{X}\hat{P}_y[\hat{Z}, \hat{P}_z] \\ &= -i\hbar\hat{Y}\hat{P}_x + i\hbar\hat{X}\hat{P}_y \\ &= i\hbar\hat{L}_z. \end{aligned}$$

A couple of sanity checks that we can perform here are that $\hat{L}_k$ is Hermitian so the commutator should be anti-Hermitian, which it is, and also the units of the commutator are angular momentum squared and $\hbar$ and $\hat{L}_z$ also both have units of angular momentum so the units match. Noting that $[\hat{L}_i, \hat{L}_j] = -[\hat{L}_j, \hat{L}_i]$ and $[\hat{L}_i, \hat{L}_i] = 0$ we have that

$$[\hat{L}_i, \hat{L}_j] = i\hbar \varepsilon_{ijk} \hat{L}_k.$$

Since this is non-zero when $i \neq j$ we see that measuring the angular moment in one direction and then another changes the probability for measuring the angular momentum in the first direction.

12.2.3   Angular Momentum Differential Operators

As with the other operators we have introduced we can view $\hat{L}_i$ as differential operators acting on the wave function. In this case

$$\hat{L}_x = \varepsilon_{1jk}\hat{X}_j\hat{P}_k = \hat{X}_2\hat{P}_3 - \hat{X}_3\hat{P}_2 = -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right).$$

Similarly

$$\hat{L}_y = -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right), \qquad \text{and} \qquad \hat{L}_z = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right).$$

Or more compactly

$$\hat{L}_i = -i\hbar\varepsilon_{ijk}x_j\frac{\partial}{\partial x_k}.$$

We can use these to derive the commutation relationships that we derived in the last section using the canonical commutation relationship:

$$\begin{aligned}
\hat{L}_x\hat{L}_y\psi &= -\hbar^2\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)\psi \\
&= -\hbar^2\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)\left(z\frac{\partial\psi}{\partial x} - x\frac{\partial\psi}{\partial z}\right) \\
&= -\hbar^2\left[y\frac{\partial}{\partial z}\left(z\frac{\partial\psi}{\partial x}\right) - y\frac{\partial}{\partial z}\left(x\frac{\partial\psi}{\partial z}\right) - z\frac{\partial}{\partial y}\left(z\frac{\partial\psi}{\partial x}\right) + z\frac{\partial}{\partial y}\left(x\frac{\partial\psi}{\partial z}\right)\right] \\
&= -\hbar^2\left[y\frac{\partial\psi}{\partial x} + yz\frac{\partial^2\psi}{\partial z\partial x} - yx\frac{\partial^2\psi}{\partial z^2} - z^2\frac{\partial^2\psi}{\partial y\partial x} + zx\frac{\partial^2\psi}{\partial y\partial z}\right]
\end{aligned}$$

Similarly

$$\begin{aligned}
\hat{L}_y\hat{L}_x\psi &= -\hbar^2\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)\psi \\
&= -\hbar^2\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)\left(y\frac{\partial\psi}{\partial z} - z\frac{\partial\psi}{\partial y}\right) \\
&= -\hbar^2\left[z\frac{\partial}{\partial x}\left(y\frac{\partial\psi}{\partial z}\right) - z\frac{\partial}{\partial x}\left(z\frac{\partial\psi}{\partial y}\right) - x\frac{\partial}{\partial z}\left(y\frac{\partial\psi}{\partial z}\right) + x\frac{\partial}{\partial z}\left(z\frac{\partial\psi}{\partial y}\right)\right] \\
&= -\hbar^2\left[zy\frac{\partial^2\psi}{\partial x\partial z} - z^2\frac{\partial^2\psi}{\partial x\partial y} - x\frac{\partial^2\psi}{\partial z^2} + xz\frac{\partial^2\psi}{\partial z\partial y} + x\frac{\partial\psi}{\partial y}\right]
\end{aligned}$$

Hence

$$[\hat{L}_x, \hat{L}_y]\psi = -\hbar^2\left(x\frac{\partial\psi}{\partial y} - y\frac{\partial\psi}{\partial x}\right) = i\hbar\hat{L}_z\psi.$$

Angular momentum is often of importance when discussing central potentials, that is potentials of the form $V(r)$. It is often best to work spherical coordinates when this is the case:

$$\hat{L}_x = i\hbar\left[\sin\varphi\frac{\partial}{\partial\vartheta} + \cot\vartheta\cos\varphi\frac{\partial}{\partial\varphi}\right],$$

$$\hat{L}_y = i\hbar\left[-\cos\varphi\frac{\partial}{\partial\vartheta} + \cot\vartheta\sin\varphi\frac{\partial}{\partial\varphi}\right],$$

$$\hat{L}_z = -i\hbar\frac{\partial}{\partial\varphi}.$$

12.2.4   Square of the Angular Momentum

We now introduce the square of the magnitude of the angular momentum:

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 = \sum_{i=1}^{3}\hat{L}_i^2.$$

This can be written as a differential operator in spherical coordinates as

$$\hat{L}^2 = -\hbar^2 \left[\frac{1}{\sin\vartheta} \frac{\partial}{\partial\vartheta} \left(\sin\vartheta \frac{\partial}{\partial\vartheta} \right) + \frac{1}{\sin^2\vartheta} \frac{\partial^2}{\partial\varphi^2} \right]. \tag{12.1}$$

Importantly this is compatible with any of the Cartesian components of the angular momentum. For example,

$$\begin{aligned}
[\hat{L}^2, \hat{L}_z] &= [\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2, \hat{L}_z] \\
&= [\hat{L}_x^2, \hat{L}_z] + [\hat{L}_y^2, \hat{L}_z] + [\hat{L}_z^2, \hat{L}_z] \\
&= [\hat{L}_x^2, \hat{L}_z] + [\hat{L}_y^2, \hat{L}_z] \\
&= \hat{L}_x\hat{L}_x\hat{L}_z - \hat{L}_z\hat{L}_x\hat{L}_x + \hat{L}_y\hat{L}_y\hat{L}_z - \hat{L}_z\hat{L}_y\hat{L}_y \\
&= \hat{L}_x\hat{L}_x\hat{L}_z - \hat{L}_z\hat{L}_x\hat{L}_x + \hat{L}_y\hat{L}_y\hat{L}_z - \hat{L}_z\hat{L}_y\hat{L}_y \\
&= \hat{L}_x\hat{L}_z\hat{L}_x - \hat{L}_x[\hat{L}_z, \hat{L}_x] + \hat{L}_z\hat{L}_x\hat{L}_x + \hat{L}_y\hat{L}_z\hat{L}_y - \hat{L}_y[\hat{L}_z, \hat{L}_y] - \hat{L}_z\hat{L}_y\hat{L}_y \\
&= [\hat{L}_x, \hat{L}_z]\hat{L}_x - \hat{L}_x[\hat{L}_z, \hat{L}_x] + [\hat{L}_y, \hat{L}_z]\hat{L}_y - \hat{L}_y[\hat{L}_z, \hat{L}_y] \\
&= -i\hbar\hat{L}_y\hat{L}_x - i\hbar\hat{L}_x\hat{L}_y + i\hbar\hat{L}_x\hat{L}_y + i\hbar\hat{L}_y\hat{L}_x \\
&= 0.
\end{aligned}$$

Here we have used

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

to introduce the first two commutators. The second two commutators come from factorising.

12.2.5   Eigenfunctions of $\hat{L}^2$ and $\hat{L}_z$

Since $\hat{L}^2$ is compatible with all Cartesian components of the angular momentum we know that $\hat{L}^2$ and $\hat{L}_z$ must have simultaneous eigenfunctions, and a common eigenbasis. We work with $\hat{L}_z$ here as is the convention but all of this holds for $\hat{L}_x$ and $\hat{L}_y$ as well.

As with the Harmonic oscillator we will introduce the solution first and discus its properties and derive the solution later. The simultaneous eigenfunctions of $\hat{L}^2$ and $\hat{L}_z$ are the **spherical harmonics**:

$$Y_\ell^m(\vartheta, \varphi) = (-1)^m \sqrt{\frac{2\ell+1}{4\pi} \frac{(\ell-m)!}{(\ell+m)!}} P_\ell^m(\cos\vartheta) e^{im\varphi}.$$

Here $\ell = 0, 1, 2, \ldots$ is the **angular momentum quantum number** and $m = -\ell, -\ell+1, \ldots, \ell-1, \ell$ is the **magnetic quantum number**. P_ℓ^m are the **associated Legendre polynomials**. The first four polynomials for $\ell = 0, 1$ are

$$P_0^0(\cos\vartheta) = 1, \qquad P_1^0(\cos\vartheta) = \cos\vartheta, \qquad P_1^1(\cos\vartheta) = -\sin\vartheta, \qquad \text{and} \qquad P_1^{-1}(\cos\vartheta) = \frac{1}{2}\sin\vartheta.$$

The first four spherical harmonics are then

$$Y_0^0(\vartheta, \varphi) = \sqrt{\frac{1}{4\pi}}, \qquad\qquad\qquad Y_1^0(\vartheta, \varphi) = \sqrt{\frac{3}{4\pi}} \cos\vartheta$$

$$Y_1^1(\vartheta, \varphi) = -\sqrt{\frac{3}{8\pi}} \sin\vartheta e^{i\varphi} \qquad\qquad Y_1^{-1}(\vartheta, \varphi) = \sqrt{\frac{3}{8\pi}} \sin\vartheta e^{-i\varphi}.$$

The spherical harmonics are the eigenfunctions that solve the eigenvalue equations

$$\hat{L}^2 Y_\ell^m(\vartheta, \varphi) = \hbar^2 \ell(\ell+1) Y_\ell^m(\vartheta, \varphi),$$

and

$$\hat{L}_z Y_\ell^m(\vartheta, \varphi) = m\hbar Y_\ell^m(\vartheta, \ell).$$

This means that measuring the magnitude of the angular momentum of a state with wave function Y_ℓ^m we find it to be $\hbar\sqrt{\ell(\ell+1)}$. However, we still often say that this state has angular momentum ℓ when it would be more accurate to say that it has angular momentum quantum number ℓ, just be aware of this

linguistic simplification. The eigenvalues, $\hbar^2\ell(\ell+1)$, for $\hat{L}^2$ have $(2\ell+1)$-fold degeneracy. We create a CSCO by also labelling the eigenfunctions with the eigenvalues $m\hbar$ from $\hat{L}_z$. It is easy to show these are the eigenvalues of $\hat{L}_z$ as the only φ dependency is in the exponential so for a differential operator with respect to φ its action on $Y_\ell^m(\vartheta, \varphi)$ is equivalent to its action on $e^{im\varphi}$:

$$\hat{L}_z Y_\ell^m(\vartheta, \varphi) = -i\hbar\frac{\partial}{\partial\varphi}Y_\ell^m(\vartheta, \varphi) = m\hbar Y_\ell^m(\vartheta, \varphi).$$

12.3   Algebraic Solution to the Eigenvalue Problem

We have defined $\hat{L}_i$ and $\hat{L}^2$ as angular momentum operators based on the classical definition $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$. In this section we generalise this and define new operators $\hat{J}_i$ and $\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2$ without making any assumptions about the physics that these operators represent. The only thing we assume about these operators is their commutation relations:

$$[\hat{J}_k, \hat{J}_l] = i\hbar\varepsilon_{klm}\hat{J}_m, \qquad \text{and} \qquad [\hat{J}^2, \hat{J}_k] = 0.$$

These are the same as the angular momentum operators of course but this generalisation allows us to use the work in this section at a later point when we talk about spin.

We can build a common eigenbasis for $\hat{J}^2$ and $\hat{J}_z$. We denote vectors in this basis set as $|\lambda, m\rangle$ where

$$\hat{J}^2|\lambda, m\rangle = \hbar^2\lambda|\lambda, m\rangle$$

and

$$\hat{J}_z|\lambda, m\rangle = \hbar m|\lambda, m\rangle.$$

12.3.1   Raising and Lowering Operators

Raising and lowering operators were useful before and we can define something analogous here. Let

$$\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y.$$

Then

$$[\hat{J}^2, \hat{J}_\pm] = 0$$

so

$$\hat{J}^2\hat{J}_\pm|\lambda, m\rangle = \hat{J}_\pm\hat{J}^2|\lambda, m\rangle = \hbar^2\lambda\hat{J}_\pm|\lambda, m\rangle.$$

Therefore $\hat{J}_\pm|\lambda, m\rangle$ are still eigenstates of $\hat{J}^2$ with the same eigenvalue, $\hbar^2\lambda$, as $|\lambda, m\rangle$.

Computing the following commutator

$$\begin{aligned}
[\hat{J}_z, \hat{J}_\pm] &= [\hat{J}_z, \hat{J}_x \pm i\hat{J}_y] \\
&= [\hat{J}_z, \hat{J}_x] \pm i[\hat{J}_z, \hat{J}_y] \\
&= i\hbar\hat{J}_y \pm i(-i\hbar\hat{J}_x) \\
&= \pm\hbar\hat{J}_\pm
\end{aligned}$$

we see that

$$\begin{aligned}
\hat{J}_z\hat{J}_\pm|\lambda, m\rangle &= (\hat{J}_\pm\hat{J}_z - [\hat{J}_\pm, \hat{J}_z])|\lambda, m\rangle \\
&= \hat{J}_\pm\hat{J}_z|\lambda, m\rangle + [\hat{J}_z, \hat{J}_\pm]|\lambda, m\rangle \\
&= \hbar m\hat{J}_\pm|\lambda, m\rangle \pm \hbar\hat{J}_\pm|\lambda, m\rangle \\
&= \hbar(m+1)\hat{J}_\pm|\lambda, m\rangle.
\end{aligned}$$

So we see that $\hat{J}_\pm|\lambda, m\rangle$ is an eigenvector of $\hat{J}_z$ with eigenvalue $m \pm 1$. In this way we see that $\hat{J}_+$ is a raising operator as it increases the angular momentum in the z direction by $\hbar$ and $\hat{J}_-$ is a lowering operator as it lowers the angular momentum in the z direction by $\hbar$. We can write this as

$$\hat{J}_\pm|\lambda, m\rangle = c_\pm\hbar|\lambda, m \pm 1\rangle$$

where $c_\pm$ are dimensionless constants of proportionality.

12.3.2   Bounds on m

Consider

$$\langle \lambda, m|(\hat{J}^2 - \hat{J}_z^2)|\lambda, m\rangle = \langle \lambda, m|\hat{J}^2|\lambda, m\rangle - \langle \lambda, m|\hat{J}_z^2|\lambda, m\rangle$$
$$= \hbar^2 \lambda - \hbar^2 m^2,$$

assuming that $|\lambda, m\rangle$ is normalised. From the definition of $\hat{J}^2$ it is clear that

$$\hat{J}^2 - \hat{J}_z^2 = \hat{J}_x^2 + \hat{J}_y^2.$$

Thus

$$\hbar^2 \lambda - \hbar^2 m^2 = \langle \lambda, m|(\hat{J}^2 - \hat{J}_z^2)|\lambda, m\rangle = \langle \lambda, m|(\hat{J}_x^2 + \hat{J}_y^2)|\lambda, m\rangle = \langle \hat{J}_x^2 + \hat{J}_y^2\rangle \geq 0.$$

Thus

$$|m| \leq \sqrt{\lambda}.$$

So the spectrum of $\hat{J}_z$ is bounded above and below for any given value of λ. Since we have raising and lowering operators there must be states $|\lambda, m_{\max}\rangle$ and $|\lambda, m_{\min}\rangle$ such that

$$\hat{J}_+|\lambda, m_{\max}\rangle = 0, \qquad \text{and} \qquad \hat{J}_-|\lambda, m_{\min}\rangle = 0.$$

Here $m_{\max}$ and $m_{\min}$ are the maximum and minimum values that m can take for a given λ. Consider now products of the ladder operators:

$$\hat{J}_+\hat{J}_- = (\hat{J}_x + i\hat{J}_y)(\hat{J}_x - i\hat{J}_y)$$
$$= \hat{J}_x^2 + \hat{J}_y^2 + i\hat{J}_y\hat{J}_x - i\hat{J}_x\hat{J}_y$$
$$= \hat{J}_x^2 + \hat{J}_y^2 + i[\hat{J}_y, \hat{J}_x]$$
$$= \hat{J}_x^2 + \hat{J}_y^2 - i[\hat{J}_x, \hat{J}_y]$$
$$= \hat{J}_x^2 + \hat{J}_y^2 + \hbar\hat{J}_z$$
$$= \hat{J}^2 - \hat{J}_z^2 + \hbar\hat{J}_z,$$

and

$$\hat{J}_-\hat{J}_+ = (\hat{J}_x - i\hat{J}_y)(\hat{J}_x + i\hat{J}_y)$$
$$= \hat{J}_x^2 + \hat{J}_y^2 - i\hat{J}_y\hat{J}_x + i\hat{J}_x\hat{J}_y$$
$$= \hat{J}_x^2 + \hat{J}_y^2 + i[\hat{J}_x, \hat{J}_y]$$
$$= \hat{J}_x^2 + \hat{J}_y^2 - \hbar\hat{J}_z$$
$$= \hat{J}^2 - \hat{J}_z^2 - \hbar\hat{J}_z,$$

By the definition of the state $|\lambda, m_{\min}\rangle$

$$\hat{J}_+\hat{J}_-|\lambda, m_{\min}\rangle = 0,$$

by the product calculated above this is also equal to

$$\hat{J}_+\hat{J}_-|\lambda, m_{\min}\rangle = (\hat{J}^2 - \hat{J}_z^2 + \hbar\hat{J}_z)|\lambda, m_{\min}\rangle$$
$$= \hat{J}^2|\lambda, m_{\min}\rangle - \hat{J}_z^2|\lambda, m_{\min}\rangle + \hbar\hat{J}_z|\lambda, m_{\min}\rangle$$
$$= \hbar^2\lambda|\lambda, m_{\min}\rangle - \hbar^2 m_{\min}^2|\lambda, m_{\min}\rangle + \hbar^2 m_{\min}|\lambda, m_{\min}\rangle.$$

Combining these results we have that

$$\lambda - m_{\min}^2 + m_{\min} = 0 \implies \lambda = m_{\min}(m_{\min} - 1).$$

Similarly by the definition of the state $|\lambda, m_{\max}\rangle$

$$\hat{J}_-\hat{J}_+|\lambda, m_{\max}\rangle = 0,$$

by the product calculated above this is also equal to

$$\hat{J}_-\hat{J}_+|\lambda, m_{\max}\rangle = (\hat{J}^2 - \hat{J}_z^2 - \hbar\hat{J}_z)|\lambda, m_{\max}\rangle$$
$$= \hat{J}^2|\lambda, m_{\max}\rangle - \hat{J}_z^2|\lambda, m_{\max}\rangle - \hbar\hat{J}_z|\lambda, m_{\max}\rangle$$
$$= \hbar^2\lambda - \hbar^2 m_{\max}^2 - \hbar^2 m_{\max}$$

Combining these results we have that

$$\lambda - m_{\max}^2 - m_{\max} = 0 \implies \lambda = m_{\max}(m_{\max} + 1).$$

It is common to denote $m_{\max}$ by j so we will do this here. Thus

$$\lambda = j(j + 1) = m_{\min}(m_{\min} - 1).$$

Which can be rearranged to get

$$m_{\min}^2 - m_{\min} - j^2 - j = 0$$

which can be factorised to get

$$(m_{\min} + j)(m_{\min} - j - 1) = 0.$$

Hence there are two options, $m_{\min} = -j$ and $m_{\min} = j + 1$. We discard the second of these as no values of m, including $m_{\min}$, can be larger than $j = m_{\max}$. So $m_{\min} - j$. Since $m_{\min}$ and $m_{\max}$ are integers we have that

$$m_{\max} - m_{\min} = k, \qquad k = 0, 1, 2, \ldots$$

so

$$j - (-j) = 2j = k \implies j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots .$$

For a given value of j we see that m ranges over

$$m = -j, -j + 1, \ldots, j - 1, j.$$

There are a total of $2j + 1$ values that m can take. Since $\lambda = j(j + 1)$ we can equally well denote a given state as $|\lambda, m\rangle$ or $|j, m\rangle$ and the second of these is often more useful. Notice that with $\hat{L}^2$ as the angular momentum operator we had that ℓ was an integer but here we allow j to be half integer as well. The extra constraint occurs due to the physics of the angular momentum which we assumed for $\hat{L}^2$ but not $\hat{J}^2$ as we defined $\hat{J}^2$ simply by its commutation properties. The requirement that ℓ is integral is because we require the wave function to be single valued and the solutions to the angular momentum eigenequations are the spherical harmonics, Y_ℓ^m. Also $Y_\ell^m(\vartheta, \varphi) \sim e^{im\varphi}$ if we ignore ϑ dependence. Thus we require that

$$Y_\ell^m(\vartheta, \varphi + 2\pi) = Y_\ell^m(\vartheta, \varphi)$$

which only happens if $m \in \mathbb{Z}$. If m had a half integral value then this would introduce a negative into this equation.

12.3.3   Normalisation Factors

We introduced a constant, $c_\pm$, that appeared in the equation

$$\hat{J}_\pm|j, m\rangle = \hbar c_\pm|j, m \pm 1\rangle.$$

We can now find the value of this constant. First consider

$$\langle j, m|\hat{J}_\pm^\dagger\hat{J}_\pm|j, m\rangle = \langle j, m|\hat{J}_\mp\hat{J}_\pm|j, m\rangle$$
$$= \langle j, m|(\hat{J}^2 - \hat{J}_z^2 \mp \hbar\hat{J}_z)|j, m\rangle$$
$$= \hbar^2 j(j + 1) - \hbar^2 m^2 \mp \hbar^2 m$$
$$= \hbar^2[j(j + 1) - m(m \pm 1)].$$

Now consider an alternate way of calculating this:

$$\langle j, m|\hat{J}_\pm^\dagger\hat{J}_\pm|j, m\rangle = \hbar c_\pm\langle j, m|\hat{J}_\pm^\dagger|j, m \pm 1\rangle$$

$$= \hbar c_{\pm} \langle j, m \pm 1 | \hat{J}_{\pm} | j, m \rangle^*$$
$$= \hbar^2 c_{\pm} c_{\pm}^* \langle j, m \pm 1 | j, m \pm 1 \rangle$$
$$= \hbar^2 |c_{\pm}|.$$

Combining these results and choosing $c_{\pm}$ to be real and positive we have

$$c_{\pm} = \sqrt{j(j+1) - m(m \pm 1)}.$$

13   Central Potentials

13.1   What is a Central Potential

A **central potential** is any potential that depends only on the distance, r, from the origin. For example the Coulomb potential:

$$V(r) = -\frac{Ze^2}{4\pi\varepsilon_0 r}.$$

Since these potentials depend only on r the natural choice of coordinates to work in is spherical coordinates since this will result in the potential being a function of r and not ϑ or φ.

13.2   Stationary States

As with the previous potentials we have considered we want to solve the TISE for a generic central potential, $V \colon [0, \infty) \to \mathbb{R}$:

$$\left[-\frac{\hbar^2}{2\mu} \nabla^2 + V(r) \right] u(r, \vartheta, \varphi) = E u(r, \vartheta, \varphi).$$

Here we denote the mass by μ to avoid confusion with the magnetic quantum number at a later stage. The Laplacian in spherical coordinates is

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin\vartheta} \frac{\partial}{\partial \vartheta} \left(\sin\vartheta \frac{\partial}{\partial \vartheta} \right) + \frac{1}{r^2 \sin^2\vartheta} \frac{\partial^2}{\partial \varphi^2}.$$

Recall that in equation 12.1 we saw how the operator $\hat{L}^2$ could be expressed in spherical polar coordinates. We see that it matches the last two terms here up to a factor of $-\hbar^2 r^2$. So we can write the Laplacian as

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{1}{\hbar^2 r^2} \hat{L}^2.$$

Thus the TISE becomes

$$E u(r, \vartheta, \varphi) = \left[-\frac{\hbar^2}{2\mu} \left(\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{1}{\hbar^2 r^2} \hat{L}^2 \right) + V(r) \right] u(r, \vartheta, \varphi)$$
$$= \left[-\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{2\mu r^2} \hat{L}^2 + V(r) \right] u(r, \vartheta, \varphi).$$

Similarly to how we split the three-dimensional harmonic oscillator into three one dimensional harmonic oscillators we can split this into a part that depends exclusively on r and part that has angular dependence. The part with angular dependence is the term with $\hat{L}^2$. We already know that the spherical harmonics, Y_ℓ^m, are the eigenfunctions of $\hat{L}^2$ so we look for a wave function of the form

$$u(r, \vartheta, \varphi) = R(r) Y_\ell^m(\vartheta, \varphi)$$

where $R \colon [0, \infty) \to \mathbb{C}$ characterises the radial dependence of the wave function. Considering only the term with angular dependence we have

$$\frac{1}{2\mu r^2} \hat{L}^2 u(r, \vartheta, \varphi) = \frac{1}{2\mu r^2} \hat{L}^2 R(r) Y_\ell^m(\vartheta, \varphi) = \frac{1}{2\mu r^2} R(r) \hat{L}^2 Y_\ell^m(\vartheta, \varphi) = \frac{\hbar^2 \ell(\ell+1)}{2\mu r^2} R(r) Y_\ell^m(\vartheta, \varphi).$$

Thus the TISE becomes

$$\left[-\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + V(r) + \frac{\hbar^2 \ell(\ell+1)}{2\mu r^2} \right] R_\ell(r) Y_\ell^m(\vartheta, \varphi) = E R_\ell Y_\ell^m(\vartheta, \varphi).$$

Here we have introduced a subscript, ℓ, for the radial component as there is ℓ dependence in the Hamiltonian so we expect R to have ℓ dependence. Notice also that the Hamiltonian has no dependence on m. This means that each solution will have $(2\ell + 1)$-fold degeneracy as there will be one solution for each value that m can take. We can factorise out the spherical harmonics to get

$$\left[-\frac{\hbar^2}{2\mu}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + V(r) + \frac{\hbar^2\ell(\ell+1)}{2\mu r^2}\right]R_\ell(r) = ER_\ell.$$

This can be further simplified by defining

$$R_\ell(r) = \frac{\chi_\ell(r)}{r}$$

so the TISE becomes

$$
\begin{aligned}
E\frac{\chi_\ell(r)}{r} &= \left[-\frac{\hbar^2}{2\mu}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + V(r) + \frac{\hbar^2\ell(\ell+1)}{2\mu r^2}\right]\frac{\chi_\ell(r)}{r} \\
&= -\frac{\hbar^2}{2\mu}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)\frac{\chi_\ell(r)}{r} + \left[V(r) + \frac{\hbar^2\ell(\ell+1)}{2\mu r^2}\right]\frac{\chi_\ell(r)}{r} \\
&= -\frac{\hbar^2}{2\mu}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\left\{\frac{d\chi_\ell}{dr}\frac{1}{r} - \frac{\chi_\ell(r)}{r^2}\right\}\right) + \left[V(r) + \frac{\hbar^2\ell(\ell+1)}{2\mu r^2}\right]\frac{\chi_\ell(r)}{r} \\
&= -\frac{\hbar^2}{2\mu}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r\frac{d\chi_\ell}{dr} - \chi_\ell(r)\right) + \left[V(r) + \frac{\hbar^2\ell(\ell+1)}{2\mu r^2}\right]\frac{\chi_\ell(r)}{r} \\
&= -\frac{\hbar^2}{2\mu}\frac{1}{r^2}\left(\frac{d\chi_\ell}{dr} + r\frac{d^2\chi_\ell}{dr^2} - \frac{d\chi_\ell}{dr}\right) + \left[V(r) + \frac{\hbar^2\ell(\ell+1)}{2\mu r^2}\right]\frac{\chi_\ell(r)}{r} \\
&= -\frac{\hbar^2}{2\mu}\frac{1}{r}\frac{d^2\chi_\ell}{dr^2} + \left[V(r) + \frac{\hbar^2\ell(\ell+1)}{2\mu r^2}\right]\frac{\chi_\ell(r)}{r}
\end{aligned}
$$

We can write this as

$$E\chi_\ell(r) = \left[-\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial r^2} + V_{\text{eff}}(r)\right]\chi_\ell(r) \tag{13.1}$$

where we have defined an effective potential,

$$V_{\text{eff}}(r) = V(r) + \frac{\hbar^2\ell(\ell+1)}{2\mu r^2}.$$

This can be thought of as the initial potential, V, plus a centrifugal term. This is analogous to the central potential that occurs in classical mechanics when dealing with rotating reference frames.

13.3   Normalisation

The normalisation condition is the same as always in three dimensions:

$$\int d^3r\,|u(r,\vartheta,\varphi)|^2 = 1.$$

Expanding this and defining $d\Omega = \sin\vartheta\,d\vartheta\,d\varphi = d\cos\vartheta\,d\varphi$ we get

$$\int dr\,r^2|R(r)|^2\int d\Omega\,|Y_\ell^m(\vartheta,\varphi)|^2 = 1.$$

Since the spherical harmonics are already properly normalised this second integral is one. The first integral is

$$\int_0^\infty dr\,r^2\frac{|\chi_\ell(r)|^2}{r^2} = \int_0^\infty dr\,|\chi_\ell(r)|^2 = 1.$$

Which is exactly the normalisation condition that we would expect for a one-dimensional wave function for a particle constrained to have $r \in [0,\infty)$.

One thing that we have to mention at this point is that this integral must be finite for the wave function to be normalisable. Since the wave function must be finite at every point we require $\chi_\ell(0) = 0$ so that the wave function doesn't diverge.[11]

---

[11]strictly we just require that $\chi(r) \to 0$ faster than $1/r \to 0$ as $r \to 0$.

13.4   Behaviour Near the Origin

Suppose that the potential is such that

$$\lim_{r \to 0}(V(r)r^2) = 0,$$

that is the potential diverges at the origin at most as quickly as $1/r^2$. The effective potential near zero can then be approximated as

$$V_{\text{eff}}(r) \approx \frac{\hbar^2 \ell(\ell+1)}{2\mu r^2}.$$

Suppose also that $R_\ell(r) \propto r^s$ for some constant s. Then $\chi_\ell(r) \propto r^{s+1}$. Putting this in the TISE we get

$$Er^{s+1} = -\frac{\hbar^2}{2\mu}\frac{\partial}{\partial r}r^{s+1} + \frac{\hbar^2}{2\mu r^2}\ell(\ell+1)$$

$$= -\frac{\hbar^2}{2\mu}s(s+1)r^{s-1} + \frac{\hbar^2}{2\mu}\ell(\ell+1)r^{s-1}.$$

Since we are concerned with behaviour as $r \to 0$ the term Er^{s+1} goes to zero much faster than the two terms proportional to r^{s-1} so we have

$$s(s+1) = \ell(\ell+1)$$

which means that either $s = \ell$ or $s = -\ell - 1$. Clearly the second of these doesn't go to zero as $r \to 0$ so we discard it. Figure 13.1 shows $|\chi_\ell(r)|^2$ plotted for $\ell = 0, 1, 2$. We see that the larger the value of ℓ the
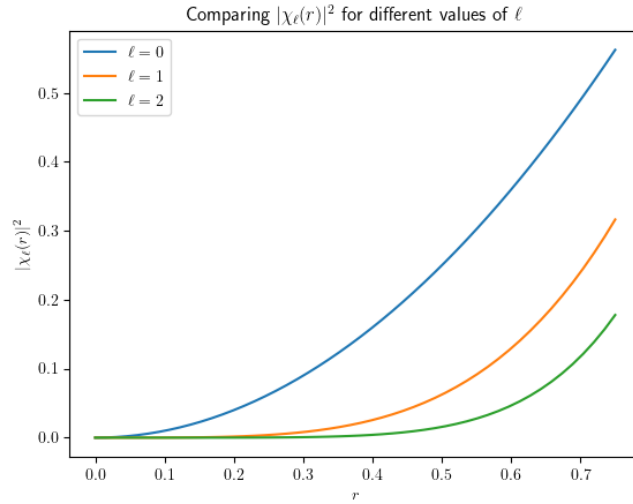
Figure 13.1: $|\chi_\ell(r)|^2$ compared for $\ell = 0, 1, 2$ and $r < 1$.

less likely the particle is to be near the origin. This can be see as analogous to the classical centrifugal potential which provides a force away from the origin. In quantum mechanics a centrifugal term in the potential instead decreases the probability of the particle being near the origin.

There are two probabilities that we may be interested in. The probability density function (PDF) for finding the particle at a specific point, $\boldsymbol{r}$ is $|u(r, \vartheta, \varphi)|^2$ so the probability that the particle is in the volume dV is $|u(r, \vartheta, \varphi)|^2 \, dV$. The probability that the particle is in a given volume, V, is then

$$\int_V |u(r, \vartheta, \varphi)|^2.$$

The PDF for finding the particle at some specific radius, $\boldsymbol{r}$ is $|\chi_\ell(r)|^2$ so the probability that the particle is at a point with radius $[r, r + dr]$ is

$$\int d\Omega \, |u(r, \vartheta, \varphi)|^2 = \int d\Omega \, |Y_\ell^m(\vartheta, \varphi)|^2 |\chi_\ell(r)|^2 \, dr = |\chi_\ell(r)|^2 \, dr.$$

The probability that the particle is at some point with a radius in $[a, b]$ is then

$$\int_a^b \mathrm{d}r \, |\chi_\ell(r)|^2.$$

14 Spin

We saw in section 12.3.2 that the angular momentum quantum number, ℓ, is always an integer despite the fact that the similarly defined j can take half integer values. This is because of the spatial dependence of the solutions. For the eigenfunctions of $\hat{L}^2$ to be single valued we require that

$$Y_\ell^m(\vartheta, \varphi) = Y_\ell^m(\vartheta, \varphi + 2\pi)$$

and we saw that the φ dependence of Y_ℓ^m was

$$Y_\ell^m(\vartheta, \varphi) \propto e^{im\varphi}.$$

In combination these two equations mean that m must be an integer. Since m takes on any value from $-\ell$ to ℓ this means that ℓ must be an integer.

Spin is an intrinsic property of a quantum system. It is like angular momentum in that the spin is associated with the operator

$$\hat{\boldsymbol{S}} = (\hat{S}_1, \hat{S}_2, \hat{S}_3),$$

with the commutation relation

$$[\hat{S}_k, \hat{S}_l] = i\hbar \varepsilon_{klm} \hat{S}_m.$$

This was the assumption that we made at the start of section 12.3 so we can use everything we derived in that section. In particular we can define

$$\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2.$$

$\hat{S}^2$ and $\hat{S}_z$ are a CSCO and have a common eigenbasis:

$$\left\{ |s, m\rangle \mid m = -s, 1 - s, \ldots, s - 1, s \text{ and } s = \frac{n}{2} \text{ where } n \in \mathbb{N} \right\}.$$

The action of $\hat{S}^2$ and $\hat{S}_z$ on these basis vectors is

$$\hat{S}^2 |s, m\rangle = s(s+1)\hbar^2 |s, m\rangle, \qquad \text{and} \qquad \hat{S}_z |s, m\rangle = m\hbar |s, m\rangle.$$

For a fixed value of s we see that this is a $(2s + 1)$-state system.

From now on we focus only on spin $1/2$ systems. That is systems where $s = 1/2$. This includes electrons, protons, neutrons, neutrinos, and quarks. So most matter that we come across day to day. In this case we have a two dimensional vector space, $\mathcal{H}_{\text{spin}}$, with a basis

$$\left\{ |\tfrac{1}{2}, \tfrac{1}{2}\rangle, |\tfrac{1}{2}, -\tfrac{1}{2}\rangle \right\}.$$

Since $s = 1/2$ for both of these it is common to introduce a shorter notation where we refer to $|\frac{1}{2}, \frac{1}{2}\rangle$ as spin up, denoted $|\frac{1}{2}\rangle$ or $|\uparrow\rangle$, and $|\frac{1}{2}, -\frac{1}{2}\rangle$ as spin down, denoted $|-\frac{1}{2}\rangle$ or $|\downarrow\rangle$. These are still eigenfunctions of both $\hat{S}^2$ and $\hat{S}_z$:

$$\hat{S}^2 |\uparrow\rangle = \frac{3}{4}\hbar^2 |\uparrow\rangle, \qquad , \hat{S}^2 |\downarrow\rangle = \frac{3}{4}\hbar^2 |\downarrow\rangle,$$

$$\hat{S}_z |\uparrow\rangle = \frac{\hbar}{2}, \qquad \text{and} \qquad \hat{S}_z |\downarrow\rangle = -\frac{\hbar}{2}|\downarrow\rangle.$$

A generic spin state, $|\psi\rangle \in \mathcal{H}_{\text{spin}}$ is given by

$$|\psi\rangle = \psi_1 |\uparrow\rangle + \psi_2 |\downarrow\rangle$$

where as usual $\psi_1 = \langle\uparrow|\psi\rangle$, $\psi_2 = \langle\downarrow|\psi\rangle$, and $|\psi_1|^2 + |\psi_2|^2 = 1$. Since spin has no spatial dependence we describe spin with vectors in $\mathcal{H}_{\text{spin}}$. If a system also has spatial dependence then we can describe this

with vectors in some other vector space $\mathcal{H}_{\text{space}}$. The full description of a state is then given by a vector in

$$\mathcal{H} = \mathcal{H}_{\text{space}} \otimes \mathcal{H}_{\text{spin}}.$$

$\mathcal{H}$ has as a basis

$$\left\{ |\boldsymbol{r}, m\rangle \mid \boldsymbol{r} \in \mathbb{R}^3 \text{ and } m = \pm\frac{1}{2} \right\}.$$

The vectors in this basis are simply the tensor products of vectors of the individual bases of $\mathcal{H}_{\text{space}}$ and $\mathcal{H}_{\text{spin}}$:

$$|\boldsymbol{r}, m\rangle = |\boldsymbol{r}\rangle|m\rangle = |\boldsymbol{r}\rangle \otimes |m\rangle.$$

This basis is complete, meaning

$$\int \mathrm{d}^3 r' \sum_m |\boldsymbol{r}', m\rangle\langle\boldsymbol{r}', m| = \mathbb{1}.$$

A generic state, $|\psi\rangle \in \mathcal{H}$, is then given by

$$|\psi\rangle = \int \mathrm{d}^3 r \sum_m |\boldsymbol{r}', m\rangle\langle\boldsymbol{r}', m|\psi\rangle$$

$$= \int \mathrm{d}^3 r' \sum_m \psi_m(\boldsymbol{r}')|\boldsymbol{r}', m\rangle$$

$$= \int \mathrm{d}^3 r' \left[\psi_1(\boldsymbol{r}')|\uparrow\rangle + \psi_2(\boldsymbol{r}')|\downarrow\rangle \right]$$

where $\psi_m(\boldsymbol{r}') = \langle\boldsymbol{r}', m|\psi\rangle$. Taking the inner product with a position eigenstate, $|\boldsymbol{r}\rangle$, we have

$$\langle\boldsymbol{r}|\psi\rangle = \int \mathrm{d}^3 r' \sum_m \langle\boldsymbol{r}|\boldsymbol{r}'\rangle|m\rangle\psi_m(\boldsymbol{r}')$$

$$= \int \mathrm{d}^3 r' \sum_m \delta(\boldsymbol{r} - \boldsymbol{r}')|m\rangle\psi_m(\boldsymbol{r}')$$

$$= \sum_m \int \mathrm{d}^3 r' \, \delta(\boldsymbol{r} - \boldsymbol{r}')|m\rangle\psi_m(\boldsymbol{r}')$$

$$= \sum_m \psi_m(\boldsymbol{r})$$

$$= \psi_1(\boldsymbol{r})|\uparrow\rangle + \psi_2(\boldsymbol{r})|\downarrow\rangle.$$

In the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$ for $\mathcal{H}_{\text{spin}}$ we can denote $\langle\boldsymbol{r}|\psi\rangle$ as a column vector:

$$\langle\boldsymbol{r}|\psi\rangle \to \begin{pmatrix} \psi_1(\boldsymbol{r}) \\ \psi_2(\boldsymbol{r}) \end{pmatrix}.$$

The interpretation of these values is as follows:

- $\left|\psi_1(\boldsymbol{r})^2\right|$ – the PDF of finding the system at $\boldsymbol{r}$ with spin up.

- $\left|\psi_2(\boldsymbol{r})^2\right|$ – the PDF of finding the system at $\boldsymbol{r}$ with spin down.

- The probability of finding the system anywhere in space with spin up is

$$P(\uparrow) = \int \mathrm{d}^3 r \, |\psi_1(\boldsymbol{r})|^2.$$

- The probability of finding the system anywhere in space with spin down is

$$P(\downarrow) = \int \mathrm{d}^3 r \, |\psi_2(\boldsymbol{r})|^2.$$

- The probability of finding the system at $\boldsymbol{r}$ with any spin is

$$P(\boldsymbol{r}) = |\psi_1(\boldsymbol{r})|^2 + |\psi_2(\boldsymbol{r})|^2.$$

14.1   Matrix Elements

$\mathcal{H}_{\mathrm{spin}}$ is a two-dimensional vector space. Therefore the spin operators can be represented by 2×2 matrices acting on column vectors. First we start with

$$\hat{S}_z|\uparrow\rangle = \frac{\hbar}{2}|\uparrow\rangle, \qquad \text{and} \qquad \hat{S}_z|\downarrow\rangle = -\frac{\hbar}{2}|\downarrow\rangle.$$

Hence

$$\langle\uparrow|\hat{S}_z|\uparrow\rangle = \frac{\hbar}{2},$$
$$\langle\uparrow|\hat{S}_z|\downarrow\rangle = 0,$$
$$\langle\downarrow|\hat{S}_z|\uparrow\rangle = 0,$$

and

$$\langle\downarrow|\hat{S}_z|\downarrow\rangle = -\frac{\hbar}{2}.$$

This can be written compactly as

$$\langle m'|\hat{S}_z|m\rangle = \mathrm{sign}(m)\delta_{m'm}\frac{\hbar}{2}.$$

Thus in this basis $\hat{S}_z$ is represented by

$$\hat{S}_z \to \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Where we follow the convention that the first row (column) corresponds to the largest value of m' (m) so that the first row (column) is labelled by $m' = 1/2$ $(m = 1/2)$ and the second row (column) is labelled by $m' = -1/2$ $(m = -1/2)$.

As before we define the operators

$$\hat{S}_+ = \hat{S}_x + i\hat{S}_y, \qquad \text{and} \qquad \hat{S}_- = \hat{S}_x - i\hat{S}_y.$$

These have the expected action that

$$\hat{S}_+|\uparrow\rangle = 0, \qquad \hat{S}_+|\downarrow\rangle = \hbar c_+|\uparrow\rangle = \hbar\sqrt{\frac{1}{2}\left(\frac{1}{2}+1\right) + \frac{1}{2}\left(-\frac{1}{2}+1\right)}|\uparrow\rangle = \hbar|\uparrow\rangle$$

$$\hat{S}_-|\downarrow\rangle = 0, \qquad \text{and} \qquad \hat{S}_-|\uparrow\rangle = \hbar c_-|\downarrow\rangle = \hbar\sqrt{\frac{1}{2}\left(\frac{1}{2}+1\right) - \frac{1}{2}\left(\frac{1}{2}+1\right)} = \hbar|\downarrow\rangle.$$

Hence

$$\langle\uparrow|\hat{S}_+|\uparrow\rangle = 0, \qquad \langle\uparrow|\hat{S}_-|\uparrow\rangle = 0,$$
$$\langle\uparrow|\hat{S}_+|\downarrow\rangle = \hbar, \qquad \langle\uparrow|\hat{S}_-|\downarrow\rangle = 0,$$
$$\langle\downarrow|\hat{S}_+|\uparrow\rangle = 0, \qquad \langle\downarrow|\hat{S}_-|\uparrow\rangle = \hbar,$$
$$\langle\downarrow|\hat{S}_+|\downarrow\rangle = 0, \qquad \langle\downarrow|\hat{S}_-|\downarrow\rangle = 0.$$

Thus in this basis these operators are represented by

$$\hat{S}_+ \to \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \hat{S}_- \to \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Now we can write

$$\hat{S}_x = \frac{1}{2}\left(\hat{S}_+ + \hat{S}_-\right), \qquad \text{and} \qquad \hat{S}_y = \frac{1}{2i}\left(\hat{S}_+ - \hat{S}_-\right).$$

Thus the representations of $\hat{S}_x$ and $\hat{S}_y$ in this basis are

$$\hat{S}_x \to \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{and} \qquad \hat{S}_y \to \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

The three matrices representing $\hat{S}_k$ in this basis are called the **Pauli spin matrices**:

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

Which allows us to compactly write

$$\hat{S}_k \to \frac{\hbar}{2}\sigma_k, \qquad \text{and} \qquad \hat{\boldsymbol{S}} = \frac{\hbar}{2}\boldsymbol{\sigma}$$

where $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$. The Pauli matrices are Hermitian, as they are associated with observables, and also unitary, meaning $\sigma_k \sigma_k^\dagger = \sigma_k^2 = \mathbb{1}$, they also all have zero trace, $\operatorname{tr}(\sigma_k) = 0$. We can then use σ as another label for the basis vectors, for example if we define

$$|\psi\rangle = \sum_\sigma \psi_\sigma |\sigma\rangle$$

then we can characterise the action of $\hat{S}_k$ as

$$|\varphi\rangle = \hat{S}_k = \sum_\sigma \varphi_\sigma |\sigma\rangle = \sum_\sigma \sum_{\sigma'} (S_k)_{\sigma\sigma'} \psi_{\sigma'} |\sigma\rangle$$

where we have used

$$\varphi_\sigma = \sum_{\sigma'} (S_k)_{\sigma\sigma'} \psi_{\sigma'}.$$

We can now use the definition of $\hat{S}^2$ to find a representation for $\hat{S}^2$ in this basis:

$$\begin{aligned}
\hat{S}^2 &= \sum_k \hat{S}_k^2 \\
&\to \sum_k \left(\frac{\hbar}{2}\sigma_k\right)^2 \\
&= \sum_k \frac{\hbar^2}{4}\sigma_k^2 \\
&= \sum_k \frac{\hbar^2}{4}\mathbb{1} \\
&= \frac{3}{4}\hbar^2 \mathbb{1} \\
&= \frac{3}{4}\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\end{aligned}$$

14.2   Eigenvectors of $\hat{S}_z$ and $\hat{S}^2$

The eigenvectors of σ_z are easy to find:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \text{and} \qquad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

So the eigenvalues of this representation of $\hat{S}_z$ are given by

$$\frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \text{and} \qquad \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{\hbar}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

We see that these have the eigenvalues $\pm\hbar/2$, which is what we would expect for a spin $1/2$ system. These eigenvectors are also simultaneously eigenvectors of $\hat{S}^2$ in this representation, this is trivial since $\hat{S}^2$ is proportional to the identity in this representation:

$$\frac{3}{4}\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{3}{4}\hbar^2 \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \text{and} \qquad \frac{3}{4}\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{3}{4}\hbar^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Thus we can identify

$$|s = \tfrac{1}{2}, m = \tfrac{1}{2}\rangle = |\uparrow\rangle \to \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \text{and} \qquad |s = \tfrac{1}{2}, m = -\tfrac{1}{2}\rangle = |\downarrow\rangle \to \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

A generic vector, $|\psi\rangle \in \mathcal{H}_{\text{spin}}$, can then be represented by

$$|\psi\rangle = \psi_1|\uparrow\rangle + \psi_2|\downarrow\rangle \to \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \psi_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \psi_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Note that in general $\psi_k \in \mathbb{C}$ so the space spanned by these eigenvectors is not $\mathbb{R}^2$ but $\mathbb{C}^2$.

14.3   Scalar Products

Now that we have representations of generic vectors in this basis we ask what is the representation of a covector, $\langle\psi| \in \mathcal{H}_{\text{spin}}^*$? The answer, as always, is that if $|\psi\rangle$ is represented by an n-dimensional column vector then $\langle\psi|$ is represented by an n-dimensional row vector with the components given by

$$|\psi\rangle \to \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \implies \langle\psi| \to \begin{pmatrix} \psi_1^* & \psi_2^* \end{pmatrix}.$$

That is $\langle\psi| = (|\psi\rangle)^\dagger$. The scalar product of two states, $|\psi\rangle$ and $|\varphi\rangle$, is then

$$\langle\varphi|\psi\rangle = \begin{pmatrix} \varphi_1^* & \varphi_2^* \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \varphi_1^*\psi_1 + \varphi_2^*\psi_2.$$

For example if $|\psi\rangle$ is a properly normalised state then

$$\langle\varphi|\psi\rangle = \begin{pmatrix} \psi_1^* & \psi_2^* \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = |\psi_1|^2 + |\varphi_2|^2 = 1.$$

14.4   Spin Along the x Direction

We can find the possible values of m_x by finding the eigenvalues of $\hbar\sigma_x/2$ in the $\hat{S}_z$ eigenbasis. We do this in the normal way:

$$\begin{aligned}
0 &= \det\left(\frac{\hbar}{2}\sigma_x - \lambda\mathbb{1}\right) \\
&= \begin{vmatrix} -\lambda & \hbar/2 \\ \hbar/2 & -\lambda \end{vmatrix} \\
&= \lambda^2 - \frac{\hbar^2}{4}
\end{aligned}$$

So $S_x = \lambda = \pm\hbar/2$. This is the same as the possible values of S_z, this is what we would expect from the symmetry of the situation, there is nothing important about the z or x directions that would mean they have different possible values of S_k. We can write a state with $m_x = 1/2$ as a combination of the basis vectors:

$$|m_x = 1/2\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle.$$

In this representation this becomes

$$\hat{S}_x|m_x = 1/2\rangle = \frac{\hbar}{2}|m_x = 1/2\rangle \to \frac{\hbar}{2}\sigma_x \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} \beta \\ \alpha \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

Hence $\alpha = \beta$ so a properly normalised state with $m_x = \hbar/2$ is

$$|m_x = 1/2\rangle = \frac{\sqrt{2}}{2}\left[|\uparrow\rangle + |\downarrow\rangle\right].$$

We can show in a similar way that if $S_x = -1/2$ then

$$|m_x = -1/2\rangle = \frac{\sqrt{2}}{2}\left[|\uparrow\rangle - |\downarrow\rangle\right].$$

It can be shown that if we measure the spin along any vector in the (x, z)-plane at an angle ϑ to the z axis,

$$\hat{\boldsymbol{n}} = \sin(\vartheta)\boldsymbol{e_x} + \cos(\vartheta)\boldsymbol{e_z},$$

then the relevant eigenstates in the the representation that we are using are the eigenvalues of

$$\boldsymbol{\sigma} \cdot \hat{\boldsymbol{n}} = \sigma_x \sin(\vartheta) + \sigma_z \cos(\vartheta) = \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ \sin\vartheta & -\cos\vartheta \end{pmatrix}.$$

15   Addition of Angular Momenta

Suppose we have two systems with independent angular momenta $\hat{\boldsymbol{J}}^{(1)}$ and $\hat{\boldsymbol{J}}^{(2)}$. This could be two particles with angular momentum, two particles with spin or one particle with angular momentum and spin. These operators satisfy the usual commutation relations,

$$\left[\hat{J}_k^{(n)}, \hat{J}_l^{(n)}\right] = i\hbar\varepsilon_{klm}\hat{J}_m^{(n)}, \qquad \text{and} \qquad \left[\hat{J}^{(n)2}, \hat{J}_k^{(n)}\right] = 0.$$

Since the two angular momenta are independent the operators associated with different angular momenta commute:

$$\left[\hat{J}_k^{(1)}, \hat{J}_l^{(2)}\right] = \left[\hat{J}^{(1)2}, \hat{J}^{(2)2}\right] = \left[\hat{J}^{(1)2}, \hat{J}_k^{(2)}\right] = 0 \qquad \text{etc.}$$

We can extend this easily to any number of independent angular momenta but we won't here.

15.1   Vector Spaces

Since $\hat{\boldsymbol{J}}^{(1)}$ and $\hat{\boldsymbol{J}}^{(2)}$ commute they form a CSCO which means we can find a common eigenbasis. Since both angular momenta are independent we can describe both with separate Hilbert spaces. Angular momentum 1 is described by vectors in a vector space $\mathcal{H}^{(1)}$. For a particular value, $j^{(1)}$, of the angular momentum quantum number this is a $2j_1 + 1$ dimensional vector space. This space has as a basis $\{|j_1, m_1\rangle\}$ Similarly angular momentum 2 is described by vectors in the $2j_2 + 1$ dimensional vector space, $\mathcal{H}^{(2)}$, with basis $|j_2, m_2\rangle$. The entire system is then described by vectors in

$$\mathcal{H} = \mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)}.$$

The dimensionality of this vector space is

$$\dim(\mathcal{H}) = \dim(\mathcal{H}^{(1)})\dim(\mathcal{H}^{(2)}) = (2j_1 + 1)(2j_2 + 1).$$

This space then has as a basis

$$\{|j_1, m_1, j_2, m_2\rangle\} = \{|j_1, m_1\rangle \otimes |j_2, m_2\rangle\}.$$

As usual we extend the definition of the operators so that they act on only their respective parts of these vectors. Strictly we should define new operators like $\hat{J}_z^{(1)\prime} = \hat{J}_z^{(1)} \otimes \mathbb{1}$ which then acts on $|j_1, m_1, j_2, m_2\rangle = |j_1, m_1\rangle \otimes |j_2, m_2\rangle$ as

$$\begin{aligned}
\hat{J}_z^{(1)\prime}|j_1, m_1, j_2, m_2\rangle &= (\hat{J}_z^{(1)} \otimes \mathbb{1})(|j_1, m_1\rangle \otimes |j_2, m_2\rangle) \\
&= \hat{J}_z^{(1)}|j_1, m_1\rangle \otimes \mathbb{1}|j_2, m_2\rangle \\
&= \hbar m_1|j_1, m_1\rangle \otimes |j_2, m_2\rangle \\
&= \hbar m_1|j_1, m_1, j_2, m_2\rangle
\end{aligned}$$

but in practice we use the same symbol for both operators, $\hat{J}_z^{(1)}\colon \mathcal{H}^{(1)} \to \mathbb{C}$ and $\hat{J}_z^{(1)\prime}\colon \mathcal{H} \to \mathbb{C}$. The action of the operators on this basis is

$$\hat{J}^{(1)2}|j_1, m_1, j_2, m_2\rangle = \hbar^2 j_1(j_1 + 1)|j_1, m_1, j_2, m_2\rangle,$$

$$\hat{J}_z^{(1)}|j_1, m_1, j_2, m_2\rangle = \hbar m_1|j_1, m_1, j_2, m_2\rangle,$$

$$\hat{J}^{(2)2}|j_1, m_1, j_2, m_2\rangle = \hbar^2 j_2(j_2 + 1)|j_1, m_1, j_2, m_2\rangle,$$

$$\hat{J}_z^{(2)}|j_1, m_1, j_2, m_2\rangle = \hbar m_2|j_1, m_1, j_2, m_2\rangle.$$

A generic state, $|\psi\rangle \in \mathcal{H}$, can be given as

$$|\psi\rangle = \sum_{j_1,m_2,j_2,m_2} c_{j_1 m_1 j_2 m_2} |j_1, m_1, j_2, m_2\rangle$$

where

$$c_{j_1 m_1 j_2 m_2} = \langle j_1, m_1, j_2, m_2 | \psi \rangle.$$

15.2   Total Angular Momentum

We can now define the total angular momentum:

$$\hat{\boldsymbol{J}} = \hat{\boldsymbol{J}}^{(1)} + \hat{\boldsymbol{J}}^{(2)}.$$

This satisfies the normal angular momentum commutation relations:

$$[\hat{J}_k, \hat{J}_l] = i\hbar\varepsilon_{klm}\hat{J}_m, \qquad \text{and} \qquad [\hat{J}^2, \hat{J}_k] = 0.$$

This means that $\hat{J}^2$ and $\hat{J}_z$ form a CSCO so we can find a simultaneous eigenbasis for $\mathcal{H}$ made of vectors of the form $\{|j, m\rangle\}$. What values can j and m take? The answer lies in the angular momentum addition theorem, which we won't prove here:

Theorem 6: Angular momentum addition theorem

Given two independent angular momenta with corresponding quantum numbers, j_1 and j_2, the allowed value of the angular momentum quantum number for the total angular momentum are

$$j = j_1 + j_2, j_1 + j_2 - 1, \ldots, |j_1 - j_2|.$$

For each of these values we also have

$$m = -j, -j + 1, \ldots, j - 1, j.$$

Before we said that $\dim(\mathcal{H}) = (2j_1 + 1)(2j_2 + 1)$. We can check that this is consistent with the angular momentum addition theorem. Without loss of generality assume that $j_1 \geq j_2$. Then

$$\begin{aligned}
\dim(\mathcal{H}) &= \sum_{j=j_1-j_2}^{j_1+j_2}(2j + 1) \\
&= \sum_{j=0}^{j_1+j_2}(2j + 1) - \sum_{j=0}^{j_1-j_2-1}(2j + 1) \\
&= 2\sum_{j=0}^{j_1+j_2}j + (j_1 + j_2) - 2\sum_{j=0}^{j_1-j_2-1}j - (j_1 - j_2 - 1) \\
&= (j_1 + j_2)(j_1 + j_2 + 1) + (j_1 + j_2) - (j_1 - j_2 - 1)(j_1 - j_2) - (j_1 - j_2 - 1) \\
&= j_1^2 + j_2^2 + 2j_1 j_2 + j_1 + j_2 + j_1 + j_2 - j_1^2 - j_2^2 + 2j_1 j_2 + j_1 - j_2 - j_1 + j_2 + 1 \\
&= 4j_1 j_2 + 2j_1 + 2j_2 + 1 \\
&= (2j_1 + 1)(2j_2 + 1).
\end{aligned}$$

Where we have used

$$\sum_{r=0}^{n} = \frac{1}{2}n(n+1),$$

which is proven in appendix B. Both bases, $\{|j_1, m_1, j_2, m_2\rangle\}$ and $\{|j, m\rangle\}$, are orthonormal so there exists a unitary transformation between the two bases. For fixed j_1 and j_2 we denote $|j, m\rangle$ as $|j, m; j_1, j_2\rangle$ to denote the specific values that j_1 and j_2 are held at. Using the completeness of the $\{|j_1, m_1, j_2, m_2\rangle\}$ basis,

$$\sum_{m_1, m_2} \langle j_1, m_1, j_2, m_2 | j_1, m_1, j_2, m_2 \rangle = \mathbb{1}$$

we have

$$|j, m; j_1, j_2\rangle = \sum_{m_1, m_2} |j_1, m_1, j_2, m_2\rangle\langle j_1, m_1, j_2, m_2|j, m; j_1, j_2\rangle.$$

The quantities $\langle j_1, m_1, j_2, m_2|j, m; j_1, j_2\rangle$ are called the Clebsch–Gordan coefficients and they are tabulated in many sources.

The basis $\{|j_1, m_1, j_2, m_2\rangle\}$ is called the **uncoupled basis** as vectors in this basis can be written as two parts, one related to angular momentum 1 and the other related to angular momentum 2, for example a basis vector can be written as $|j_1, m_1\rangle \otimes |j_2, m_2\rangle$ and a generic vector as

$$|\psi\rangle = \sum_{j_1, m_1, j_2, m_2} (\langle j_1, m_1|\psi\rangle|j_1, m_1\rangle \otimes \langle j_2, m_2|\psi\rangle|j_2, m_2\rangle).$$

The basis $\{|j, m\rangle\}$ is called the **coupled basis** as this is not possible.

Example 15.1. Consider a system composed of two spin 1/2 particles. $\mathcal{H}^{(1)}$ is a two dimensional vector space with a basis

$$\{s_1, m_1\} = \{|\tfrac{1}{2}, \tfrac{1}{2}\rangle, |\tfrac{1}{2}, -\tfrac{1}{2}\rangle\}.$$

$\mathcal{H}^{(2)}$ is also a two dimensional vector space with basis

$$\{s_2, m_2\} = \{|\tfrac{1}{2}, \tfrac{1}{2}\rangle, |\tfrac{1}{2}, -\tfrac{1}{2}\rangle\}.$$

The spin of the whole system is then characterised by vectors in $\mathcal{H} = \mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)}$ which has the basis

$$\{|s_1, m_1, s_2, m_2\rangle\} = \{|\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}\rangle, |\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}\rangle, |\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}\rangle, |\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}\rangle\}.$$

Since $s_i = 1/2$ for all states in this system we compactly write this basis as

$$\{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}.$$

Where the first arrow denotes particle 1's spin and the second arrow denotes particle 2's spin. The values that s can take, as given by the angular momentum addition theorem, is $s = s_1 + s_2, |s_1 - s_2| = 1, 0$. So the entire system has spin 1 or spin 0. To find the uncoupled basis consider the action of $\hat{S}_z$ on the coupled basis

$$\hat{S}_z|\uparrow\uparrow\rangle = (\hat{S}_z^{(1)} + \hat{S}_z^{(2)})|\uparrow\uparrow\rangle = \frac{\hbar}{2}|\uparrow\uparrow\rangle + \frac{\hbar}{2}|\uparrow\uparrow\rangle = \hbar|\uparrow\uparrow\rangle.$$

Hence $m = 1$. Since $m > 0$ we know that $s \neq 0$ so $s = 1$. Similarly

$$\hat{S}_z|\downarrow\downarrow\rangle = (\hat{S}_z^{(1)} + \hat{S}_z^{(2)})|\downarrow\downarrow\rangle = -\frac{\hbar}{2}|\downarrow\downarrow\rangle + -\frac{\hbar}{2}|\downarrow\downarrow\rangle = -\hbar|\downarrow\downarrow\rangle.$$

Hence $m = 1$ and $s = 1$ for this state. So far we have

$$|1, 1\rangle = |\uparrow\uparrow\rangle, \qquad \text{and} \qquad |1, -1\rangle = |\downarrow\downarrow\rangle.$$

We don't yet know what $|1, 0\rangle$ or $|0, 0\rangle$ are. The difficulty in finding these states is that they aren't uniquely identified purely by the value of m. However we do know how to construct them from other states using raising and lowering operators. The lowering operator for the total angular momentum is simply

$$\hat{S}_- = \hat{S}_-^{(1)} + \hat{S}_-^{(2)}$$

so by identifying that we expect

$$\hat{S}_-|1, 1\rangle \propto |1, 0\rangle$$

we know that

$$|1, 0\rangle \propto \hat{S}_-|\uparrow\uparrow\rangle = (\hat{S}_-^{(1)} + \hat{S}_-^{(2)})|\uparrow\uparrow\rangle = \hat{S}_-^{(1)}|\uparrow\uparrow\rangle + \hat{S}_-^{(2)}|\uparrow\uparrow\rangle \propto |\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle.$$

Using the fact that the uncoupled basis vectors are orthonormal we have

$$|1, 0\rangle = \frac{\sqrt{2}}{2}(|\uparrow\downarrow\rangle + |\uparrow\downarrow\rangle).$$

Finally $|0, 0\rangle$ must be a linear combination of $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$ and it must be orthogonal to $|1, 0\rangle$. We are then free to chose that

$$|0, 0\rangle = \frac{\sqrt{2}}{2}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle).$$

This agrees with the Clebsch–Gordan coefficients which are tabulated in table 15.1.

---

m_1, m_2 \ j	1
$1/2, 1/2$	1

(a) $m = 1$

m_1, m_2 \ j	1
$-1/2, -1/2$	1

(b) $m = -1$

m_1, m_2 \ j	1	0
$1/2, -1/2$	$\sqrt{2}/2$	$\sqrt{2}/2$
$-1/2, 1/2$	$\sqrt{2}/2$	$-\sqrt{2}/2$

(c) $m = 0$

Table 15.1: Clebsch–Gordan coefficients for two spin $1/2$ particles.

16   Identical Particles

In classical mechanics if we have two particles that have all the same properties, like mass, charge, angular momentum, etc. then we can still tell them apart from the fact that one of them is 'here' and the other is 'there'. Even at some later time we can still tell them apart as we can track their trajectories back to a time when we could tell them apart. In quantum mechanics (QM) this is not possible. If two particles have the same properties, meaning identical quantum numbers, n, ℓ, s, m, etc. then we label one 'particle 1' and the other 'particle 2' we lose this information as soon as we ascribe it as the system evolves and since there is no concept of trajectory we can't tell which particle was which at some earlier time.

To think about this mathematically we start with the states, $|\xi_1\rangle, |\xi_2\rangle \in \mathcal{H}$, which describe each particle. The two particle system is then described by $|\xi_1, \xi_2\rangle = |\xi_1\rangle \otimes |\xi_2\rangle \in \mathcal{H} \otimes \mathcal{H}$.

Here particle 1 and 2 are labels that we assign that distinguish the particles for the sake of the maths. There is no way to look at a particle and tell if it is particle 1 or 2. We define an operator, $\hat{\mathscr{P}}_{12}$ which swaps particle 1 and 2. That is

$$\hat{\mathscr{P}}_{12}|\xi_1, \xi_2\rangle = |\xi_2, \xi_1\rangle.$$

Since quantum states are defined up to a phase factor and swapping the two particles doesn't change the state of the system we must have that

$$|\xi_1, \xi_2\rangle = e^{i\alpha}|\xi_2, \xi_1\rangle$$

for some $\alpha \in \mathbb{R}$. By swapping the two states again we get back to the original state and we pick up another phase factor:

$$|\xi_1, \xi_2\rangle = e^{2i\alpha}|\xi_1, \xi_2\rangle.$$

This can only be true if $e^{2i\alpha} = 1$ which means that $e^{i\alpha} = \pm 1$, thus $\alpha = n\pi$ for $n \in \mathbb{Z}$. This means that this state is either symmetric or antisymmetric under swapping the particles. It turns out that there is a theorem that allows us to be even more precise than this, which we will not prove here. It is called the spin statistics theorem:

> **Theorem 7: Spin Statistics Theorem**
>
> Particles with integer spin, $s = 0, 1, 2, \ldots$, are symmetric under swapping identical particles:
>
> $$|\xi_1, \xi_2\rangle = |\xi_2, \xi_1\rangle.$$
>
> Such particles are called **bosons** and are described by Bose–Einstein statistics.
> Particles with half integer spin, $s = 1/2, 3/2, 5/2, \ldots$, are antisymmetric under swapping identical particles:
>
> $$|\xi_1, \xi_2\rangle = -|\xi_2, \xi_1\rangle.$$
>
> Such particles are called **fermions** and are described by Fermi–Dirac statistics.

16.1   Helium Atom

A helium atom, He, has two electrons. These electrons are identical particles. There is also a proton. The interaction of one electron and the proton has a Hamiltonian given by

$$\hat{H}_i = \frac{\hat{\boldsymbol{P}} \cdot \hat{\boldsymbol{P}}}{2\mu} - \frac{2e^2}{r_i}$$

where $i = 1, 2$ denotes the particle number, μ is the mass of the electron, e is the magnitude of the charge of the electron and proton, and r_i denotes the distance of electron i from the proton. The entire system can then be described as the Hamiltonian for each electron plus a term characterising the interaction of the two electrons:

$$\hat{H} = \hat{H}_1 + \hat{H}_2 + \frac{e^2}{|\boldsymbol{r_1} - \boldsymbol{r_2}|}.$$

Notice that $\hat{H}$ is symmetric under $\hat{\mathscr{P}}_{12}$. That is $\hat{H}(1,2) = \hat{H}(2,1)$ where $\hat{H}(1,2)$ is $\hat{H}$ as defined above and $\hat{H}(2,1) = \hat{\mathscr{P}}_{12}\hat{H}(1,2)$. The TISE for this system is

$$\hat{H}(1,2)|\xi_1, \xi_2\rangle = E|\xi_1, \xi_2\rangle,$$

where we now require that $|\xi_1, \xi_2\rangle$ is an energy eigenstate. We now rename 1 to 2 and 2 to 1, this changes nothing other than labels:

$$\hat{H}(2,1)|\xi_2, \xi_1\rangle = E|\xi_2, \xi_1\rangle.$$

Finally using the fact that $\hat{H}(1,2) = \hat{H}(2,1)$ we have that

$$\hat{H}(1,2)|\xi_2, \xi_2\rangle = E|\xi_2, \xi_2\rangle.$$

Thus any given eigenstate, $|\xi_1, \xi_2\rangle$, of $\hat{H}(1,2)$, gives us another eigenstate, $|\xi_2, \xi_1\rangle$. We can define a symmetric and antisymmetric state, $|+\rangle$ and $|-\rangle$ respectively, in the usual way:

$$|\psi\pm\rangle = \frac{\sqrt{2}}{2}(|\xi_1, \xi_2\rangle \pm |\xi_2, \xi_1\rangle).$$

The action of $\hat{\mathscr{P}}_{12}$ on these is

$$\hat{\mathscr{P}}_{12}|\pm\rangle = \pm|\pm\rangle.$$

So $|\pm\rangle$ are eigenstates of $\hat{\mathscr{P}}_{12}$, with eigenvalues ± 1 and $\hat{H}$, with eigenvalues E. Hence $[\hat{H}, \hat{\mathscr{P}}_{12}] = 0$. This means that the symmetry properties of the system are preserved in time since this means that $\hat{\mathscr{P}}_{12}$ commutes with the time evolution operator.

16.2   Two Electron State

Consider now a system made of two electrons. There are sets of degrees of freedom to consider. The spatial degrees of freedom, described by vectors in $\mathcal{H}_{\text{space}}$, and the spin degrees of freedom, described by vectors in $\mathcal{H}_{\text{spin}}$. Both electrons have $s_i = 1/2$ so by the angular momenta addition theorem the total spin of the system is $s = 0, 1$. In the case where $s = 0$ we must also have $m = 0$ so $|0, 0\rangle$ is the only state with $s = 0$, we call this a singlet state. In the case where $s = 1$ we have $m = -1, 0, 1$ so there are three states, $|1, -1\rangle$, $|1, 0\rangle$, and $|1, 1\rangle$, with $s = 1$, we call this a triplet state. We have previously shown that in the uncoupled basis these states are

$$|0, 0\rangle = \frac{\sqrt{2}}{2}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle), \qquad |1, 1\rangle = |\uparrow\uparrow\rangle, \qquad |1, -1\rangle = |\downarrow\downarrow\rangle, \qquad \text{and} \qquad |1, 0\rangle = \frac{\sqrt{2}}{2}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle).$$

From this we can see that $|0, 0\rangle$ is antisymmetric under swapping of the two particles and all the other states are symmetric. Let $|\psi_{12}\rangle \mathcal{H}_{\text{space}}$ be a generic state describing the spatial degrees of freedom. The a generic state, $|\Psi\rangle \in \mathcal{H} = \mathcal{H}_{\text{space}} \otimes \mathcal{H}_{\text{spin}}$ with $s = 1$ is given by

$$|\Psi\rangle = |\psi_{12}\rangle \otimes |s = 1, m\rangle.$$

Using the fact that $s = 1$ implies that $|s = 1, m\rangle$ is symmetric under $\hat{\mathscr{P}}_{12}$ the action of swapping the two particles is given by

$$\begin{aligned}
\hat{\mathscr{P}}_{12}|\Psi\rangle &= (\hat{\mathscr{P}}_{12}|\psi_{12}\rangle) \otimes (\hat{\mathscr{P}}_{12}|s = 1, m\rangle) \\
&= (\hat{\mathscr{P}}_{12}|\psi_{12}\rangle) \otimes |s = 1, m\rangle \\
&= -|\Psi\rangle.
\end{aligned}$$

Here we have used that electrons are spin $1/2$ particles so by the spin statistics theorem the overall state describing them is antisymmetric. The only way to have this be true is to have $|\psi_{12}\rangle$ be antisymmetric under $\hat{\mathscr{P}}_{12}$. Similarly a generic state with $s = 0$ is described by

$$|\Psi\rangle = |\psi_{12}\rangle \otimes |s = 0, m = 0\rangle.$$

Again this is antisymmetric under $\hat{\mathscr{P}}_{12}$ but this time $|s = 0, m = 0\rangle$ is antisymmetric under $\hat{\mathscr{P}}_{12}$ so $|\psi_{12}\rangle$ must be symmetric under $\hat{\mathscr{P}}_{12}$. These two results are tabulated in table 16.1.

s	$\lvert s, m \rangle$	$\lvert \psi_{12} \rangle$	$\lvert \Psi \rangle$
0	A	S	A
1	S	A	A

Table 16.1: The symmetry of various parts of a state under $\hat{\mathscr{P}}_{12}$. S denotes symmetric and A denotes antisymmetric. The symmetry of the total system, $\lvert \Psi \rangle$, is set by the spin statistics theorem. The symmetry of the spin part, $\lvert s, m \rangle$, follows from the expression in the uncoupled basis. The symmetry of the spatial part, $\lvert \psi_{12} \rangle$, is then fixed to make the symmetries of the other parts match.

16.3   Pauli Exclusion Principle

One consequence of the spin statistics theorem is the **Pauli exclusion principle**, that no two fermions can be in the same state. The reason for this is that if particles 1 and 2 are identical fermions and $\lvert \xi_1 \rangle = \lvert \xi_2 \rangle$ then by the spin statistics theorem

$$\lvert \xi_1, \xi_2 \rangle = -\lvert \xi_2, \xi_1 \rangle = -\lvert \xi_1, \xi_2 \rangle$$

where for the first equality we use the spin statistics theorem and for the second the fact that $\lvert \xi_1 \rangle = \lvert \xi_2 \rangle$. The only solution to this equation is $\lvert \xi_1, \xi_2 \rangle = 0$ but this is not a valid state, it isn't normalisable for one. Therefore we cannot have $\lvert \xi_1 \rangle = \lvert \xi_2 \rangle$ if the spin statistics theorem holds.

17   The Hydrogen Atom

In this section we will give a non-relativistic treatment of the hydrogen atom. We model the hydrogen atom as an electron and a proton interacting due to a Coulomb potential. The important constants are the mass of a proton, m_p, the mass of the electron, m_e, and the elementary charge, q, which is the charge of the proton and the negative of the charge of the electron. The values of these are

$$m_p = 1.7 \times 10^{-27} \, \text{kg}, \qquad m_e = 9.1 \times 10^{-31} \, \text{kg}, \qquad q = 1.6 \times 10^{-19} \, \text{C}, \qquad \text{and} \qquad \frac{m_p}{m_e} = 1836.15.$$

As with two body systems in classical physics it is often useful to work with the reduced mass, μ, defined by

$$\frac{1}{\mu} = \frac{1}{m_p} + \frac{1}{m_e} = \frac{m_e + m_p}{m_p m_e}.$$

The Coulomb potential is

$$V(r) = -\frac{q^2}{4\pi\varepsilon_0 r} = -\frac{e^2}{r} \qquad \text{where} \qquad e^2 = \frac{q^2}{4\pi\varepsilon_0}.$$

Here r is the distance between the proton and the electron, that is we choose the position of the proton as the origin. We can view e as a new variable tidying away some constant prefactor or as the charge in Gaussian units. The Hamiltonian for the system is then

$$\hat{H} = \frac{\hat{\boldsymbol{P}} \cdot \hat{\boldsymbol{P}}}{2\mu} + V(r).$$

17.1   Stationary States

As usual we look for the eigenstates of the Hamiltonian. To do this we write the Hamiltonian as a differential operator:

$$\hat{H} = -\frac{\hbar^2}{2\mu}\nabla^2 - \frac{e^2}{r}$$

and then solve the TISE:

$$\left[-\frac{\hbar^2}{2\mu} - \frac{e^2}{r} \right] \psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}).$$

Since V is a central potential we work in spherical coordinates and we look for a solution of the form

$$\psi(r, \vartheta, \varphi) = R(r)Y_\ell^m(\vartheta, \varphi).$$

As we did in section 13.2 we note that the Laplacian can be written as

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{1}{\hbar^2 r^2}\hat{L}^2.$$

We define $\chi_\ell(r) = rR_\ell(r)$ and from equation 13.1 we have

$$\left[-\frac{\hbar^2}{2\mu}\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{\hbar^2\ell(\ell+1)}{2\mu r^2} - \frac{e^2}{r}\right]E\chi_\ell(r) = E_\ell\chi_\ell(r).$$

Figure 17.1 shows the potential. For $r \to 0$ the potential goes as r^{-2} and for $r \to \infty$ the potential goes as r^{-1}. We are looking for states where the electron is bound to the proton so the total energy, E_ℓ, must be less than zero.
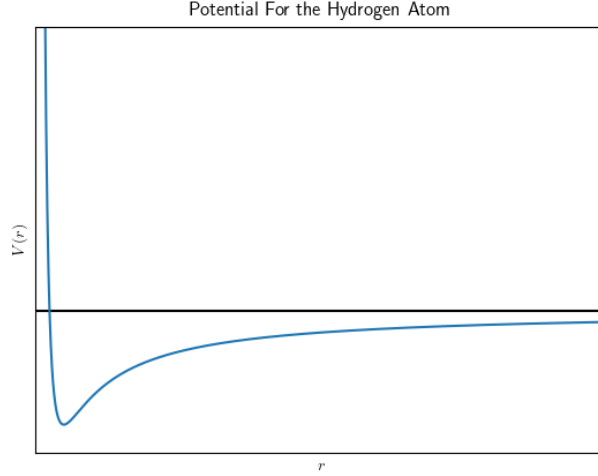
Figure 17.1: The potential for the hydrogen atom.

To tidy away even more constants we define

$$a_0 = \frac{\hbar^2}{\mu e^2} \approx 0.52\,\text{Å}, \qquad \text{and} \qquad E_I = \frac{\hbar^2}{2\mu a_0^2} = \frac{\mu e^4}{2\hbar^2} \approx 13.6\,\text{eV}.$$

a_0 is called the Bohr radius. We will see the physical significance of these values later. We now use these quantities to rescale the physical properties of the system:

$$\rho = \frac{r}{a_0}, \qquad \text{and} \qquad \lambda_\ell = \sqrt{-\frac{E_\ell}{E_I}}.$$

Note that since $E_\ell < 0$ for a bound electron $\lambda \in \mathbb{R}$. We also define $u_\ell(\rho) = \chi_\ell(\rho)$. Using this we have

$$\begin{aligned}
-\lambda_\ell^2 E_I\chi(a_0\rho) &= -\lambda_\ell^2 E_I u_\ell(\rho) \\
&= \left[-\frac{\hbar^2}{2\mu}\frac{1}{a_0^2}\frac{\mathrm{d}^2}{\mathrm{d}\rho^2} + \frac{\hbar^2\ell(\ell+1)}{2\mu a_0^2\rho^2} - \frac{e^2}{a_0\rho}\right]\chi_\ell(a_0\rho) \\
&= \left[-\frac{\hbar^2}{2\mu}\frac{1}{a_0^2}\frac{\mathrm{d}^2}{\mathrm{d}\rho^2} + \frac{\hbar^2\ell(\ell+1)}{2\mu a_0^2\rho^2} - \frac{e^2}{a_0\rho}\right]u_\ell(\rho).
\end{aligned}$$

Which gives us

$$\left[-\frac{\hbar^2}{2\mu}\frac{1}{a_0^2}\frac{\mathrm{d}^2}{\mathrm{d}\rho^2} + \frac{\hbar^2\ell(\ell+1)}{2\mu a_0^2\rho^2} - \frac{e^2}{a_0\rho} + \lambda_\ell^2 E_I\right]u_\ell(\rho) = 0.$$

Dividing through by $-\hbar^2/2\mu a_0^2$ we have

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}\rho^2} - \frac{\ell(\ell+1)}{\rho^2} + \frac{2\mu a_0 e^2}{\hbar^2\rho} - \lambda_\ell^2\frac{2\mu a_0^2 E_I}{\hbar^2}\right]u_\ell(\rho) = 0.$$

Using the definitions of a_0 and E_I this becomes

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}\rho^2} - \frac{\ell(\ell+1)}{\rho^2} + \frac{2}{\rho} - \lambda_\ell^2\right]u_\ell(\rho) = 0.$$

17.1.1   Solution to the Radial Equation

To motivate an ansatz we will look at the limiting behaviour of the radial equation. First consider the case that $\rho \to \infty$, in this case the radial equation reduces to

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}\rho^2} - \lambda^2\right] u_\ell(\rho) = 0.$$

The solutions to this are

$$u_\ell(\rho) = e^{\pm\lambda_\ell\rho}.$$

We discard the exponential growth solution as it is non-normalisable so we are left only with the exponential decay solution. It makes sense for the radial solution to decay exponentially as it becomes increasingly unlikely for the electron to be found as we move away from the proton. We make the ansatz that

$$u_\ell(\rho) = e^{-\lambda_\ell\rho}\eta_\ell(\rho).$$

Substituting this into the radial equation we have

$$
\begin{aligned}
0 &= \frac{\mathrm{d}^2}{\mathrm{d}\rho^2}[e^{-\lambda_\ell\rho}\eta_\ell(\rho)] + \left[\frac{2}{\rho} - \frac{\ell(\ell+1)}{\rho^2} - \lambda_\ell^2\right]e^{-\lambda_\ell\rho}\eta_\ell(\rho) \\
&= \frac{\mathrm{d}^2\eta_\ell}{\mathrm{d}\rho^2}e^{-\lambda_\ell\rho} + 2\frac{\mathrm{d}\eta_\ell}{\mathrm{d}\rho}\frac{\mathrm{d}}{\mathrm{d}\rho}e^{-\lambda_\ell\rho} + \eta_\ell(\rho)\frac{\mathrm{d}^2}{\mathrm{d}\rho^2}e^{-\lambda_\ell\rho} + \left[\frac{2}{\rho} - \frac{\ell(\ell+1)}{\rho^2} - \lambda_\ell^2\right]e^{-\lambda_\ell\rho}\eta_\ell(\rho) \\
&= \frac{\mathrm{d}^2\eta_\ell}{\mathrm{d}\rho^2}e^{-\lambda_\ell\rho} - 2\lambda_\ell 2\frac{\mathrm{d}\eta_\ell}{\mathrm{d}\rho}e^{-\lambda_\ell} + \lambda_\ell^2\eta_\ell(\rho)e^{-\lambda_\ell\rho} + + \left[\frac{2}{\rho} - \frac{\ell(\ell+1)}{\rho^2} - \lambda_\ell^2\right]e^{-\lambda_\ell\rho}\eta_\ell(\rho).
\end{aligned}
$$

Dividing through by $e^{-\lambda_\ell\rho}$ we have

$$\frac{\mathrm{d}^2\eta_\ell}{\mathrm{d}\rho^2} - 2\lambda_\ell\frac{\mathrm{d}\eta_\ell}{\mathrm{d}\rho} + \left[\frac{2}{\rho} - \frac{\ell(\ell+1)}{\rho^2}\right]\eta_\ell(\rho) = 0. \tag{17.1}$$

Now as $\rho \to 0$ we know from section 13.4 that $u_\ell(\rho) \sim \rho^{\ell+1}$. We expand η_ℓ as a power series in ρ as

$$\eta_\ell(\rho) = \rho^{\ell+1}\sum_{q=0}^{\infty}c_q\rho^q.$$

Inserting this into equation 17.1 we have

$$
\begin{aligned}
0 &= \sum_q\left[(q+\ell+1)(q+\ell)c_q\rho^{q+\ell-1} - 2\lambda_\ell(q+\ell+1)c_q\rho^{\ell+1} + 2c_q\rho^{q+\ell} - \ell(\ell+1)c_q\rho^{q+\ell-1}\right] \\
&= \sum_q q(q+2\ell+1)c_q\rho^{q+\ell-1} - \sum_q 2(\lambda_\ell[q+\ell+1]-1)c_q\rho^{q+\ell}
\end{aligned}
$$

relabelling, $q \to q-1$ in the second sum we have

$$= \sum_q\left[q(q+2\ell+1)c_q - 2[\lambda_\ell(q+\ell)-1]c_{q-1}\right]\rho^{q+\ell-1}.$$

Since we require this to hold for all ρ we have that for all q

$$q(q+2\ell+1)c_q - 2[\lambda_\ell(q+\ell)-1]c_{q-1} = 0. \tag{17.2}$$

This gives us a recursion relation between the coefficients of the Taylor expansion of $\eta_\ell(\rho)/\rho^{\ell+1}$. For large q we have

$$\frac{c_q}{c_{q-1}} = \frac{2\lambda_\ell(q+1)-2}{q(q+2\ell+1)} \to \frac{2\lambda}{q}$$

as $q \to \infty$ we have

$$c_q \sim c_{q-1}\frac{2\lambda}{q} \sim \frac{(2\lambda_\ell)^q}{q!}.$$

This is the asymptotic behaviour that we expect to give exponential growth in a Taylor series so we expect $\eta_\ell(\rho) \sim \rho^{\ell+1}e^{2\lambda\rho}$. However this is not normalisable. The only way for this to be normalisable is if for some $q = n_r$ we have $c_{n_r} = 0$ as this means that $c_q = 0$ for all $q \geq n_r$. This means that η_ℓ is a polynomial of order n_r. Setting $q = n_r$ in equation 17.2 we have that

$$2[\lambda_{n\ell}(n_r + 1) - 1]c_{n_r-1} = 0 \implies \lambda_{n\ell} = \frac{1}{n_r + \ell} = \frac{1}{n}.$$

Here we introduce n, the **principle quantum number**. We see that $\lambda_{n\ell}$ is quantised as a result of requiring the wave function be normalisable.

17.1.2   Physical Interpretation

Since $\lambda_{n\ell}$ is quantised the energy is also quantised as

$$E_{n\ell} = -\frac{E_I}{(n_r + \ell)^2} = -\frac{E_I}{n^2}.$$

We see that E_I is the energy needed to remove an electron from the ground state, $n = 1$. That is E_I is the ionisation energy of hydrogen. Going back to the definition of E_I we can write

$$E_I = \frac{\mu e^4}{2\hbar^2} = \frac{1}{2}\alpha^2\mu c^2$$

where

$$\alpha = \frac{e^2}{\hbar c^2} = \frac{q^2}{4\pi\varepsilon_0\hbar c} \approx \frac{1}{137}$$

is the **fine-structure constant**. Notice that $\mu \approx m_e$ and therefore μc^2 is approximately the rest energy of the electron. The energy levels are on the scale of $\alpha^2\mu c^2/2 \approx \mu c^2 2.7 \times 10^{-5}$. Since the energy levels are on a scale so much smaller than the rest mass of the electron (which is in turn much less than the rest energy of the proton) we are justified in using a non-relativistic approach. Any relativistic corrections will typically be $\mathcal{O}\alpha$. Further justification is provided by Heisenberg's uncertainty principle. Taking a_0 to be a typical length scale in hydrogen we have that $\Delta p \approx a_0\hbar/2 = \mu e^2/\hbar$. We can also use the symmetry to state that positive and negative momenta are equally likely so $\Delta p = \sqrt{\langle p^2 \rangle - \langle p \rangle^2} = \sqrt{\langle p^2 \rangle}$ from which we have

$$v \sim \frac{p}{\mu} \sim \frac{e^2}{\hbar} = \alpha c \approx \frac{1}{137}c \ll c.$$

So we see that once again relativity only comes into play on scales less than a_0. Any relativistic corrections are going to be on the order of $\mathcal{O}v/c = \mathcal{O}\alpha$.

We can write the energy levels further as

$$E_n = -\frac{e^2}{2n^2 a_0} = -\frac{\mu e^4}{2n^2\hbar^2} = -\frac{\mu q^4}{32\hbar^2\pi^2\varepsilon_0^2 n^2}.$$

From $n = n_r + \ell$ since $n, n_r \in \mathbb{N}_{>0}$ we have that $\ell = 0, 1, \ldots, n - 1$, which corresponds to $n_r = n, n - 1, \ldots, 1$. For a particular value of ℓ we have $(2\ell + 1)$-fold degeneracy. This means that the total degeneracy of the eigenvalue E_n is

$$g_n = \sum_{\ell=0}^{n-1}(2\ell + 1) = 2\sum_{\ell=0}^{n-1}\ell + \sum_{\ell=0}^{n-1}1 = (n-1)n + n = n^2.$$

Here we have used the result proven in section B. The fact that $E_n = E_I/n^2$ was first proposed by Niels Bohr in 1913 before the full quantum picture that we have seen here was known. However quantum mechanics is still needed to find the degeneracy of E_n and other important properties.

The full radial wave function is

$$R_{n\ell}(r) = \frac{\chi_{n\ell}(r)}{r} = \frac{u_{n\ell}(a_0\rho)}{a_0\rho} = \frac{1}{a_0\rho}e^{-\lambda\rho}\eta_{n\ell}(\rho) = \frac{1}{a_0\rho}e^{-\lambda\rho}\sum_{q=0}^{n_r}c_q\rho^{q+\ell+1}.$$

The polynomials in the last term are called the **associated Laguerre polynomials**. They can be looked up in standard texts. If we do this we find that

$$R_{1,0}(r) = \frac{2}{a^{3/2}} e^{-r/a_0}$$

$$R_{2,0}(r) = \frac{1}{2\sqrt{2}} \frac{1}{a^{3/2}} \left[2 - \frac{r}{a_0} \right] e^{-r/(2a_0)}$$

$$R_{2,1}(r) = \frac{1}{2\sqrt{6}} \frac{1}{a^{3/2}} \frac{r}{a_0} e^{-r/(2a_0)}.$$

Notice that $R \sim a_0^{-3/2}$ so $|R|^2 \sim a_0^3$. This means that $|R|$ has units of probability per unit volume, which is what we want since $\psi = R Y_\ell^m$ is a probability density and Y_ℓ^m is dimensionless.

18   Non-degenerate Time-Independent Perturbation Theory

Many problems in quantum mechanics involve solving the TISE. In most cases however an analytic solution does not exist. If this is the case then the best we can do is find an estimate. One of the most common ways to do this is perturbation theory. This gives us a way to solve the TISE for a system with a Hamiltonian of the form

$$\hat{H} = \hat{H}_0 + \varepsilon \hat{V}.$$

Here $\hat{H}$ is the Hamiltonian we wish to find a solution for, $\hat{H}_0$ is a Hamiltonian for which we know the solution to the TISE and $\hat{V}$ is known as a perturbation. We will assume that $\hat{H}$ and $\hat{H}_0$ have discrete, non-degenerate eigenvalues. Let $|\varphi_n\rangle$ be the eigenstate of $\hat{H}_0$ with eigenvalue $E_n^{(0)}$, i.e.

$$\hat{H}_0 |\varphi_n\rangle = E_n^{(0)} |\varphi_n\rangle.$$

We want to find $|\psi_n\rangle$ and E_n such that

$$\hat{H} |\psi_n\rangle = E_n |\psi_n\rangle.$$

We look for a solution expanded in terms of ε. We can expand the energy as

$$E_n = E_{0n} + \varepsilon E_{1n} + \varepsilon^2 E_{2n} + \mathcal{O}\varepsilon^3 = \sum_{k=1}^{\infty} \varepsilon^k E_{kn},$$

where E_{kn} are to be found. Similarly we can expand the eigenstate as

$$|\psi_n\rangle = |\psi_{0n}\rangle + \varepsilon |\psi_{1n}\rangle + \varepsilon^2 |\psi_{2n}\rangle + \mathcal{O}\varepsilon^3 = \sum_{k=1}^{\infty} \varepsilon^k |\psi_{kn}\rangle,$$

where $|\psi_{kn}\rangle$ are to be found. Inserting these expansions and the definition of $\hat{H}$ into the TISE we find that

$$(\hat{H}_0 + \varepsilon \hat{V}) \sum_{i=1}^{\infty} |\psi_{in}\rangle = \left[\sum_{j=1}^{\infty} E_{jn} \right] \left[\sum_{k=1}^{\infty} |\psi_{kn}\rangle \right].$$

Expanding these sums up to second order in ε we have

$$\hat{H}_0 |\psi_{0n}\rangle + \varepsilon(\hat{H}_0 |\psi_{1n}\rangle + \hat{V} |\psi_{0n}\rangle) + \varepsilon^2(\hat{H}_0 |\psi_{2n}\rangle + \hat{V} |\psi_{1n}\rangle) + \mathcal{O}\varepsilon^3 =$$
$$E_{0n} |\psi_{0n}\rangle + \varepsilon(E_{0n} |\psi_{1n}\rangle + E_{1n} |\psi_{0n}\rangle) + \varepsilon^2(E_{0n} |\psi_{2n}\rangle + E_{1n} |\psi_{1n}\rangle + E_{2n} |\psi_{0n}\rangle) + \mathcal{O}\varepsilon^3.$$

18.1   Zeroth Order

Equating coefficients of terms of zeroth order in ε we have

$$\hat{H}_0 |\psi_{0n}\rangle = E_{0n} |\psi_{0n}\rangle.$$

Notice that this is exactly what we would get if $\varepsilon = 0$ so this is just the unperturbed Hamiltonian meaning that $|\psi_{0n}\rangle = |\varphi_n\rangle$ and $E_{0n} = E_n^{(0)}$.

18.2   First Order

18.2.1   Energy Perturbation

Equating coefficients of terms of first order in ε we have

$$\hat{H}_0|\psi_{1n}\rangle + \hat{V}|\psi_{0n}\rangle = E_{0n}|\psi_{1n}\rangle + E_{1n}|\psi_{0n}\rangle.$$

Rearranging this we have

$$(\hat{H} - E_{0n})|\psi_{1n}\rangle - (\hat{V} - E_{1n})|\psi_{0n}\rangle = 0. \tag{18.1}$$

We then take an inner product with $|\psi_{0n}\rangle$:

$$\langle\psi_{0n}|\hat{H}_0|\psi_{1n}\rangle - E_{0n}\langle\psi_{0n}|\psi_{1n}\rangle - \langle\psi_{0n}|\hat{V}|\psi_{0n}\rangle + E_{1n}\langle\psi_{0n}|\psi_{0n}\rangle = 0.$$

Consider the first term of this sum:

$$\begin{aligned}
\langle\psi_{0n}|\hat{H}_0|\psi_{1n}\rangle &= (\langle\psi_{1n}|\hat{H}_0^\dagger|\psi_{0n}\rangle)^* \\
&= (\langle\psi_{1n}|\hat{H}_0|\psi_{0n}\rangle)^* \\
&= E_{0n}(\langle\psi_{1n}|\psi_{0n}\rangle)^* \\
&= E_{0n}\langle\psi_{0n}|\psi_{1n}\rangle.
\end{aligned}$$

So the first term cancels with the second, note that we make no assumptions about $|\psi_{kn}\rangle$ being orthonormal. Rearranging the remaining two terms we have

$$E_{1n} = \frac{\langle\psi_{0n}|\hat{V}|\psi_{0n}\rangle}{\langle\psi_{0n}|\psi_{0n}\rangle} = \frac{\langle\varphi_n|\hat{V}|\varphi_n\rangle}{\langle\varphi_n|\varphi_n\rangle} = V_{00},$$

where V is a matrix representation of $\hat{V}$ and we assume that $|\psi_n\rangle$ are orthonormal. Note that since E_{0n} is a shift in the energy to first order it is sometimes written as ΔE_{0n}.

18.2.2   Eigenstate Perturbation

We can express $|\psi_{1n}\rangle$ in the energy eigenbasis of $\hat{H}_0$:

$$|\psi_{1n}\rangle = \sum_m c_m|\varphi_m\rangle,$$

where as usual $c_m = \langle\varphi_m|\psi_{1n}\rangle$. We can take a scalar product of equation 18.1 with $|\varphi_m\rangle$ and we get

$$\langle\varphi_m|\hat{H}_0|\psi_{1n}\rangle - E_{0n}\langle\varphi_m|\psi_{1n}\rangle + \langle\varphi_m|\hat{V}|\varphi_n\rangle + E_{1n}\langle\varphi_m|\varphi_n\rangle = 0.$$

We will first consider the case when $m \neq n$, thus the last term is zero. Analogously to how we worked with the first term for the energy perturbation we can show that the first term is simply $E_{0m}\langle\varphi_m|\psi_{1n}\rangle$. Thus

$$(E_{0m} - E_{0n})\langle\varphi_m|\psi_{1n}\rangle + \langle\varphi_m|\hat{V}|\varphi_n\rangle.$$

Hence

$$c_m = \langle\varphi_m|\psi_{1n}\rangle = \frac{\langle\varphi_m|\hat{V}|\varphi_n\rangle}{E_{0n} - E_{0m}} = \frac{V_{mn}}{E_{0n-E_{0m}}}.$$

Note that $E_{0n} = E_n^{(0)}$ and $E_{0m} = E_m^{(0)}$ are simply the nth and mth energy eigenvalues of $\hat{H}_0$ so are known quantities. Knowing this we can express the nth perturbed eigenstate, $|\psi_n\rangle$, as

$$|\psi_n\rangle = |\varphi_n\rangle + \varepsilon \sum_{m \neq n} \frac{V_{mn}}{E_{0n} - E_{0m}}|\varphi_m\rangle + \varepsilon c_n|\varphi_n\rangle + \mathcal{O}\varepsilon^2.$$

So to find the adjustment needed at first order we just need to compute c_n which we can do by imposing normalisation, taking the inner product of this with $|\psi_n\rangle$ we have

$$1 = \langle\psi_n|\psi_n\rangle = 1 + 2\varepsilon c_n\langle\varphi_n|\varphi_n\rangle + \mathcal{O}\varepsilon^2,$$

where we have used the fact that $|\varphi_n\rangle$ are orthonormal so $\langle\varphi_n|\varphi_m\rangle = \delta_{nm}$ meaning that the inner product with the sum term gives zero as $|\varphi_n\rangle$ is explicitly excluded from that sum. From this we conclude that we must have $c_n = 0$, so $|\psi_{1n}\rangle$ has no component in the $|\varphi_n\rangle$ direction. Thus

$$|\psi_{1n}\rangle = \sum_{m \neq n} \frac{V_{mn}}{E_{0n} - E_{0m}}|\varphi_m\rangle.$$

18.3   Second Order

Equating coefficients of terms of second order in ε we have

$$\hat{H}_0|\psi_{2n}\rangle + \hat{V}|\psi_{1n}\rangle = E_{0n}|\psi_{2n}\rangle + E_{1n}|\psi_{1n}\rangle + E_{2n}|\psi_{0n}\rangle.$$

Rearranging this we have

$$(\hat{H}_0 - E_{0n})|\psi_{2n}\rangle + (\hat{V} - E_{1n})|\psi_{1n}\rangle - E_{2n}|\psi_{0n}\rangle = 0.$$

We then take the inner product with $|\varphi_n\rangle$ and we have

$$\langle\varphi_n|\hat{H}_0|\psi_{2n}\rangle - E_{0n}\langle\varphi_n|\psi_{2n}\rangle + \langle\varphi_n|\hat{V}|\psi_{1n}\rangle - E_{1n}\langle\varphi_n|\psi_{1n}\rangle - E_{2n}\langle\varphi_n|\psi_{0n}\rangle = 0.$$

The first term is equal to $E_{0n}\langle\varphi_n|\psi_{2n}\rangle$ so combines with the second term to give zero. The fourth term is zero as $|\psi_{1n}\rangle$ has no $|\varphi_n\rangle$ component. The scalar product in the fifth term is simply one as $|\psi_{0n}\rangle = |\varphi_n\rangle$. Thus

$$\begin{aligned}
E_{2n} &= \langle\varphi_n|\hat{V}|\psi_{1n}\rangle \\
&= \langle\varphi_n|\hat{V}\sum_{m\neq n}\frac{V_{mn}}{E_{0n}-E_{0m}}|\varphi_m\rangle \\
&= \sum_{m\neq n}\frac{V_{mn}}{E_{0n}-E_{0m}}\langle\varphi_n|\hat{V}|\varphi_m\rangle \\
&= \sum_{m\neq n}\frac{V_{mn}V_{nm}}{E_{0n}-E_{0m}} \\
&= \sum_{m\neq n}\frac{|V_{mn}^2|}{E_{0n}-E_{0m}}.
\end{aligned}$$

18.4   Perturbation Examples

18.4.1   Infinite Well With a Cosine Floor

Consider the potential

$$V(x) = \begin{cases} V_0\cos\left(\frac{\pi x}{2a}\right), & |x| \leq a, \\ \infty, & |x| > a. \end{cases}$$

If we define $\hat{H}'$ as

$$V_0\cos\left(\frac{\pi x}{2a}\right)$$

then we can view the Hamiltonian for this system as

$$\hat{H} = \hat{H}_0 + \hat{H}'$$

where $\hat{H}_0$ is the Hamiltonian for an infinite square well. The solution to the TISE for the infinite square well is discussed in section 9.3. In particular the nth energy level is

$$E_n^{(0)} = \frac{\hbar^2\pi^2 n^2}{8ma^2},$$

and the corresponding eigenfunction is

$$\varphi_{(n)}(x) = \begin{cases} \frac{1}{\sqrt{a}}\cos\left(\frac{n\pi x}{2a}\right), & \text{for odd } n, \\ \frac{1}{\sqrt{a}}\sin\left(\frac{n\pi x}{2a}\right), & \text{for even } n. \end{cases}$$

Perturbation theory only gives a good estimate if the energy shift due to the addition of $\hat{H}'$ to the Hamiltonian is small, in this case this means that we require $V_0 \ll E_0^{(2)} - E_0^{(1)}$. The energy shift is then

$$\Delta E_1 = \langle\varphi^{(1)}|\hat{H}'|\varphi^{(1)}\rangle$$

$$= \int_{-a}^{a} \varphi^{(1)}(x)^* V_0 \cos\left(\frac{\pi x}{2a}\right) \varphi^{(1)}(x) \, dx$$

$$= \frac{V_0}{a} \int_{-a}^{a} \cos^3\left(\frac{\pi x}{2a}\right) dx$$

$$= \frac{8V_0}{3\pi}.$$

So to first order the energy of the ground state is

$$E_1 = E_1^{(0)} + \Delta E_1 = \frac{\hbar^2 \pi^2 n^2}{8ma^2} + \frac{8V_0}{3\pi}.$$

18.4.2   Helium

Helium is a two electron atom. If we ignore the interaction between the electrons then the Hamiltonian for a helium atom is

$$\hat{H}_0 = \frac{\hbar^2}{2m}\nabla_1^2 - \frac{Ze^2}{r_1} + \frac{\hbar^2}{2m}\nabla_2^2 - \frac{Ze^2}{r_2} = \hat{H}_0^{(1)} + \hat{H}_0^{(2)}.$$

Here m is the reduced mass of an electron and the nucleus,

$$\frac{1}{m} = \frac{1}{m_e} + \frac{1}{2m_p} \approx \frac{1}{m_e}.$$

The distances r_1 and r_2 are the distances from the nucleus to the relevant electron, where we model the four nucleon nucleus as a point charge of charge $Z = 2e$. The Laplacian operator, ∇_1^2 or ∇_2^2, acts only on the coordinates of the relevant electron.

The Hamiltonian is independent of the spin of any one part of the system. This means that we can treat the spin of the electrons as we do for a normal two particle, spin 1, system. A generic state, $|\Psi\rangle \in \mathcal{H}$ can be described by $|\psi\rangle \in \mathcal{H}_{\text{space}}$ and $|s, s_z\rangle \in \mathcal{H}_{\text{spin}}$ as

$$|\Psi\rangle = |\psi\rangle \otimes |s, s_z\rangle = |\psi; s, s_z\rangle.$$

As discussed in section 15.2

$$|s = 1, s_z = 1\rangle = |\uparrow\uparrow\rangle, \qquad |s = 1, s_z = -1\rangle = |\downarrow\downarrow\rangle, \qquad |s = 1, s_z = 0\rangle = \frac{\sqrt{2}}{2}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle),$$

$$\text{and} \qquad |s = 0, s_z = 0\rangle = \frac{\sqrt{2}}{2}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle).$$

The spatial wave function is given by

$$\langle \boldsymbol{r_1}, \boldsymbol{r_2} | \psi \rangle = \psi(\boldsymbol{r_1}, \boldsymbol{r_2}) = \psi^{(1)}(\boldsymbol{r_1})\psi^{(2)}(\boldsymbol{r_2}),$$

where we assume in the last step that a separable solution exists. Each $\psi^{(k)}$ must satisfy the TISE for the relevant Hamiltonian:

$$\hat{H}_0^{(k)}\psi^{(k)}(\boldsymbol{r_k}) = \left[\frac{\hbar^2}{2m}\nabla_k^2 - \frac{Ze^2}{r_k}\right]\psi^{(k)}(\boldsymbol{r_k}) = E^{(k)}\psi^{(k)}(\boldsymbol{r_k}).$$

This is simply the TISE that we solved for the hydrogen atom but with the nucleus charge changed to $Z = 2e$ instead of $Z = e$. The solutions to this are the states $u_{n\ell m}$, where the ground state is

$$\psi^{(k)}(\boldsymbol{r_k}) = u_{100}(\boldsymbol{r_k}) = \frac{1}{\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{3/2} e^{-r/a_0}.$$

Thus

$$\langle \boldsymbol{r_1}, \boldsymbol{r_2} | \Psi \rangle = u_{100}(\boldsymbol{r_1})u_{100}(\boldsymbol{r_2}) \sum_{\alpha=1}^{4} c_\alpha |\alpha\rangle$$

where

$$\{|\alpha\rangle\} = \{|s, s_z\rangle \mid s = 0, 1 \text{ and } s_z = -s, \ldots, s\}.$$

Electrons are fermions. This means that $|\Psi\rangle$ should be antisymmetric under exchanging of 1 and 2. The radial part, $u_{100}(\boldsymbol{r_1})u_{100}(\boldsymbol{r_2})$, is symmetric under exchanging of 1 and 2 so the spin part must be antisymmetric. The only antisymmetric spin state available is $|s = 0, s_z = 0\rangle$.

We are now in a position to consider what happens when we include the interaction between the electrons. This interaction is characterised by the potential

$$\hat{V} = \frac{e^2}{|\boldsymbol{r_1} - \boldsymbol{r_2}|}.$$

The energy shift due to this perturbation is

$$\Delta E_1 = \langle \Psi | \hat{V} | \Psi \rangle$$
$$= \langle \psi; 0, 0 | \hat{V} | \psi; 0, 0 \rangle$$

using the fact that $\hat{V}$ acts only on the radial part this becomes

$$= \langle 0, 0 | 0, 0 \rangle \langle \psi | \hat{V} | \psi \rangle$$

considering the ground state, $\psi(\boldsymbol{r}) = u_{100}(\boldsymbol{r})$, or equivalently $|\psi\rangle = |100\rangle$, this becomes

$$= \langle 100, 100 | \hat{V} | 100, 100 \rangle$$
$$= \int \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \, u_{100}^*(\boldsymbol{r_1}) u_{100}^*(\boldsymbol{r_2}) \frac{e^2}{|\boldsymbol{r_1} - \boldsymbol{r_2}|} u_{100}(\boldsymbol{r_1}) u_{100}(\boldsymbol{r_2})$$
$$= \left(\frac{eZ^3}{\pi a_0^3} \right)^2 \int \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \exp\left(-\frac{2Z(r_1 + r_2)}{a_0} \right) \frac{1}{|\boldsymbol{r_1} - \boldsymbol{r_2}|}$$
$$= \frac{5}{4} Z E_I$$

where $E_I \approx 13.6\,\mathrm{eV}$ is defined as it was in section 17.1.

Part V

Recap

19   Recap Part One

<div style="border:1px solid green; padding:8px;">

Postulate 1

Every possible physical state of a given system corresponds to some state vector, $|\Psi(t)\rangle \in \mathcal{H}$, from which all possible predictions of the physical properties of the system can be obtained.

</div>

The state vector can often be represented by a wave function, $\Psi(\boldsymbol{r}, t) = \langle \boldsymbol{r} | \Psi(t) \rangle$, however this is not always the case. For example the spin of a system cannot be represented in this way and we instead represent spin with column vectors.

<div style="border:1px solid green; padding:8px;">

Postulate 2

Every physical observable is represented by a Hermitian operator. For each observable, $\mathcal{A}$, there is a Hermitian operator, $\hat{A}$, with a complete orthonormal set of eigenfunctions, $\{u_i\}$, or eigenvectors, $\{|u_i\rangle\}$, or generically eigenstates, which have a corresponding set of real eigenvalues, $\{A_i\}$, such that

$$\hat{A} u_i(\boldsymbol{r}) = A_i u_i(\boldsymbol{r}), \qquad \text{or} \qquad \hat{A} | u_i \rangle = A_i | u_i \rangle.$$

The set of possible values that a measurement of $\mathcal{A}$ can yield is identically $\{A_i\}$.

</div>

The Hermitian conjugate of an operator, $\hat{O}$, is $\hat{O}^\dagger$, which is defined by

$$\left(\int \psi^* \hat{O} \varphi \, \mathrm{d}^3 r \right)^* = \int \varphi^* \hat{O}^\dagger \psi \, \mathrm{d}^3 r \qquad \text{or} \qquad \langle \psi | \hat{O} | \varphi \rangle^* = \langle \varphi | \hat{O}^\dagger | \psi \rangle,$$

where $|\varphi\rangle, |\psi\rangle \in \mathcal{H}$ are arbitrary states. The operator $\hat{O}$ is Hermitian if $\hat{O}^\dagger = \hat{O}$.

States, $\{u_i\}$ or $\{|u_i\rangle\}$, are orthonormal if

$$\int u_i^*(\boldsymbol{r}) u_j(\boldsymbol{r}) \, \mathrm{d}^3 r = \delta_{ij}, \qquad \text{or} \qquad \langle u_i | u_j \rangle = \delta_{ij}.$$

A set of states, $\{u_i\}$ or $\{|u_i\rangle\}$, is complete if an arbitrary state, $\Psi(\boldsymbol{r}, t)$ or $|\Psi(t)\rangle$, can be represented by

$$\Psi(\boldsymbol{r}, t) = \sum_i c_i(t) u_i(\boldsymbol{r}), \qquad \text{or} \qquad |\Psi(t)\rangle = \sum_i c_i(t) |u_i\rangle.$$

The coefficients $\{c_i\}$ are given by

$$c_i(t) = \int u_i^*(\boldsymbol{r}) \Psi(\boldsymbol{r}, t) \, \mathrm{d}^3 r, \qquad \text{or} \qquad c_i(t) = \langle u_i | \Psi(t) \rangle.$$

Compare this to some vector $\boldsymbol{v} \in \mathbb{R}^n$ where $\boldsymbol{v} = \sum_i v_i \boldsymbol{e_i}$ with $v_i = \boldsymbol{e_i} \cdot \boldsymbol{v}$. The set of eigenfunctions, $\{u_i\}$, or eigenvectors, $\{|u_i\rangle\}$, is referred to as the eigenbasis of $\hat{A}$. Comparing the equations for $\Psi(\boldsymbol{r}, t)$ or $|\Psi(t)\rangle$ and $c_i(t)$ we see that

$$|\Psi(t)\rangle = \sum_i \langle u_i | \langle u_i | \Psi(t) \rangle$$

and so the completeness relation,

$$\sum_i |u_i\rangle \langle u_i| = \mathbb{1}$$

where $\mathbb{1}$ is the identity operator defined by

$$\mathbb{1}|\Psi\rangle = |\Psi\rangle$$

for all $|\Psi\rangle$.

The spectrum of eigenvalues of $\hat{A}$ may be discrete (which we assumed was the case above) or continuous. For example the angular momentum operators $\hat{L}^2$ and $\hat{L}_z$ have discrete spectra whereas the position, $\hat{X}$ and momenta $\hat{P}$ have continuous spectra and the energy, $\hat{H}$, has a spectrum that is continuous in some cases and discrete in other cases.

Observables can be represented by differential operators acting on the wave function or as Hermitian matrices acting on column vectors, for example

$$\hat{P} \to -i\hbar \nabla, \qquad \text{and} \qquad \hat{S}_z \to \frac{1}{2}\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Postulate 3

If the observable $\mathcal{A}$ is measured on a system which, immediately prior to the measurement, is in the state $|\Psi(t)\rangle$ then the strongest predictive statement we can make about the result is that the probability of measuring $\mathcal{A}$ to be A_i is

$$P(A_i) = |\langle u_i | \Psi(t) \rangle|^2 = |c_i(t)|^2.$$

This assumes that the state $|\Psi(t)\rangle$ is normalised so that

$$|\langle \Psi(t) | \Psi(t) \rangle|^2 = \sum_i |\langle u_i | \Psi(t) \rangle|^2 = \sum_i |c_i(t)|^2 = 1.$$

We also assume that measurements are ideal and yield a single, errorless, number.

The coefficients, $c_i(t)$, are called the probability amplitudes. In general we cannot predict with certainty the outcome of a measurement unless $|\Psi(t)\rangle$ happens to coincide with an eigenstate of $\hat{A}$.

19.1   Repeated Measurements

Repeated measurements are measurements made on the same state, $|\Psi(t)\rangle$. We imagine creating an assembly of an infinite number of identical systems and performing each measurement on a different one of these systems. The average value of this process, if all measurements are completed at time t, is

$$\langle \hat{A}\rangle_t = \int \Psi^*(\boldsymbol{r},t)\hat{A}\Psi(\boldsymbol{r},t)\,\mathrm{d}^3r\,, \qquad \text{or} \qquad \langle \hat{A}\rangle_t = \langle\Psi(t)|\hat{A}|\Psi(t)\rangle.$$

If we expand $|\Psi(t)\rangle$ in the eigenstates of $\hat{A}$ then we see that

$$
\begin{aligned}
\langle \hat{A}\rangle_t &= \sum_{i,j} c_i^*(t)c_j(t)\langle u_i|\hat{A}|u_j\rangle \\
&= \sum_{i,j} c_i^*(t)c_j(t)A_j\langle u_i|u_j\rangle \\
&= \sum_{i,j} c_i^*(t)c_j(t)A_j\delta_{ij} \\
&= \sum_{i} |c_i|^2 A_i \\
&= \sum_{i} P(A_i)A_i.
\end{aligned}
$$

This is the usual definition of the mean of a discrete variable $\mathcal{A}$ which can take values in $\{A_i\}$.

19.2   Uncertainty

The uncertainty is the root mean square deviation about the mean of an infinite amount of repeated measurements. It is given at time t by

$$\Delta\hat{A}_t\sqrt{\langle\hat{A}^2\rangle_t - \langle\hat{A}\rangle_t^2}.$$

This uncertainty is not due to the way the experiment is conducted or due to missing data. It is an intrinsic limit to what we can know.

The generalised uncertainty relation states that two observables $\mathcal{A}$ and $\mathcal{B}$ have uncertainties $\Delta\hat{A}_t$ and $\Delta\hat{B}_t$ respectively which satisfy

$$\Delta\hat{A}_t\Delta\hat{B}_t \geq \frac{1}{2}\Big|[\hat{A},\hat{B}]\Big|$$

where

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

is the commutator. The most important case of this relationship is

$$\Delta\hat{X}_t\Delta\hat{P} \geq \frac{\hbar}{2}.$$

20   Recap Part Two

20.1   Successive Measurements

Successive measurements refers to measurements made on the same system in quick succession.

> **Postulate 4**
>
> A measurement of an observable, $\mathcal{A}$, generally causes a drastic, uncontrollable change in the state of the system. Regardless of the state, $|\Psi(t)\rangle$, before the measurement immediately after the measurement the state vector will coincide with the eigenstate $|u_k\rangle$ corresponding to the measured value, A_k, of $\mathcal{A}$.

This is referred to as collapse of the wave function or reduction of state. We say that the state is forced or projected into the eigenstate. This assumes that there is no degeneracy. If instead A_k is g-fold degenerate and the g eigenstates $\{|u_i^{(k)}\rangle\}$ all have eigenvalue A_k then if a system is measured to have $\mathcal{A} = A_k$ we will find immediately after this measurement that the state is a linear superposition of the degenerate states:

$$\frac{1}{\sqrt{\sum\limits_{i=1}^{g}\left|c_i^{(k)}\right|^2}} \sum_{i=1}^{n} c_i^{(k)}|u_i^{(k)}\rangle.$$

20.2   Compatible Observables

If $\mathcal{A}$ and $\mathcal{B}$ are observables and we perform the following series of measurements,

1. Measure $\mathcal{A}$,

2. Measure $\mathcal{B}$,

3. Measure $\mathcal{A}$,

then $\mathcal{A}$ and $\mathcal{B}$ are said to be compatible if and only if the result of the third measurement is guaranteed to be the same as the result of the first measurement.

Theorem 8: Compatibility Theorem

If $\mathcal{A}$ and $\mathcal{B}$ are observables represented by Hermitian operators $\hat{A}$ and $\hat{B}$ respectively then the following are equivalent:
- $\mathcal{A}$ and $\mathcal{B}$ are compatible observables.
- $\hat{A}$ and $\hat{B}$ have a common eigenbasis.
- $\hat{A}$ and $\hat{B}$ commute.

A set of variables, $\{\mathcal{A}^{(i)}\}$ are called a complete set of commuting/compatible observables (CSCO) if the eigenvalues of each observable, $A_k^{(i)}$, uniquely identify an eigenvector in the common eigenbasis. For example for a hydrogen atom the set energy, angular momentum squared and the z component of angular momentum are a CSCO as a state $|n, \ell, m\rangle$ is uniquely identified by the numbers n, ℓ, and m.

20.3   Time Development

Postulate 5

The time development of a quantum system is determined by the time dependent Schrödinger equation (TDSE):

$$\hat{H}\Psi(\boldsymbol{r}, t) = i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{r}, t), \qquad \text{or} \qquad \hat{H}|\Psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle.$$

$\hat{H}$ is the Hamiltonian operator which is formed from the corresponding classical Hamiltonian replacing the variables with corresponding operators. $\hat{H}$ represents the total energy of the system.

The wave function can be interpreted as $|\Psi(\boldsymbol{r}, t)|^2 \, \mathrm{d}^3 r$ giving the probability that a measurement of position will give a result in the infinitesimal volume element $\mathrm{d}^3 r$ at $\boldsymbol{r}$ where the volume element in Cartesian and spherical coordinates is $\mathrm{d}^3 r = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = r^2 \sin\vartheta\,\mathrm{d}r\,\mathrm{d}\vartheta\,\mathrm{d}\varphi$. Hence $|\Psi(\boldsymbol{r}, t)|$ is a probability density.

The classical Hamiltonian is $H = T + V$. When constructing the quantum Hamiltonian the momentum, $\boldsymbol{p}$, is substituted for the momentum operator $\hat{\boldsymbol{P}}$ which has the following representation:

$$\hat{P}_i \to -i\hbar\frac{\partial}{\partial x_i} \implies \hat{\boldsymbol{P}} \to -i\hbar\nabla.$$

For a single particle $T = p^2/2m = \boldsymbol{p}\cdot\boldsymbol{p}/2m$ and therefore

$$\hat{T} = \frac{\hat{P}^2}{2m} \to \frac{\hbar^2}{2m}\nabla^2.$$

Most of the time we consider the potential, V, to be a function of position only, $V = V(\boldsymbol{r})$. The position in the classical Hamiltonian is replaced by the position operator, $\hat{\boldsymbol{X}}$, which has the following representation:

$$\hat{X}_i = x_i \implies \hat{\boldsymbol{X}} = \boldsymbol{r}.$$

Thus

$$\hat{V} \to V(\boldsymbol{r}).$$

So the Hamiltonian operator is given by

$$\hat{H} = \hat{T} + \hat{V} = \frac{\hat{P}^2}{2m} + V(\hat{\boldsymbol{X}}) \to -\frac{\hbar^2}{2m} + V(\boldsymbol{r}).$$

The TDSE is then

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r}) \right] \Psi(\boldsymbol{r}, t) = i\hbar \frac{\partial}{\partial t}\Psi(\boldsymbol{r}, t).$$

This has separable solutions of the form

$$\Psi_n(\boldsymbol{r}, t) = u_n(\boldsymbol{r})\exp(-iE_n t/\hbar).$$

This gives us the time independent Schrödinger equation (TISE):

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r}) \right] u_n(\boldsymbol{r}) = E_n u_n(\boldsymbol{r}), \qquad \text{or} \qquad \hat{H}|n\rangle = E_n|n\rangle.$$

This is an eigenvalue equation for $\hat{H}$ so we see that the stationary states, $\{|u_n\rangle\}$, correspond to the energy levels.

It can be shown that for an observable $\mathcal{A}$ represented by the Hermitian operator $\hat{A}$

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \hat{A} \rangle_t = \frac{i}{\hbar}\langle [\hat{H}, \hat{A}] \rangle_t.$$

Importantly if $\hat{A}$ and $\hat{H}$ commute then $\langle \hat{A} \rangle_t$ is constant.

20.4 Angular Momentum

The classical angular momentum, $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$, under operator substitution becomes the angular momentum operator

$$\hat{\boldsymbol{L}} = \hat{\boldsymbol{X}} \times \hat{\boldsymbol{P}}$$

which has components

$$\hat{L}_x \to -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y} \right), \qquad \hat{L}_y \to -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z} \right), \qquad \text{and} \qquad \hat{L}_z \to -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} \right).$$

More compactly:

$$\hat{L}_k \to -i\hbar\varepsilon_{ijk}x_i\frac{\partial}{\partial x_j}.$$

These satisfy the commutation relations

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z, \qquad [\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x, \qquad \text{and} \qquad [\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y.$$

More compactly

$$[\hat{L}_i, \hat{L}_j] = i\varepsilon_{ijk}\hat{L}_k.$$

We also define

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2.$$

Which satisfies the commutation relation

$$[\hat{L}^2, \hat{L}_k] = 0.$$

This means that $\hat{L}^2$ and $\hat{L}_z$ have a common eigenbasis. We denote these by $\{|\ell, m\rangle\}$ where

$$\hat{L}^2|\ell, m\rangle = \ell(\ell+1)\hbar^2|\ell, m\rangle, \qquad \text{and} \qquad \hat{L}_z|\ell, m\rangle = m\hbar|\ell, m\rangle.$$

As wave functions these states are represented by the spherical harmonics, Y_ℓ^m, so

$$\hat{L}^2 Y_\ell^m(\vartheta, \varphi) = \ell(\ell+1)\hbar^2 Y_\ell^m(\vartheta, \varphi), \qquad \text{and} \qquad \hat{L}_z Y_\ell^m(\vartheta, \varphi) = m\hbar Y_\ell^m(\vartheta, \varphi).$$

For angular momentum we have two quantum numbers. The orbital angular momentum number, $\ell = 0, 1, 2, \ldots$ and the magnetic quantum number, m, which for a particular value of ℓ can take any of the $2\ell + 1$ integer values $-\ell, -\ell + 1, \ldots, \ell - 1, \ell$.

20.5   Spin

Many particles also have an intrinsic angular momentum known as spin. This is similar to the angular momentum but we remove the initial definition of the angular momentum operator and only consider the algebra of the operators. We define operators $\hat{S}_k$ which satisfy

$$[\hat{S}_i, \hat{S}_j] = i\hbar\varepsilon_{ijk}\hat{S}_k.$$

We also define

$$\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2$$

which has the property that

$$[\hat{S}^2, \hat{S}_k] = 0.$$

The simultaneous eigenstates of $\hat{S}^2$ an $\hat{S}_z$ are denoted $|s, m\rangle$ and are such that

$$\hat{S}^2|s, m\rangle = s(s+1)\hbar^2|s, m\rangle \qquad \text{and} \qquad \hat{S}_z|s, m\rangle = m\hbar|s, m\rangle.$$

The spin quantum number, s, can be shown to take integer and half integer values, $s = -, 1/2, 1, 3/2, 2, \ldots$ and for a specific value of s there are $2s + 1$ possible values of m which run in integer steps from $-s$ to s meaning that m is an integer if s is and m is a half integer if s is.

Spin states don't have a wave function representation. Instead we represent them by matrices with matrix elements

$$\langle s, m'|\hat{S}_z|s, m\rangle = m\hbar\langle s, m'|s, m\rangle = m\hbar\delta_{mm'}.$$

This will, for a specific value of s, give a $(2s+1) \times (2s+1)$ matrix. In the case that $s = 1/2$ we have two options for $m = \pm 1/2$. We use

$$|s = 1/2, m = 1/2\rangle \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \text{and} \qquad |s = 1/2, m = -1/2\rangle \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

which we call spin up and down respectively. Then the Pauli matrices,

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \text{and} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

can be used to represent the spin operators, $\hat{S}_k \rightarrow \hbar\sigma_k/2$.

21   Recap Part Three

21.1   Addition of Angular Momenta

Suppose a system has two independent angular momenta with corresponding operators $\hat{\boldsymbol{L}}$ and $\hat{\boldsymbol{S}}$. We can define the total angular momentum with the operator

$$\hat{\boldsymbol{J}} = \hat{\boldsymbol{L}} + \hat{\boldsymbol{S}}.$$

Similarly we can define $\hat{J}_k$ operators with the following commutation relations

$$[\hat{J}_i, \hat{J}_j] = i\hbar\hat{J}_k.$$

We define the square of the total angular momentum through the operator

$$\hat{J}^2 = \hat{\boldsymbol{J}} \cdot \hat{\boldsymbol{J}} = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2.$$

This operator satisfies the commutation relationship

$$[\hat{J}^2, \hat{J}_k] = 0.$$

Thus we have a simultaneous basis for $\hat{J}^2$ and $\hat{J}_z$. We denote this basis $\{|j, m_j\rangle\}$ where

$$\hat{J}^2|j, m_j\rangle = j(j+1)\hbar^2|j, m_j\rangle$$
$$\hat{J}_z|j, m_j\rangle = m\hbar|j, m_j\rangle.$$

First note that since $\hat{L}^2$ and $\hat{S}^2$ act on completely different spaces they commute. Therefore

$$\hat{J}^2 = (\hat{\boldsymbol{L}} + \hat{\boldsymbol{S}}) \cdot (\hat{\boldsymbol{L}} + \hat{\boldsymbol{S}})$$
$$= \hat{L}^2 + \hat{S}^2 + 2\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{S}}.$$

Using this we see that

$$[\hat{J}^2, \hat{L}_z] = \underbrace{[\hat{L}^2, \hat{L}_z]}_{=0} + \underbrace{[\hat{S}^2, \hat{L}_z]}_{=0} + 2[\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{S}}, \hat{L}_z]$$
$$= 2[\hat{L}_x\hat{S}_x, \hat{L}_z] + 2[\hat{L}_y\hat{S}_y, \hat{L}_z] + 2\underbrace{[\hat{L}_z\hat{S}_z, \hat{L}_z]}_{=0}$$
$$= 2\left(\hat{L}_x\hat{S}_x\hat{L}_z - \hat{L}_z\hat{L}_x\hat{S}_x + \hat{L}_y\hat{S}_y\hat{L}_z - \hat{L}_z\hat{L}_y\hat{S}_y\right)$$
$$= 2\hat{S}_x[\hat{L}_x, \hat{L}_z] + 2\hat{S}_y[\hat{L}_y, \hat{L}_z]$$
$$= -2i\hbar\hat{L}_y + 2i\hbar\hat{L}_x$$
$$\neq 0.$$

So $\hat{J}^2$ doesn't commute with $\hat{L}_z$, similarly we can show that $\hat{J}^2$ doesn't commute with $\hat{S}_z$.

There are two sets of commuting observables that we may consider. The first is $\{\hat{L}^2, \hat{L}_z, \hat{S}^2, \hat{S}_z\}$, which has a common eigenbasis, $\{|\ell, m_\ell, s, m_s\rangle\}$, called the uncoupled basis. The second is the set $\{\hat{L}^2, \hat{S}^2, \hat{J}^2, \hat{J}_z\}$, which has a common eigenbasis, $\{|\ell, s, j, m_j\rangle\}$, called the coupled basis. These two bases are linearly related and the expansion coefficients are called the Clebsch–Gordan coefficients.

The angular momentum addition theorem tells us that for given ℓ and s the allowed values of j are

$$\ell + s, \ell + s - 1, \ldots, |\ell - s + 1|, |\ell - s|.$$

For example an electron with $\ell = 1$ and $s = 1/2$ has two possible values of j:

$$j = \ell + s = 1 + \frac{1}{2} = \frac{3}{2}, \qquad \text{and} \qquad j = \ell - s = 1 - \frac{1}{2} = \frac{1}{2}.$$

A two electron system where each electron has spin $1/2$ will have total spin quantum number

$$s = s_1 + s_2 = \frac{1}{2} + \frac{1}{2} = 1, \qquad \text{and} \qquad s = s_1 - s_2 = \frac{1}{2} - \frac{1}{2} = 0.$$

21.2   Degeneracy

Up until now we have largely ignored degeneracy in this recap. Degeneracy is when two or more eigenstates have the same eigenvalue. For example the first excited energy state of the hydrogen atom is four fold degenerate as u_{200}, u_{211}, u_{210}, and u_{21-1} all have the same energy eigenvalue, E_2. The degeneracy is the reason that we need the orbital angular momentum quantum number, ℓ, and even then we still have $(2\ell + 1)$-fold degeneracy which we lift by introducing the magnetic quantum number, m_ℓ.

If we have degenerate states then we need to alter the third and fourth postulates. In the following we will consider a system with an observable, $\mathcal{A}$, which is associated with the operator $\hat{A}$, which has an eigenbasis $\{|u_i\rangle\}$ where $|u_1\rangle$ and $|u_2\rangle$ have a common eigenvalue A. The rest of the eigenstates have an eigenvalue other than A but may be degenerate still with these eigenvalues.

> **Postulate 3 (with degeneracy)**
>
> If the observable $\mathcal{A}$ is measured on a system which, immediately prior to the measurement, is in the state $|\Psi(t)\rangle = \sum_i c_i(t)|u_i\rangle$, then the probability of getting the result A is
>
> $$P(A) = |\langle u_1|\Psi(t)\rangle|^2 + |\langle u_2|\Psi(t)\rangle|^2 = |c_1(t)|^2 + |c_2(t)|^2.$$

This generalises to g-fold degeneracy for A by summing over all g states with A as an eigenvalue of $\hat{A}$. Any linear combination of degenerate eigenstates is another degenerate eigenstate, for example

$$k_1|u_1\rangle + k_2|u_2\rangle,$$

is an eigenstate with eigenvalue A. The set of all degenerate eigenstates spans a subspace of degenerate eigenstates. In this case it is a two dimensional subspace. We use this freedom to assume that $|u_1\rangle$ and $|u_2\rangle$ are orthonormal as if they aren't then we can always construct an orthonormal pair of states, $|u_1'\rangle$ and $|u_2'\rangle$, defined by

$$|u_1'\rangle = |u_2'\rangle, \qquad \text{and} \qquad |u_2'\rangle = |u_2\rangle - \frac{\langle u_1'|u_2\rangle}{\langle u_1'|u_1'\rangle}|u_1\rangle.$$

From now we will assume that $|u_1\rangle$ and $|u_2\rangle$ are orthonormal.

> **Postulate 4**
>
> If the result of a measurement of $\mathcal{A}$ is A then immediately after the measurement the state vector will coincide with the normalised eigenstate
>
> $$\frac{c_1(t)|u_1\rangle + c_2(t)|u_2\rangle}{\sqrt{|c_1(t)|^2 + |c_2(t)^2|}}.$$

Again this generalises to g-fold degeneracy simply by extending the sums to be over all g degenerate states. Notice that, unlike the non-degenerate case, the state after measurement is not necessarily one eigenbasis vectors.

21.3 Complete Sets of Commuting Operators

We distinguish eigenstates by their eigenvalues. Clearly we cannot separate $|u_1\rangle$ and $|u_2\rangle$ by the eigenvalue A. Instead we must introduce a new observable, $\mathcal{B}$, which is compatible with $\mathcal{A}$, and then we use the eigenvalue of $\hat{B}$ to distinguish the states. For example $\hat{L}^2$ has $2\ell + 1$ degenerate eigenstates for a given value of ℓ. We distinguish these states by the eigenvalue of $\hat{L}_z$ which gives us a distinct eigenvalue, m_ℓ, for each of the degenerate states. We end up labelling a state with two eigenvalues, $|\ell, m_\ell\rangle$.

We also use this to lift the degeneracy of the stationary states of the hydrogen atom. The first operator we use is the Hamiltonian which assigns each eigenstate an energy value, E_n, and we use n as a label. We then introduce $\hat{L}^2$ which allows some separation between states of the same energy with the extra label ℓ. It isn't until we introduce $\hat{L}_z$ with the label m that we can completely separate the states which we then label $|u_{n\ell m}\rangle$ or $|n, \ell, m\rangle$.

We keep introducing new compatible observables until we have a complete set of commuting/compatible observables (CSCO), $\{\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots\}$, with corresponding operators, $\{\hat{A}, \hat{B}, \hat{C}, \dots\}$, which have a common eigenbasis which we label $\{|A, B, C, \dots\rangle\}$. If $\{A, B, C, \dots\}$ uniquely identify a state then we call them good quantum numbers. For example one set of good quantum numbers for the hydrogen atom is $\{n, \ell, m\}$. A good set of quantum numbers for a two electron system is $\{s_1, s_2, s, m_s\}$ or $\{s_1, m_{s_1}, s_2, m_{s_2}\}$.

A central potential with no external fields and a spin $1/2$ system has the Hamiltonian

$$\hat{H} = \frac{\hat{P}^2}{2m} + \hat{V}(r).$$

Any pair of $\{\hat{H}, \hat{L}^2, \hat{L}_z, \hat{S}^2, \hat{S}_z\}$ commute and the common eigenbasis is $\{|n, \ell, m_\ell, s, m_s\rangle\}$ with good quantum numbers $\{n, \ell, m_\ell, s, m_s\}$. Alternatively we could work in the coupled basis, $\{|n, \ell, s, j, m_j\rangle\}$, with good quantum numbers $\{n, \ell, s, j, m_j\}$.

A central potential spin 1/2 system with spin-orbit interactions has the Hamiltonian

$$\hat{H} = \frac{\hat{P}^2}{2m} + \hat{V}(r) + a(r)\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{S}}.$$

In this case $\hat{L}_z$ and $\hat{S}_z$ don't commute with $\hat{H}$ due to the spin-orbit term containing $\hat{L}_x$, $\hat{L}_y$, $\hat{S}_x$, and $\hat{S}_y$. Therefore m_ℓ and m_s are no longer good quantum numbers. However we can still use the coupled basis, $\{|n, \ell, s, j, m_j\rangle\}$ with good quantum numbers $\{n, \ell, s, j, m_j\}$.

To know exactly what state a system is in we must measure all of the observables in a CSCO. This is called a **maximal measurement**. It enables us to know exactly which eigenstate a series of measurements forces a state into by examining the set of eigenvalues of that CSCO.

Part VI

Approximation Methods

22   Time Independent, Non-Degenerate, Perturbation Theory

Many problems in quantum mechanics cannot be solved analytically. In these cases often the best we can do is approximate. One of the principle ways that we do this is we try to treat the unsolvable problem as a solvable problem with an extra term. This is done by defining the Hamiltonian

$$\hat{H} = \hat{H}_0 + \hat{H}'$$

where $\hat{H}_0$ is the Hamiltonian for a system which we can solve exactly and $\hat{H}'$, known as the perturbation Hamiltonian, is a 'small' (exactly what we mean by small will be made clear later) perturbation.

Since the TISE can be solved exactly for $\hat{H}_0$ we assume that the eigenvalues, $\{E_n^{(0)}\}$, are known as well as the corresponding eigenstates, $\{n^{(0)}\}$. The aim is to find the eigenvalues and eigenstates of $\hat{H}$. Since $\hat{H}'$ is 'small' we expect these to be close to the eigenvalues and eigenstates of $\hat{H}_0$.

As with many other problems when we have to approximate we attempt a solution in the term of a series. We introduce a parameter, λ, which allows us to keep track of powers of $\hat{H}'$:

$$\hat{H} = \hat{H}_0 + \lambda\hat{H}'.$$

This parameter will be set to one later on.

We will assume that the eigenstates of $\hat{H}$ and $\hat{H}_0$ are discrete. We will also assume that the eigenstates $|n^{(0)}\rangle$ are non-degenerate and therefore all uniquely identified by their energy. In later sections we will look at what happens if $|n^{(0)}\rangle$ aren't degenerate or $\hat{H}'$ is time dependent.

Currently we know that

$$\hat{H}_0|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle.$$

The goal is, as usual, to solve the TISE for $\hat{H}$:

$$\hat{H}|n\rangle = E_n|n\rangle.$$

We assume that E_n and $|n\rangle$ can be expanded in a series:

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda E_n^{(2)} + \cdots$$
$$|n\rangle + |n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \lambda^2|n^{(2)}\rangle + \cdots$$

Where the correction terms, $E^{(i)}$ and $|n^{(i)}\rangle$, $i \geq 1$, are of successively smaller size. We keep track of this with λ but remember that λ will be set to one so it is not actually controlling the size. Notice also that the correction terms, $|n^{(i)}\rangle$, are not normalised.

Substituting these series into the TISE we get

$$\hat{H}|n\rangle = (\hat{H}_0 + \lambda\hat{H}')|n\rangle$$

$$= (\hat{H}_0 + \lambda \hat{H}')(|n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \lambda^2|n^{(2)}\rangle + \cdots)$$
$$= (E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots)(|n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \lambda^2|n^{(2)}\rangle + \cdots).$$

Equating terms of the same order in λ we see that

$$\lambda^0 : \qquad \hat{H}_0|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle$$
$$\lambda^1 : \qquad \hat{H}_0|n^{(1)}\rangle + \hat{H}'|n^{(0)}\rangle = E_n^{(0)}|n^{(1)}\rangle + E_n^{(1)}|n^{(0)}\rangle$$
$$\lambda^2 : \qquad \hat{H}_0|n^{(2)}\rangle + \hat{H}'|n^{(1)}\rangle = E_n^{(0)}|n^{(2)}\rangle + E_n^{(1)}|n^{(1)}\rangle + E_n^{(2)}|n^{(0)}\rangle.$$

Collecting terms by state these become

$$\lambda^0 : \qquad \hat{H}_0|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle$$
$$\lambda^1 : \qquad (\hat{H}_0 - E_n^{(0)})|n^{(1)}\rangle = (E_n^{(1)} - \hat{H}')|n^{(0)}\rangle$$
$$\lambda^2 : \qquad (\hat{H}_0 - E_n^{(0)})|n^{(2)}\rangle = (E_n^{(1)} - \hat{H}')|n^{(1)}\rangle + E_n^{(2)}|n^{(0)}\rangle.$$

The first of these, λ^0, tells us nothing new, it is simply a restatement of $|n^{(0)}\rangle$ as an eigenstate of $\hat{H}_0$.

Considering the first order terms take an inner product with the kth unperturbed eigenstate:

$$\langle k^{(0)}|(\hat{H}_0 - E_n^{(0)})|n^{(1)}\rangle = \langle k^{(0)}|(E_n^{(1)} - \hat{H}')|n^{(0)}\rangle. \tag{22.1}$$

The first term of the left hand side gives us

$$\langle k^{(0)}|\hat{H}_0|n^{(1)}\rangle = \langle n^{(1)}|\hat{H}_0^\dagger|k^{(0)}\rangle^*$$
$$= \langle n^{(1)}|\hat{H}_0|k^{(0)}\rangle^*$$
$$= E_k^{(0)}\langle n^{(1)}|k^{(0)}\rangle^*$$
$$= E_k^{(0)}\langle k^{(0)}|n^{(1)}\rangle.$$

Thus equation 22.1 becomes

$$(E_k^{(0)} - E_n^{(0)})\langle k^{(0)}|n^{(1)}\rangle = E_n^{(1)}\delta_{kn} - \langle k^{(0)}|\hat{H}'|n^{(0)}\rangle.$$

We chose k to be arbitrary, we now choose $k = n$ and this becomes

$$0 = E_n^{(1)} - H_{nn}$$

$$\boxed{\Delta E^{(1)} = E_n^{(1)} = H_{nn} = \langle n^{(0)}|\hat{H}'|n^{(0)}\rangle.}$$

This tells us the shift in the energy of the nth eigenvalue to first order. In words this tells us: the shift in energy induced by a perturbation is given to first order by the expectation value of the perturbation with respect to the unperturbed state.

Since $E_n^{(1)}$ represents the shift in energy to first order it is sometimes written as $\Delta E_n^{(1)}$ or ΔE_n. If instead
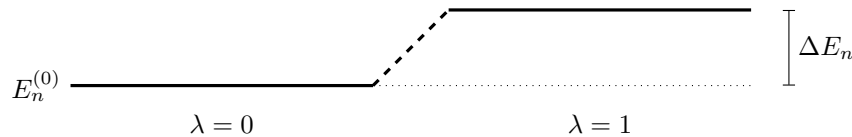

Figure 22.1: The first order effect of a perturbation on a non-degenerate energy level of energy $E_n^{(0)}$.

we choose $k \neq n$ then we have

$$(E_k^{(0)} - E_n^{(0)})\langle k^{(0)}|n^{(1)}\rangle = -\langle k^{(0)}|\hat{H}'|n^{(0)}\rangle, \qquad k \neq n.$$

Since $|n^{(1)}\rangle$ is just a state it can be expanded in terms of the unperturbed eigenbasis:

$$|n^{(1)}\rangle = \sum_m c^{(1)}_{nm}|m^{(0)}\rangle.$$

Taking an inner product with $\langle k^{(0)}|$ we find

$$\begin{aligned}\langle k^{(0)}|n^{(0)}\rangle &= \sum_m c^{(1)}_{nm}\langle k^{(0)}|m^{(0)}\rangle \\ &= \sum_m c^{(1)}_{nm}\delta_{km} \\ &= c^{(1)}_{nk}.\end{aligned}$$

Hence

$$(E^{(0)}_k - E^{(0)}_n)c^{(1)}_{nk} = -\langle k^{(0)}|\hat{H}'|n^{(0)}\rangle \implies c^{(1)}_{nk} = \frac{\langle k^{(0)}|\hat{H}'|n^{(0)}\rangle}{E^{(0)}_n - E^{(0)}_k} = \frac{H'_{kn}}{E^{(0)}_n - E^{(0)}_k}, \qquad k \neq n.$$

So we have now determined all of the coordinates of $|n^{(0)}\rangle$ apart from $c^{(1)}_{nn}$. However it turns out that if we choose the correction terms to be orthogonal to the unperturbed states then $c^{(1)}_{nn} = 0$. Thus we have fully specified $|n^{(1)}\rangle$ in the $\{|k^{(0)}\rangle\}$ basis and we find that

$$\boxed{|n^{(1)}\rangle = \sum_{k\neq n} \frac{\langle k^{(0)}|\hat{H}|n^{(0)}\rangle}{E^{(0)}_n - E^{(0)}_k}|k^{(0)}\rangle = \sum_{k\neq n} \frac{H'_{kn}}{E^{(0)}_n - E^{(0)}_k}|k^{(0)}\rangle.}$$

We say that the perturbation mixes the unperturbed states since the perturbed states have a small component from each unperturbed state. Unlike with the energy shift we have a potentially infinite sum to evaluate. This, along with the initial assumption that the series expansions exist, means that several things have to be 'small':

- For first order changes to the eigenstate to be small we require

$$|H'_{kn}| \ll |E^{(0)}_n - E^{(0)}_k| \forall k \neq n.$$

- For first order changes to the energy levels to be small we require

$$|E^{(1)}_n| \ll \min\left\{|E^{(0)}_n| - E^{(0)}_k\right\}.$$

Note that these conditions break down if there is degeneracy in the unperturbed system, for example some terms of the sum for $|n^{(1)}\rangle$ will become undefined if $E^{(0)}_k = E^{(0)}_n$ for $k \neq n$. However we only need the particular energy level for the shift that we are calculating to be non-degenerate for this analysis to be correct. Other energy levels can be degenerate but we will need to treat them specially when we come to calculate their energy shift.

Example 22.1. *Consider the potential*

$$V(x) = \begin{cases} V_0 \cos\left(\frac{\pi x}{2a}\right), & |x| \leq a, \\ \infty, & |x| > a. \end{cases}$$

Calculate the ground state energy to first order in perturbation theory. This corresponds to an infinite square well with the floor of the well bulging upwards in the centre. We can consider the Hamiltonian for this system to be the Hamiltonian for the square well plus a perturbative cosine term. We already know the solutions for the infinite square well:

$$E^{(0)}_n = \frac{\pi^2\hbar^2 n^2}{8ma^2},$$
$$u^{(0)}_n(x) = \frac{1}{\sqrt{a}}\cos\left(\frac{n\pi x}{2a}\right), \qquad n \text{ odd},$$

$$u_n^{(0)}(x) = \frac{1}{\sqrt{a}} \sin\left(\frac{n\pi x}{2a}\right), \qquad n \text{ even},$$

The perturbation is then

$$\hat{H}' = V_0 \cos\left(\frac{\pi x}{2a}\right).$$

This is 'small' as long as $V_0 \ll E_2^{(0)} - E_1^{(0)}$. The first order correction to the energy is then

$$\Delta E = E_1^{(1)} = H_{11}' = \int_{-\infty}^{\infty} u_1^{(0)} \hat{H}' u_1^{(0)} \, \mathrm{d}x = \frac{V_0}{a} \int_{-a}^{a} \cos^3\left(\frac{\pi x}{2a}\right) \mathrm{d}x = \frac{8V_0}{3\pi} \approx 0.85 V_0.$$

22.1   Higher Orders

Sometimes the first order correction is zero. If this is the case then we need to go to second order. Going back to the result for the second order terms in λ we had

$$(\hat{H}_0 - E_n^{(0)})|n^{(2)}\rangle = (E_n^{(1)} - \hat{H}')|n^{(1)}\rangle + E_n^{(2)}|n^{(0)}\rangle.$$

Taking an inner product with an unperturbed eigenstate gives

$$\langle k^{(0)}|(\hat{H}_0 - E_n^0)|n^{(2)}\rangle = \langle k^{(0)}|(E_n^{(1)} - \hat{H}')|n^{(1)}\rangle + E_n^{(2)}\delta_{kn}.$$

As before the first term on the left hand side can be shown to be $E_k^{(0)}\langle k^{(0)}|n^{(2)}\rangle$ so

$$(E_k^{(0)} - E_n^{(0)})\langle k^{(0)}|n^{(2)}\rangle = \langle k^{(0)}|(E_n^{(1)} - \hat{H}')|n^{(1)}\rangle + E_n^{(2)}\delta_{kn}.$$

In the case that $k = n$ we find that

$$E_n^{(2)} = \langle n^{(0)}|\hat{H}'|n^{(1)}\rangle - E_n^{(1)}\langle n^{(0)}|n^{(1)}\rangle$$

we choose the perturbative term $|n^{(1)}\rangle$ to be orthogonal to the eigenstate $|n^{(0)}\rangle$ leaving

$$= \langle n^{(0)}|\hat{H}'|n^{(1)}\rangle$$

$$= \sum_{m \neq n} c_{nm}^{(1)} \langle n^0|\hat{H}'|m^{(0)}\rangle$$

$$= \sum_{m \neq n} c_{nm}^{(1)} H_{nm}'$$

$$= \sum_{m \neq n} \frac{H_{mn}' H_{nm}'}{E_n^{(0)} - E_m^{(0)}}$$

$$= \sum_{m \neq n} \frac{|H_{mn}'|}{E_n^{(0)} - E_n^{(0)}}.$$

Here we have used the expansion of $|n^{(1)}\rangle$ in the $\{|m^{(0)}\rangle\}$ eigenbasis, the formula we found for $c_{nm}^{(1)}$, and the fact that H' is Hermitian so $H_{nm}' = H_{nm}'^*$.

We see that for the second order energy shift we have to calculate an infinite sum. To find the second order correction to the state we would need a double infinite sum and so on. Sometimes symmetry dictates that only a finite number of terms will be non-zero, sometimes we just have to truncate the sum after some number of terms.

22.2   Convergence

This section is non-examinable.

The series expansions that we come up with in perturbation theory often aren't actually convergent. This is because even though each term is small each correction contains multiple infinite sums which dominate and cause the series to diverge. Formally these series are asymptotic series. An asymptotic series for f is a series whose partial sums, that is the sum up to and including the nth term, evaluated at x can be made arbitrarily close to $f(x)$ by taking sufficiently large x. This differs from the usual definition of a convergent sum where the sequence of nth partial sums tends to $f(x)$ as $n \to \infty$ for all x.

23 Degenerate, Time Independent, Perturbation Theory

Suppose that the unperturbed Hamiltonian, $\hat{H}_0$, has one eigenvalue which is g-fold degenerate. We label the states such that the first g-states are degenerate so

$$\hat{H}_0|E_n^{(0)}\rangle = E^{(0)}|E_n^{(0)}\rangle, \qquad\qquad n = 1, \ldots, g,$$
$$\hat{H}_0|E_n^{(0)}\rangle = E_n^{(0)}|E_n^{(0)}\rangle, \qquad\qquad n > g.$$

Here $E^{(0)}$ is the degenerate eigenvalue and $E^{(0)} \neq E_n^{(0)}$ for $n > g$. All other energy levels are non-degenerate so $E_n^{(0)} \neq E_m^{(0)}$ for $n \neq m$ and $n, m > g$.

We can choose a basis such that the g degenerate states are mutually orthogonal. Any linear combination of degenerate states, such as

$$|E^{(0)}\rangle = \sum_{n=1}^{g} b_n |E_n^{(0)}\rangle,$$

will be another eigenstate of $\hat{H}_0$ with eigenvalue $E^{(0)}$.

In the non-degenerate case we assumed that the perturbed eigenstates were close to the unperturbed states. More formally there was a one-to-one correspondence between the unperturbed states, $\{n^{(0)}\}$, and the perturbed states $|n\rangle$, and it was possible to continuously move from a perturbed state to an unperturbed state by allowing $\lambda \to 0$. This cannot be assumed to be the case for a degenerate system because of the freedom we have in picking the basis for the degenerate subspace. In the limit $\lambda \to 0$ we cannot be sure which of the infinite set of degenerate eigenstates the series will converge to. We aim to find a linear combination of eigenstates, $\{|E^{(0)}\rangle\}$, such that these eigenstates are close to the perturbed states.

As before we start with

$$(\hat{H}_0 + \lambda \hat{H}')|E\rangle = E|E\rangle$$

where

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \cdots,$$
$$|E\rangle = |E^{(0)}\rangle + \lambda|E^{(1)}\rangle + \lambda^2|E^{(2)}\rangle + \cdots.$$

As before by equating coefficients of λ we see that

$$\lambda^0 : \qquad\qquad\qquad (\hat{H}_0 - E^{(0)})|E^{(0)}\rangle = 0,$$
$$\lambda^1 : \qquad\qquad (\hat{H}_0 - E^{(0)})|E^{(1)}\rangle + (\hat{H}' - E^{(1)})|E^{(0)}\rangle = 0,$$
$$\lambda^2 : \qquad (\hat{H}_0 - E^{(0)})|E^{(2)}\rangle + (\hat{H}' - E^{(1)})|E^{(1)}\rangle - E^{(2)}|E^{(0)}\rangle = 0.$$

The first of these simply restates that $|E^{(0)}\rangle$ is an eigenstate with eigenvalue $E^{(0)}$. For the second we take an inner product with the kth unperturbed state:

$$\langle E_k^{(0)}|(\hat{H}_0 - E^{(0)})|E^{(1)}\rangle + \langle E_k^{(0)}|(\hat{H}' - E^{(1)})|E^{(0)}\rangle.$$

As before we use the fact that $\langle E_k^{(0)}|\hat{H}_0 = E_k^{(0)}\langle E_k^{(0)}|$ for the first term and for the second term we use

$$\langle E^{(0)}| = \sum_{n=1}^{g} b_n |E_n^{(0)}\rangle$$

to get

$$(E_k^{(0)} - E^{(0)})\langle E_k^{(0)}|E^{(0)}\rangle + \sum_{n=1}^{g} b_n(H'_{kn} - E^{(1)}\delta_{kn})$$

where

$$H'_{kn} = \langle E_k^{(0)}|\hat{H}'|E_n^{(0)}\rangle, \qquad \text{and} \qquad \langle E_k^{(0)}|E_n^{(0)}\rangle = \delta_{kn}.$$
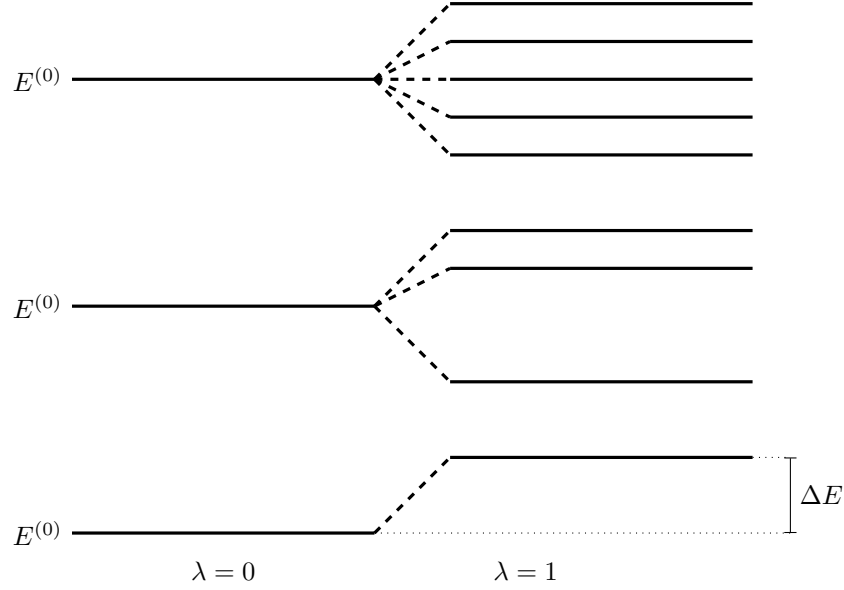
Figure 23.1: The three possible cases for roots of the eigenvalue problem with 5-fold degeneracy, top to bottom: all roots distinct, some roots equal, all roots equal.

We chose k arbitrarily and we now consider the case when $k \leq g$. In this case $E_k^{(0)} = E^{(0)}$ so this equation reduces to

$$\sum_{n=1}^{g}(H'_{kn} - E^{(1)}\delta_{kn})b_n = 0, \qquad k = 1, \ldots, g.$$

This is simply an eigenvalue problem in the g dimensional degenerate subspace. We can make this more explicit if we define a $g \times g$ matrix, H', with components H'_{kn}, and the g-dimensional identity matrix, $\mathbb{1}$, which has components δ_{kn}. Then this equation becomes

$$(H' - E^{(1)}\mathbb{1})\boldsymbol{b} = 0$$

where $\boldsymbol{b}$ is some g-dimensional eigenvector to be determined.

For this to have non-trivial solutions the determinant must be zero, so

$$\begin{vmatrix} H'_{11} - E^{(1)} & H'_{12} & \ldots & H'_{1g} \\ H'_{21} & H'_{22} - E^{(1)} & \ldots & H'_{2g} \\ \vdots & \vdots & \ddots & \vdots \\ H'_{g1} & H'_{g2} & \ldots & H'_{gg} - E^{(1)} \end{vmatrix} = 0.$$

The determinant is a g degree polynomial in $E^{(1)}$ and therefore has g roots. Each root corresponds to a set of coefficients, $\{b_n\}$. The roots are real since H' is a Hermitian matrix. There are three cases that we consider.

- *All roots distinct* – In this case the g-fold degeneracy is completely lifted at first order and the degenerate energy level splits into g distinct levels.

- *Some roots equal* – In this case the g-fold degeneracy is partly lifted at first order and the degenerate energy level splits into some distinct levels.

- *All roots equal* – In this case the degeneracy remains completely at first order, as long as the roots are non-zero there is still a first order energy shift, it just doesn't separate the energy levels.

These are shown in figure 23.1.

23.1 Special Cases

For a diagonal matrix the determinant is simply the product along the diagonal. If H' is diagonal then the determinant of $H' - E^{(1)}\mathbb{1}$ will be

$$(H'_{11} - E^{(1)})(H'_{22} - E^{(1)}) \cdots (H'_{gg} - E^{(1)}) = 0$$

which trivially has roots

$$\Delta E_n^{(1)} = E^{(0)} = H'_{nn} = \langle E_n^{(0)} | \hat{H}' | E_n^{(0)} \rangle, \qquad n = 1, \ldots, g.$$

We see that this is exactly the same as if $\hat{H}'$ was non-degenerate. Normally to diagonalise a matrix we need to know the eigenvalues and eigenvectors, which requires us to solve the eigenvalue problem so doesn't actually help us solve the eigenvalue problem. However here we can use the compatible observables theorem to know that the matrix is diagonal without actually computing it. Suppose we can find an observable, $\mathcal{A}$, represented by an operator, $\hat{A}$, such that

$$[\hat{H}_0, \hat{A}] = [\hat{H}', \hat{A}] = 0.$$

Notice that this doesn't necessarily mean that $[\hat{H}_0, \hat{H}'] = 0$, in fact we assume this isn't the case as if it was then we wouldn't need perturbation theory. We use the fact that $\hat{H}_0$ and $\hat{A}$ have a simultaneous eigenbasis, $\{|E_n^{(0)}, A_i\rangle\}$, and in this basis

$$\hat{H}_0 | E^{(0)}, A_i \rangle = E^{(0)} | E^{(0)}, A_i \rangle$$
$$\hat{A} | E^{(0)}, A_i \rangle = A_i | E^{(0)}, A_i \rangle.$$

If the eigenvalues, $\{A_i\}$, for $i = 1, \ldots, g$, are distinct then we choose these states as our basis states, that is

$$|E_n^{(0)}\rangle = |E^{(0)}, A_n\rangle, \qquad n = 1, \ldots, g.$$

We then use the fact that $\hat{H}'$ and $\hat{A}$ commute so

$$\langle E_k^{(0)} | [\hat{H}', \hat{A}] | E_n^{(0)} \rangle = 0.$$

Writing this out in full gives

$$\langle E_k^{(0)} | \hat{H}' \hat{A} | E_n^{(0)} \rangle - \langle E_k^{(0)} | \hat{A} \hat{H}' | E_n^{(0)} \rangle = (A_n - A_k) \langle E_k^{(0)} | \hat{H}' | E_n^{(0)} \rangle = (A_n - A_k) H'_{kn}.$$

Since this must always be zero and we assume that $A_n \neq A_k$ for $n \neq k$ we must have that $H'_{kn} = 0$ for $k \neq n$. Thus in this eigenbasis $\hat{H}'$ is diagonalised and so we can use the same solution as we would for the non-degenerate case.

If A_i are not distinct then we look for another new operator, $\hat{B}$, such that $\{\hat{H}_0, \hat{A}, \hat{B}\}$ and $\{\hat{H}', \hat{A}, \hat{B}\}$ are mutually commuting sets of operators, and so on.

Example 23.1. Consider a central potential, $\hat{V}(r)$, along with spin $1/2$ spin-orbit interaction, that is an interaction between the dipole of an electron and the magnetic field it sees due to moving in an electric field. The Hamiltonian for this system can be written as

$$\hat{H} = \frac{\hat{P}^2}{2m} + \hat{V}(r) + f(r)\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{S}} = \hat{H}_0 + \hat{H}'.$$

where $\hat{H}' = f(r)\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{S}}$. Since $\hat{H}'$ has $\hat{L}_x$ and $\hat{S}_x$ components we expect $[\hat{H}', \hat{L}_z] \neq 0 \neq [\hat{H}', \hat{S}_z]$. However we do have $[\hat{H}', \hat{J}_z] = 0$ where $\hat{\boldsymbol{J}} = \hat{\boldsymbol{L}} + \hat{\boldsymbol{S}}$. Thus in the coupled basis, $|n, \ell, s, j, m_j\rangle$, we can use non-degenerate perturbation theory to compute the first order energy shifts. Note that $[\hat{H}_0, \hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{S}}] = 0$ but $[\hat{H}_0, \hat{H}'] \neq 0$ due to the $f(r)$ term in $\hat{H}'$ and the differential operator in $\hat{H}_0$ due to $\hat{P}^2 = -\hbar^2 \nabla^2$.

---

24 Hydrogen Fine Structure

As an example of a real life use of perturbation theory we consider the fine structure of hydrogen. This results from considering relativistic effects. We won't derive these but we will state and explain what they correct for. Recall that the unperturbed Hamiltonian for a hydrogen like atom with Z protons is

$$\hat{H}_0 = \frac{\hat{P}^2}{2m} + \hat{V}(r), \qquad \text{where} \qquad \hat{V}(r) = -\frac{Ze^2}{4\pi\varepsilon_0 r}.$$

The eigenvalues are given by

$$E_n^{(0)} = -\frac{m}{2\hbar^2}\left(\frac{Ze^2}{4\pi\varepsilon_0}\right)^2\frac{1}{n^2} = -\frac{e^2}{2(4\pi\varepsilon_0)a_0}\frac{Z^2}{n^2} = -mc^2\frac{(Z\alpha)}{2n^2}$$

where a_0 is the Bohr radius and α is the fine structure constant defined by

$$a_0 = \frac{4\pi\varepsilon_0\hbar^2}{me^2}, \qquad \text{and} \qquad \alpha = \frac{e^2}{4\pi\varepsilon_0\hbar c} \approx \frac{1}{137}.$$

One commonly used unit in this context is the rydberg unit of energy defined by

$$1\,\mathrm{Ry} = \frac{e^2}{2(4\pi\varepsilon_0)a_0} = 13.6\,\mathrm{eV}.$$

The ground state of hydrogen is then $-1\,\mathrm{Ry} = -13.6\,\mathrm{eV}$.

24.1   Perturbed Hamiltonian

The perturbation to the Hamiltonian, $\hat{H}'$, can be further decomposed into three parts:

$$\hat{H}' = \hat{H}'_{\mathrm{KE}} + \hat{H}'_{\mathrm{SO}} + \hat{H}'_{\mathrm{Darwin}}.$$

The first term is a relativistic correction to the kinetic energy, which in special relativity is given by $(\gamma - 1)mc^2$. Expanding this we see that

$$(\gamma - 1)mc^2 = \frac{1}{2}mv^2 + \frac{3}{8}m\frac{v^4}{c^2} + \mathcal{O}v^6.$$

For the speeds that an electron travels this fourth order term is small but not entirely negligible. It leads to the perturbation

$$\hat{H}'_{\mathrm{KE}} = -\frac{\hat{P}^4}{8m^3c^2}.$$

The second term is the spin-orbit term which accounts for the interaction between the intrinsic magnetic dipole moment of the electron and the magnetic field that the electron sees due to moving in the electric field created by the nucleus. The form of this perturbation is

$$\hat{H}'_{\mathrm{SO}} = f(r)\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{S}}$$

where

$$f(r) = \frac{2}{m^2c^2r}\frac{\mathrm{d}V}{\mathrm{d}r}.$$

The third term, known as the Darwin term[12] is a relativistic correction to the potential energy given by

$$\hat{H}'_{\mathrm{Darwin}} = \frac{\hbar^2}{8m^2c^2}\nabla^2 V(\boldsymbol{r}) = \frac{\pi\hbar^2}{2m^2c^2}\frac{Ze^2}{4\pi\varepsilon_0}\delta(\boldsymbol{r})$$

where δ is the Dirac delta distribution.

24.2   Kinetic Energy Correction

The unperturbed energy level $E_n^{(0)}$ for some specific n is $2n^2$-fold degenerate. However since $\hat{H}'_{\mathrm{KE}}$ contains no spin components we know it commutes with $\hat{S}^2$ and $\hat{S}_z$. It can also be shown that $\hat{H}'_{\mathrm{KE}}$ commutes with $\hat{L}^2$ and $\hat{L}_z$. Therefore in the uncoupled basis, $\{|n, \ell, m_\ell, s, m_s\rangle\}$, $\hat{H}'_{\mathrm{KE}}$ will be diagonal. This means that if we work in the uncoupled basis we can use non-degenerate perturbation theory. Since $\hat{H}'_{\mathrm{KE}}$ has no effect on the spin components we drop s and m_s from the notation. The first order energy shift is

$$\Delta E_{\mathrm{KE}} = \langle n, \ell, m_\ell|\left(-\frac{\hat{P}^4}{8m^3c^2}\right)|n, \ell, m_\ell\rangle$$

---

[12] Named after Charles Galton Darwin who was the grandson of Charles Robert Darwin of evolution fame.

$$= -\frac{1}{2mc^2}\langle n, \ell, m_\ell | \hat{T}^2 | n, \ell, m_\ell \rangle.$$

Here we have used the fact that

$$\hat{T} = \frac{\hat{P}^2}{2m} \implies \hat{T}^2 = \frac{\hat{P}^4}{4m^2}.$$

Rather than try to compute the fourth derivative, $\nabla^2\nabla^2\Psi(\boldsymbol{r})$, we use the fact that $\hat{T} = \hat{H}_0 - \hat{V}(r)$ and therefore

$$\hat{T}^2 = \hat{H}_0^2 - \hat{H}_0\hat{V}(r) - \hat{V}(r)\hat{H}_0 + \hat{V}(r)$$

and then ΔE_{KE} is proportional to the expectation value of this:

$$\Delta E_{\mathrm{KE}} = -\frac{1}{2mc^2}(E_n^{(0)\,2} - 2E_n^{(0)}\langle \hat{V} \rangle_{n\ell m_\ell} + \langle \hat{V}^2 \rangle_{n\ell m_\ell}).$$

Since $\hat{V} \propto 1/r$ to evaluate the right hand side we simply need $\langle 1/r \rangle$ and $\langle 1/r^2 \rangle$ which can be shown to be

$$\left\langle \frac{1}{r} \right\rangle_n = \frac{Z}{a_0}\frac{1}{n^2}, \qquad \text{and} \qquad \left\langle \frac{1}{r^2} \right\rangle_{n\ell} = \frac{Z^2}{a_0^2}\frac{1}{n^3(\ell+1/2)}.$$

So we find that

$$\Delta E_{\mathrm{KE}} = -E_n^{(0)}\frac{(Z\alpha)^2}{n^2}\left[\frac{3}{4} - \frac{n}{\ell+1/2}\right].$$

As a quick sanity check we see that this is proportional to $\alpha^2 E_n^{(0)} \approx 5 \times 10^{-5} E_n^{(0)}$ so the energy shift is indeed small compared to the size of the energy levels, $E_n^{(0)}$.

24.3   Spin-Orbit Correction

For given values of n and ℓ the $E_n^{(0)}$ energy level is $2(2\ell+1)$-fold degenerate. This is due to the $2\ell+1$ different values of m_ℓ for each value of ℓ and the two possible values of m_s, i.e. $\pm 1/2$.

However if we work in the coupled basis, $\{n, \ell, s, j, m_j\}$, then we can use non-degenerate theory. To see that this is valid first note that

$$\hat{H}'_{\mathrm{SO}} = f(r)\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{S}} = \frac{1}{2}f(r)[\hat{J}^2 - \hat{L}^2 - \hat{S}^2]$$

and

$$[\hat{J}^2 - \hat{L}^2 - \hat{S}^2]|n, \ell, s, j, m_j\rangle = [j(j+1) - \ell(\ell+1) - s(s+1)]\hbar^2|n, \ell, s, j, m_j\rangle.$$

Since this is just an eigenvalue equation we see that $\hat{H}'_{\mathrm{SO}}$ is diagonal in the coupled basis. Thus the energy shift is simply the expectation value in this basis:

$$\begin{aligned}\Delta E_{\mathrm{SO}} &= \langle n, \ell, s, j, m_j | \hat{H}'_{\mathrm{SO}} | n, \ell, s, j, m_j \rangle \\ &= \frac{1}{2}[j(j+1) - \ell(\ell+1) - s(s+1)]\hbar^2\langle f(r) \rangle.\end{aligned}$$

This simplifies since we are considering an electron we know that $s = 1/2$ and by the angular momentum addition theorem this means that for a given value of ℓ we have $j = \ell + 1/2$ or $j = \ell - 1/2$. This means that a state of given n and ℓ is split into a doublet by the spin-orbit interaction. Since $f(r)$ is independent of ϑ and φ and of the spin the expectation value can be written as

$$\langle f(r) \rangle = \frac{1}{2m^2c^2}\frac{Ze^2}{4\pi\varepsilon_0}\int_0^\infty \frac{1}{r^3}|R_{n\ell}(r)|^2 r^2 \,\mathrm{d}r\,.$$

Note that the $1/r^3$ comes from f and the r^2 comes from the integration measure in polar coordinates. It can be shown that

$$\left\langle \frac{1}{r^3} \right\rangle_{n\ell} = \frac{Z^3}{a_0^3}\frac{1}{n^3\ell(\ell+1/2)(\ell+1)}.$$

For $\ell \neq 0$ we then find

$$\Delta E_{\mathrm{SO}} = \begin{cases} -E_n^{(0)}\frac{(Z\alpha)^2}{2n}\frac{\ell}{\ell(\ell+1/2)(\ell+1)}, & j = \ell + \frac{1}{2}, \\ E_n^{(0)}\frac{(Z\alpha)^2}{2n}\frac{1}{\ell(\ell+1/2)}, & j = \ell - \frac{1}{2}. \end{cases}$$

24.4   Darwin Correction

The Darwin correction is non-zero only at $\boldsymbol{r} = \boldsymbol{0}$ as it is proportional to $\delta(\boldsymbol{r})$. The radial function, $R_{n\ell}(r)$, vanishes at the origin when $\ell \neq 0$. Therefore we only need to consider this term at the origin when $\ell = 0$ and consequently $m_\ell = 0$. This term also contains no spin components so has no effect on the spin so again we drop s and m_s from the notation. We find that

$$\begin{aligned}
\Delta E_{\mathrm{Darwin}} &= \frac{\pi\hbar^2}{2m^2c^2}\frac{Ze^2}{4\pi\varepsilon_0}\langle n,0,0|\delta(\boldsymbol{r})|n,0,0\rangle \\
&= \frac{\pi\hbar^2}{2m^2c^2}\frac{Ze^2}{4\pi\varepsilon_0}\int u_{n00}^*(\boldsymbol{r})\delta(\boldsymbol{r})u_{n00}\,\mathrm{d}^3r \\
&= \frac{\pi\hbar^2}{2m^2c^2}\frac{Ze^2}{4\pi\varepsilon_0}|u_{n00}(\boldsymbol{0})|^2
\end{aligned}$$

where in the last step we used the sifting property of the Dirac delta distribution. This simplifies as

$$|u_{n00}(\boldsymbol{0})| = \left|Y_0^0(\vartheta,\varphi)\right|^2|R_{n\ell}(0)|^2 = \frac{1}{4\pi}|R_{n\ell}(0)|^2 = \frac{Z^3}{\pi a_0^3 n^3}.$$

Hence

$$\Delta E_{\mathrm{Darwin}} = -E_n^{(0)}\frac{(Z\alpha)^2}{n}.$$

24.5   Total Correction

Combining all three terms it can be shown that

$$\Delta E_{nj} = E_n^{(0)}\frac{(Z\alpha)^2}{n^2}\left[\frac{n}{j+1/2} - \frac{3}{4}\right].$$
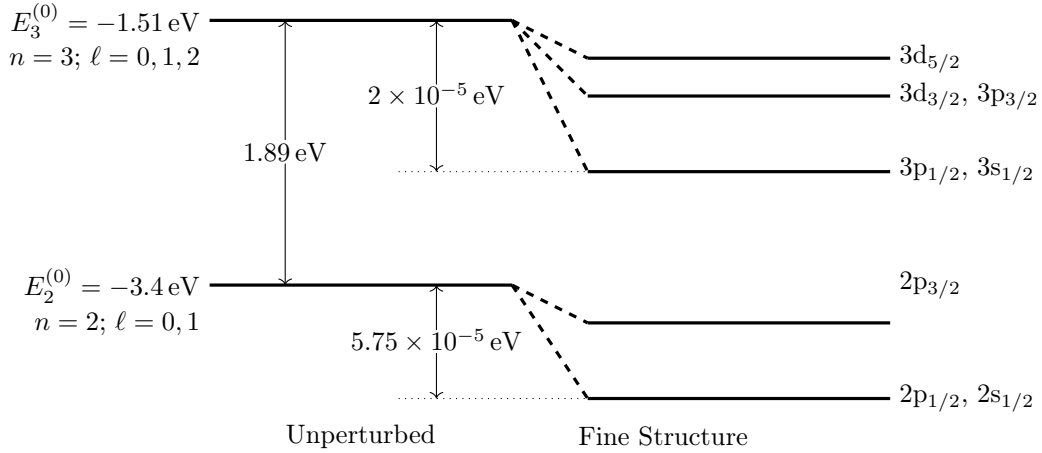

Figure 24.1: The fine structure of the $n = 2, 3$ energy levels of hydrogen (not to scale).

The resulting energy levels that occur when we account for relativistic effects are shown (not to scale) in figure 24.1. Since $\Delta E \propto \alpha^2$ the energy shifts are therefore on the order of 10^{-4} eV. The degeneracy between the states with $\ell = j \pm 1/2$, such as $2\mathrm{p}_{1/2}$ and $2\mathrm{s}_{1/2}$, is lifted if we take into account effects from quantum electrodynamics (QED). This shift, called the Lamb shift, was discovered experimentally in 1947 by Lamb and Rutherford. In atomic hydrogen this shift is another two orders of magnitude smaller than the relativistic shifts, on the order of 4×10^{-6} eV.

24.6   Term Notation

Term notation, also know as Russell–Saunders notation is a way of denoting a particular state with orbital angular momentum ℓ, spin s, and total angular momentum j. Such states are denoted

$$^{2s+1}\mathrm{l}_j.$$

ℓ	0	1	2	3	4	5	6	7
l	s	p	d	f	g	h	i	k

Table 24.1: The orbital angular momentum quantum number, ℓ, and the corresponding letter used in term notation. From f onwards the letters are simply alphabetical skipping j. The first few letters come from the names for the corresponding spectral lines, namely sharp, primary, diffuse, and fundamental

Here l is a letter corresponding to a particular value of ℓ as given in table 24.1 If the principle quantum number, n, is important then it is prepended:

$$n^{2s+1}\mathrm{l}_j.$$

Since for one electron atoms like hydrogen the spin is always $s = 1/2$ often the spin is dropped from the notation.

For example the ground state of hydrogen has $n = 1$, $\ell = 0$, $s = 1/2$, and $j = 1/2$. Therefore the corresponding state can be denoted as any of the following:

$$1^2\mathrm{s}_{1/2}, \qquad 1\mathrm{s}_{1/2}, \qquad {}^2\mathrm{s}_{1/2}, \qquad \mathrm{s}_{1/2} \qquad \text{or} \qquad \mathrm{s}.$$

25   Helium Atom

25.1   The Hamiltonian

Helium, He, is the second element on the periodic table. It has two protons and two electrons. We label these electrons 1 and 2. The Hamiltonian for a He atom is

$$\hat{H} = \hat{H}_1 + \hat{H}_2 + \frac{e^2}{4\pi\varepsilon|\boldsymbol{r_1} - \boldsymbol{r_2}|}$$

where

$$\hat{H}_i = \frac{\hat{P}_i^2}{2m} - \frac{2e^2}{4\pi\varepsilon_0}.$$

This Hamiltonian is the sum of the a Hamiltonian for hydrogen like atom as well as an electron–electron interaction term. Notice that the Hamiltonian is symmetric under exchanging of the two electrons. That is if we change all the terms labelled 1 to be labelled 2 and vice versa the Hamiltonian wouldn't change. This is important as electrons are identical so it should be able to swap them without changing the physics. Our goal in this section is to apply perturbation theory to find the energy of the ground state.

25.2   Electron Wave Functions

Before we can apply perturbation theory we need to know the ground state of the unperturbed system. We first consider the spin components. Recall that since electrons are spin $1/2$ particles the entire system is thus a spin 1 system. We will use a condensed notation where

$$\alpha = |\uparrow\rangle, \qquad \beta = |\downarrow\rangle, \qquad \chi_{SM_S} = |s_1 = 1/2, s_2 = 1/2, S, M_S\rangle.$$

We will also use subscripts for α and β to denote the electron that the state corresponds to. As well as this tensor products will simply be written by putting the two vectors next to each other without the usual $\otimes$ symbol.

Recall that there are three triplet states:

$$\chi_{11} = \alpha_1\alpha_2, \qquad \chi_{10} = \frac{1}{\sqrt{2}}[\alpha_1\beta_2 + \beta_1\alpha_2], \qquad \text{and} \qquad \chi_{1,-1} = \beta_1\beta_2,$$

and one singlet state:

$$\chi_{00} = \frac{1}{\sqrt{2}}[\alpha_1\beta_2 - \beta_1\alpha_2].$$

Notice the triplet states are symmetric under exchange of the two electrons whereas the singlet state is antisymmetric. This is important because the spin statistics theorem tells us that the symmetry of a wave function for the whole system must be antisymmetric under exchange of the two electrons. That is

$$\Psi(1,2) = \psi(\boldsymbol{r_1}, \boldsymbol{r_2})\chi = -\Psi(2,1).$$

We start by treating neglecting the electron–electron Coulomb interaction and consider the system as two independent hydrogenic atoms with $Z = 2$ which have a combined Hamiltonian

$$\hat{H}_0 = \hat{H}_1 + \hat{H}_2.$$

We already know the eigenfunctions for $\hat{H}_i$ are

$$\hat{H}_i u_{n_i \ell_i m_{\ell_i}}(\boldsymbol{r_i}) = E_{n_i} u_{n_i \ell_i m_{\ell_i}}(\boldsymbol{r_i}).$$

Since the two Hamiltonians only act on their respective electron's:

$$\hat{H}_i u_{n_1 \ell_1 m_{\ell_1}}(\boldsymbol{r_1}) u_{n_2 \ell_2 m_{\ell_2}}(\boldsymbol{r_2}) = E_{n_i} u_{n_1 \ell_1 m_{\ell_1}}(\boldsymbol{r_1}) u_{n_2 \ell_2 m_{\ell_2}}(\boldsymbol{r_2})$$

wave function the action of the total Hamiltonian on the combined state is simply

$$\begin{aligned}
\hat{H}_0 u_{n_1 \ell_1 m_{\ell_1}}(\boldsymbol{r_1}) u_{n_2 \ell_2 m_{\ell_2}}(\boldsymbol{r_2}) &= (\hat{H}_1 + \hat{H}_2) u_{n_1 \ell_1 m_{\ell_1}}(\boldsymbol{r_1}) u_{n_2 \ell_2 m_{\ell_2}}(\boldsymbol{r_2}) \\
&= (E_{n_1} + E_{n_2}) u_{n_1 \ell_1 m_{\ell_1}}(\boldsymbol{r_1}) u_{n_2 \ell_2 m_{\ell_2}}(\boldsymbol{r_2}) \\
&= E_n u_{n_1 \ell_1 m_{\ell_1}}(\boldsymbol{r_1}) u_{n_2 \ell_2 m_{\ell_2}}(\boldsymbol{r_2}).
\end{aligned}$$

So the energy eigenvalues are given by

$$E_n = E_{n_1} + E_{n_2}, \qquad \text{where} \qquad E_{n_i} = -\frac{m}{2\hbar^2}\left(\frac{2e^2}{4\pi\varepsilon_0}\right)^2 \frac{1}{n_i^2}.$$

25.3   Unperturbed Ground State

We are only interested in the ground state, that is $n = 1$, which corresponds to both constituent systems also being in their ground states, $n_i = 1$. The energy of the ground state, ignoring the effects of the perturbation, is

$$E_{n=1} = E_{n_1=1} + E_{n_2=1} = 2E_{n_2=1}.$$

Now we use the fact that for hydrogen with $Z = 1$ the ground state energy is $-13.6\,\text{eV}$ and helium has the same form of energy levels but with $Z = 2$. Since Z appears squared in the energy level we expect $E_{n_i} = 4 \cdot (-13.6\,\text{eV})$. Thus the total energy of the ground state is

$$E_n = 2 \cdot 4 \cdot (-13.6\,\text{eV}) = -108.8\,\text{eV}.$$

Experimentally we actually measure the ground state energy to be $-78.957\,\text{eV}$ which is higher. The reason this is higher is the repulsive force between the electrons which we have not yet considered.

The wave function in the ground state has $n_1 = n_2 = 1$ and $\ell_1 = \ell_2 = m_{\ell_1} = m_{\ell_2} = 0$ which gives the joint spatial wave function in the ground state

$$\psi_{\text{GS}}(\boldsymbol{r_1}, \boldsymbol{r_2}) = u_{100}(\boldsymbol{r_1}) u_{100}(\boldsymbol{r_2}).$$

At this point we remark that the electronic configuration of He is typically denoted 1s^2 where the 1 refers to the value of n, s corresponds to $\ell = 1$ and 2 tells us that there are two electrons in this state. Notice that the ground state wave function is symmetric under exchange of the electrons and therefore for the total ground state wave function, $\Psi_{\text{GS}} = \psi_{\text{GS}}\chi_{\text{GS}}$, to be antisymmetric χ_{GS} must be antisymmetric. Therefore we conclude that $\chi_{\text{GS}} = \chi_{00}$ as this is the only antisymmetric spin state. The total ground state wave function is then

$$|\Psi_{\text{GS}}\rangle \to u_{100}(\boldsymbol{r_1}) u_{100}(\boldsymbol{r_2})\chi_{00}.$$

25.4   The Perturbation

We are now in a position to apply perturbation theory to estimate the ground state energy. As usual we start with the Hamiltonian

$$\hat{H} = \hat{H}_0 + \hat{H}'$$

where

$$\hat{H}_0 = \hat{H}_1 + \hat{H}_2, \qquad \text{and} \qquad \hat{H}' = \frac{e^2}{4\pi\varepsilon_0 |\boldsymbol{r_1} - \boldsymbol{r_2}|}.$$

For brevity of notation let $r_{12} = |\boldsymbol{r_1} - \boldsymbol{r_2}|$. The first order correction to the ground state is then

$$
\begin{aligned}
\Delta E_1 &= \langle \Psi_{\text{GS}} | \hat{H}' | \Psi_{\text{GS}} \rangle \\
&= \langle \psi, \chi_{00} | \hat{H}' | \psi, \chi_{00} \rangle \\
&= (\langle \psi | \otimes \langle \chi_{00} |) \hat{H}' (|\psi\rangle \otimes |\chi_{00}\rangle) \\
&= \langle \psi | \hat{H}' | \psi \rangle \otimes \langle \chi_{00} | \hat{H}' | \chi_{00} \rangle \\
&= \langle \psi | \hat{H}' | \psi \rangle \otimes \langle \chi_{00} | \chi_{00} \rangle \\
&= \langle \psi | \hat{H}' | \psi \rangle
\end{aligned}
$$

Here we have used the fact that $\hat{H}'$ has no spin components and therefore doesn't act on $|\chi_{00}\rangle$ so assuming that $|\chi_{00}\rangle$ is properly normalised the spin terms become one and we drop them. Next we use the full form of the 1s electron wave function:

$$u_{100}(\boldsymbol{r}) = \frac{1}{\sqrt{\pi}} \left(\frac{Z}{a_0} \right)^{3/2} e^{-Zr/a_0}.$$

Which gives the integral

$$\Delta E_1 = \frac{e^2}{4\pi\varepsilon_0} \left(\frac{Z^3}{\pi a_0^3} \right)^2 \int \frac{1}{r_{12}} \exp\left[-\frac{2Z(r_1 + r_2)}{a_0} \right] \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \,.$$

This integral can then be shown to give

$$\Delta E_1 = \frac{5}{4} Z \, \text{Ry} = \frac{5}{2} \, \text{Ry} = 34 \, \text{eV}.$$

So to first order we estimate the ground state energy to be

$$E_1 = -108.8 \, \text{eV} + 34 \, \text{eV} = -74.8 \, \text{eV} = -5.5 \, \text{Ry}.$$

This is respectably close to the measured value of $-78.957 \, \text{eV}$.

One thing we do have to consider is if perturbation theory is even valid in this case. The perturbation is on the order of $10 \, \text{eV}$. This is only one order of magnitude smaller than the unperturbed energy which on the order of $100 \, \text{eV}$. As well as this the fine structure corrections will be on the order of $10 \, \text{eV}$. One justification for using perturbation theory is simply how close the first order energy is to the experimental value.

26   Multi-Electron Atoms

26.1   Excited States of Helium

Recall that the Hamiltonian for Helium is

$$\hat{H} = \hat{H}_1 + \hat{H}_2 + \frac{e^2}{4\pi\varepsilon_0 |\boldsymbol{r_1} - \boldsymbol{r_2}|}$$

where

$$\hat{H}_i = \frac{\hat{P}_i^2}{2m} - \frac{2e^2}{4\pi\varepsilon_0 r_i}.$$

Ignoring the electron repulsion term the first excited state for helium corresponds to exciting an electron to the first hydrogen-like excited state. That is one of n_1 or n_2 becomes 2 and the corresponding angular orbital momentum, ℓ_i becomes 0 or 1. The corresponding electron configuration is then 1s2s or 1s2p with $\ell_i = 0, 1$ respectively.

Since the electrons are indistinguishable it doesn't make sense to say which electron is in the excited state. Instead we consider a superposition of each electron in the excited state. This allows us to construct the spatial wave function

$$\psi(\boldsymbol{r_1}, \boldsymbol{r_2}) = \frac{1}{\sqrt{2}}[u_{2\ell m_\ell}(\boldsymbol{r_1})u_{100}(\boldsymbol{r_2}) \pm u_{100}(\boldsymbol{r_1})u_{2\ell m_\ell}(\boldsymbol{r_2})].$$

This is symmetric if we choose $+$ and antisymmetric if we choose $-$. From this we can construct overall antisymmetric wave functions, the antisymmetry being dictated by spin statistics:

$$\Psi^{\text{singlet}}_{LM_LSM_S}(1,2) = \frac{1}{\sqrt{2}}[u_{2\ell m_\ell}(\boldsymbol{r_1})u_{100}(\boldsymbol{r_2}) + u_{100}(\boldsymbol{r_1})u_{2\ell m_\ell}(\boldsymbol{r_2})] \otimes \chi_{S=0}$$

$$\Psi^{\text{triplet}}_{LM_LSM_S}(1,2) = \frac{1}{\sqrt{2}}[u_{2\ell m_\ell}(\boldsymbol{r_1})u_{100}(\boldsymbol{r_2}) - u_{100}(\boldsymbol{r_1})u_{2\ell m_\ell}(\boldsymbol{r_2})] \otimes \chi_{S=1}$$

Here L denotes the total angular momentum quantum number. It is 0 for the 1s2s state and 1 for the 1s2p state.

We can see the degeneracy of various states by considering all possible combinations of quantum numbers allowed by the angular momentum addition theorem. First consider the case $S = 0$. There are two options for L, it can be 0 or 1. Recall that by the addition of angular momentum theorem the total angular momentum number, J, can be any value from $L + S$ to $|L - S|$ decreasing in integer steps. The degeneracy, g_J, of this state is then $2J + 1$. The possible values of the quantum numbers the first two excited states are given in table 26.1. We can use perturbation theory to calculate the energy shift due

S	L	J	g_J	Terms
0	0	0	1	$^1S_0$
	1	1	3	$^1P_1$
1	0	1	3	$^3S_1$
	1	0	1	$^3P_0$
	1	1	3	$^3P_1$
	1	2	5	$^3P_2$

Table 26.1: The degeneracy of the first few excited states of helium.

to the electron-electron interaction. The perturbation commutes with $\hat{L}_z$ and therefore is independent of m_ℓ. This means we can calculate for a specific m_ℓ and use the result for any m_ℓ. We choose $m_\ell = 0$. As before the spin components commute with the perturbation and their inner product is one. The energy shift is then

$$\Delta E^{\text{singlet}}_{\text{triplet}} = \frac{e^2}{2(4\pi\varepsilon_0)} \int [u_{2\ell 0}(\boldsymbol{r_1})u_{100}(\boldsymbol{r_2}) \pm u_{100}(\boldsymbol{r_1})u_{2\ell 0}(\boldsymbol{r_2})]^* \frac{1}{r_{12}}$$
$$[u_{2\ell 0}(\boldsymbol{r_1})u_{100}(\boldsymbol{r_2}) \pm u_{100}(\boldsymbol{r_1})u_{2\ell 0}(\boldsymbol{r_2})]\, \mathrm{d}^3r_1\, \mathrm{d}^3r_2\,.$$

Here $r_{12} = |\boldsymbol{r_1} - \boldsymbol{r_2}|$. Expanding this gives

$$\Delta E^{\text{singlet}}_{\text{triplet}} = \frac{e^2}{2(4\pi\varepsilon_0)} \left[\int \frac{1}{r_{12}} u^*_{2\ell 0}(\boldsymbol{r_1})u^*_{100}(\boldsymbol{r_2})u_{2\ell 0}(\boldsymbol{r_1})u_{100}(\boldsymbol{r_2})\, \mathrm{d}^3r_1\, \mathrm{d}^3r_2 \right.$$
$$\pm \int \frac{1}{r_{12}} u^*_{2\ell 0}(\boldsymbol{r_1})u^*_{100}(\boldsymbol{r_2})u_{100}(\boldsymbol{r_1})u_{2\ell 0}(\boldsymbol{r_2})\, \mathrm{d}^3r_1\, \mathrm{d}^3r_2$$
$$\pm \int \frac{1}{r_{12}} u^*_{100}(\boldsymbol{r_1})u^*_{2\ell 0}(\boldsymbol{r_2})u_{2\ell 0}(\boldsymbol{r_1})u_{100}(\boldsymbol{r_2})\, \mathrm{d}^3r_1\, \mathrm{d}^3r_2$$
$$\left. + \int \frac{1}{r_{12}} u^*_{100}(\boldsymbol{r_1})u^*_{2\ell 0}(\boldsymbol{r_2})u_{100}(\boldsymbol{r_1})u_{2\ell 0}(\boldsymbol{r_2})\, \mathrm{d}^3r_1\, \mathrm{d}^3r_2 \right]$$
$$= \frac{e^2}{2(4\pi\varepsilon_0)} \left[\int \frac{1}{r_{12}} |u_{2\ell 0}(\boldsymbol{r_1})|^2 |u_{100}(\boldsymbol{r_2})|^2\, \mathrm{d}^3r_1\, \mathrm{d}^3r_2 \right.$$

$$\pm \int \frac{1}{r_{12}} u_{2\ell 0}^*(\boldsymbol{r_1}) u_{100}^*(\boldsymbol{r_2}) u_{100}(\boldsymbol{r_1}) u_{2\ell 0}(\boldsymbol{r_2}) \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2$$

$$\pm \int \frac{1}{r_{12}} u_{100}^*(\boldsymbol{r_1}) u_{2\ell 0}^*(\boldsymbol{r_2}) u_{2\ell 0}(\boldsymbol{r_1}) u_{100}(\boldsymbol{r_2}) \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2$$

$$+ \int \frac{1}{r_{12}} |u_{100}(\boldsymbol{r_1})|^2 |u_{2\ell 0}(\boldsymbol{r_2})|^2 \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \Bigg]$$

We now use the fact that the perturbation is symmetric under exchagne of the two electrons which allows us to swap $\boldsymbol{r_1}$ and $\boldsymbol{r_2}$ as long as we do it for every occurence within a term

$$\Delta E_{\text{triplet}}^{\text{singlet}} = \frac{e^2}{2(4\pi\varepsilon_0)} \left[\int \frac{1}{r_{12}} |u_{2\ell 0}(\boldsymbol{r_1})|^2 |u_{100}(\boldsymbol{r_2})|^2 \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \right.$$

$$\pm \int \frac{1}{r_{12}} u_{2\ell 0}^*(\boldsymbol{r_1}) u_{100}^*(\boldsymbol{r_2}) u_{100}(\boldsymbol{r_1}) u_{2\ell 0}(\boldsymbol{r_2}) \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2$$

$$\pm \int \frac{1}{r_{12}} u_{100}^*(\boldsymbol{r_2}) u_{2\ell 0}^*(\boldsymbol{r_1}) u_{2\ell 0}(\boldsymbol{r_2}) u_{100}(\boldsymbol{r_1}) \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2$$

$$\left. + \int \frac{1}{r_{12}} |u_{100}(\boldsymbol{r_2})|^2 |u_{2\ell 0}(\boldsymbol{r_1})|^2 \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \right]$$

Notice now the repeated terms which we can combine and cancel the factor of $1/2$ in the constant at the start

$$\Delta E_{\text{triplet}}^{\text{singlet}} = \frac{e^2}{4\pi\varepsilon_0} \left[\int \frac{1}{r_{12}} |u_{2\ell 0}(\boldsymbol{r_1})|^2 |u_{100}(\boldsymbol{r_2})|^2 \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \right.$$

$$\left. \pm \int \frac{1}{r_{12}} u_{2\ell 0}^*(\boldsymbol{r_1}) u_{100}^*(\boldsymbol{r_2}) u_{100}(\boldsymbol{r_1}) u_{2\ell 0}(\boldsymbol{r_2}) \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \right]$$

The first term looks similar to the ground state shift and is due to the interaction between the electrons. The second term has no classical interpretation. It is called the exchange contribution. Its sign depends on whether the total spin is 0 or 1. What we see is that although the perturbation doesn't depend on the spin the energy shift does as the symmetry of the wave function depends on the spin.

In general the energy shift for the energy level with quantum numbers n and ℓ is

$$\Delta E_{n\ell}^{\text{singlet}} = A_{n\ell} + B_{n\ell}, \qquad \text{or} \qquad \Delta E_{n\ell}^{\text{triplet}} = A_{n\ell} - B_{n\ell}.$$

Here $A_{n\ell}$ represents the electron interaction and $B_{n\ell}$ represents the exchange contribution. Clearly $A_{n\ell} > 0$ and it turns out that $B_{n\ell} > 0$ to and $B_{n\ell} < A_{n\ell}$ so we find that

$$\Delta_{n\ell}^{\text{triplet}} < \Delta_{n\ell}^{\text{singlet}}.$$

This should be of no suprise to chemists as it essentially says that the p orbitals are of higher energy than the s orbitals. This is shown if figure 26.1. The reason for this can be seen if we look at the original spatial wave functions. The antisymmetric spatial wave function is zero if $\boldsymbol{r_1} = \boldsymbol{r_2}$ and very small if they are close. We interpret this as the electrons being, on average, further apart in this state meaning that the repulsive effect of the electron interaction is smaller in this state and hence the energy shift is smaller.

26.2   Multi-Electron Atoms

Consider now a generic atom with a nuclear charge of Ze and N electrons. As usual we assume the nucleus is infinitely heavy and neglect all interactions apart from the Coulomb interactions. The Hamiltonian for this system is

$$\hat{H} = \sum_{i=1}^{N} \left[\frac{\hat{P}_i}{2m} - \frac{Ze^2}{4\pi\varepsilon_0 r_i} \right] + \sum_{i>j=1}^{N} \frac{e^2}{4\pi\varepsilon_0 r_{ij}}.$$

Here $r_{ij} = |\boldsymbol{r_i} - \boldsymbol{r_j}|$. Notice that the final sum simply sums over pairs of electrons without double counting. The goal is to solve the TISE but as is often the case this isn't possible so we approximate.

One way of approximating is called the central field approximation. We treat each electron as an independent particle moving in a central potential which is a function of r, the distance from the nucleus
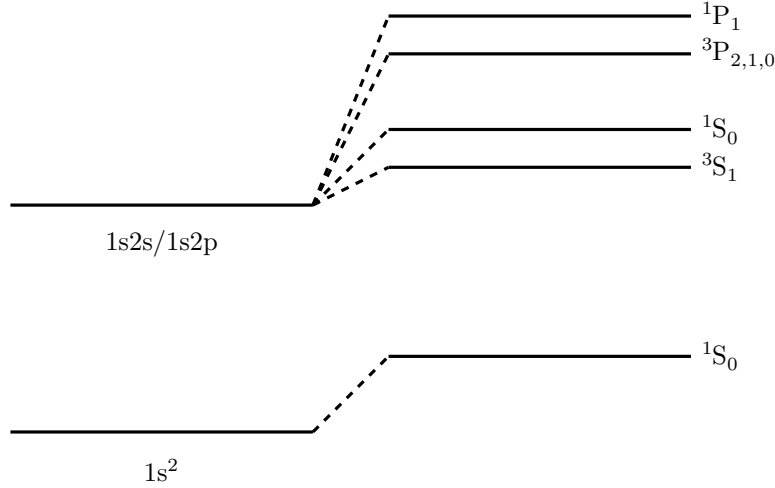
Figure 26.1: The effect of Coulomb repulsion on the helium spectrum. Notice that this shows two possible excited states, 1s2s and 1s2p.

(which is assumed to be a point mass). The task is to come up with a potential that approximates the true potential. We start by considering the limiting cases of $r \to 0$ and $r \to \infty$. In the limit of $r \to 0$ there will be no other electrons nearby and therefore we expect the only potential to be the interaction between the nucleus and electron:

$$\lim_{r \to 0} V(r) = -\frac{Ze^2}{4\pi\varepsilon_0 r}.$$

In the limit that $r \to \infty$ we expect all of the other $N - 1$ electrons to be between the electron and the nucleus diluting the effective potential:

$$\lim_{r \to \infty} V(r) = -\frac{[Z - N + 1]e^2}{4\pi\varepsilon_0 r}.$$

Finding the intermediate form of V is not easy and is generally done by an iterative process called the Hartree–Fock method or the self-consistent field method. We will not discuss this here.

We write the full Hamiltonian as

$$\hat{H} = \hat{H}_c + \hat{H}'$$

where $\hat{H}_c$ corresponds to the central field approximation:

$$\hat{H}_c = \sum_{i=1}^{N} \left[\frac{\hat{P}_i^2}{2m} + V(r_i) \right] = \sum_{i=1}^{N} \hat{H}_i$$

where

$$\hat{H}_i = \frac{\hat{P}_i^2}{2m} + V(r_i).$$

$\hat{H}'$ is the rest of the Hamiltonian.

Considering only the central field Hamiltonian we can view it as the sum of N single particle Hamiltonians, $\hat{H}_i$, which will have energy eigenfunctions satisfying

$$\hat{H}_i u_{n_i \ell_i m_{\ell_i}}(\boldsymbol{r_i}) = E_{n_i \ell_i} u_{n_i \ell_i m_{\ell_i}}(\boldsymbol{r_i}).$$

The energy eigenvalues are independent of m_{ℓ_i} due to the spherical symmetry of $\hat{H}_i$ but they depend on ℓ_i. The fact that the Hydrogen energy levels *don't* depend on ℓ turns out to be an exception rather than a rule. The spherical symmetry of the single particle Hamiltonians means we can decompose the energy eigenstates as

$$u_{n_i \ell_i m_{\ell_i}}(\boldsymbol{r_i}) = R_{n_i \ell_i}(r_i) Y_{\ell_i}^{m_{\ell_i}}(\vartheta_i, \varphi_i).$$

We can then account for spin by including a spin term, $\chi_{s_i m_{s_i}}$ where, since these are single electrons $s_i = 1/2$ and so $m_{s_i} = \pm 1/2$. So the total wave function is

$$u_{n_i \ell_i m_{\ell_i} s_i m_{s_i}}(i) = u_{n_i \ell_i m_{\ell_i}}(\boldsymbol{r_i}) \otimes \chi_{s_i m_{s_i}}.$$

The total energy of the atom in this approximation is simply the sum of the individual energies:

$$E_c = \sum_{i=1}^{N} E_{n_i \ell_i}.$$

The wave function for the entire system is totally antisymmetric, as spin statistics demands, and is a linear combination of single electron spin orbitals given by the Slater determinant:

$$\Psi(1, 2, \ldots, N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} u_\alpha(1) & u_\beta(1) & \ldots & u_\nu(1) \\ u_\alpha(2) & u_\beta(2) & \ldots & u_\nu(2) \\ \vdots & \vdots & \ddots & \vdots \\ u_\alpha(N) & u_\beta(N) & \ldots & u_\nu(N) \end{vmatrix}$$

Here α, β and ν represent different sets of the four quantum numbers (n, ℓ, m_ℓ, m_s) so that the final state satisfies the Pauli exclusion principle.

The order of the individual energy levels doesn't depend on the potential. The order is

$$1s2s2p3s3p[4s3d]4p[5s4d]5p[6s4f5d].$$

Orbitals in square brackets are close enough in energy that the exact order can change from atom to atom. Electrons with the same value of n are said to be in the same shell, electrons with the same values of n and ℓ are said to be in the same sub shell. The maximum number of electrons in the same subshell is $2(2\ell+1)$. The electron configuration of a given atom is given using a superscript to denote the number of electrons in a given subshell. For example in its ground state carbon has the electron configuration $1s^2 2s^2 2p^2$ which can also be written as $[\text{He}]2s^2 2p^2$ where $[\text{He}]$ denotes the electron configuration of helium in its ground state which is $1s^2$.

27 Variational Methods

27.1 The Theory

Variational methods, also known as Rayleigh–Ritz methods, are a way of approximating the ground state energy of a given Hamiltonian, $\hat{H}$. Suppose $\hat{H}$ has a complete set of eigenstates, $\{|n\rangle\}$, and corresponding energy eigenvalues, $\{E_n\}$, such that (E_n) is an increasing sequence, that is

$$E_1 \leq E_2 \leq \cdots \leq E_i \leq E_{i+1} \leq \cdots$$

Then we can expand any state, $|\psi\rangle$, as a linear superposition of these states as

$$|\psi\rangle = \sum_n c_n |\psi\rangle, \qquad \text{where} \qquad c_n = \langle n|\psi\rangle.$$

The expectation energy of a system in state $|\psi\rangle$ is then

$$\langle E \rangle = \frac{\langle \psi|\hat{H}|\psi\rangle}{\langle \psi|\psi\rangle} = \frac{\sum_n |c_n|^2 E_n}{\sum_n |c_n|^2} \geq \frac{\sum_n |c_n|^2 E_1}{\sum_n |c_n|^2} = E_1.$$

Where we have used the ordering of (E_n) to replace all E_n with $E_1 \leq E_n$. We see that $\langle E \rangle$ provides an upper bound for the ground state energy, E_1. Using this fact the variational method simply gives us a way to find a close upper bound.

To apply the variational method to a given Hamiltonian, $\hat{H}$, choose a trial state, $|\psi_T\rangle$, which depends on the parameters $\alpha_1, \ldots, \alpha_r$. We then calculate

$$E(\alpha_1, \ldots, \alpha_r) = \frac{\langle \psi_T|\hat{H}|\psi_T\rangle}{\langle \psi_T|\psi_T\rangle}.$$

To find the upper bound on the ground state energy we then minimise E with respect to its parameters by finding $\{\alpha_i^*\}$ such that

$$\left.\frac{\partial E}{\partial \alpha_i}\right|_{\alpha_i^*} = 0.$$

This then gives us a best estimate for the upper bound of the ground state energy.

This method works best when $|\psi_T\rangle$ is as close to the true ground state as possible. This means we should consider symmetry requirements when choosing $|\psi_T\rangle$. For example if our system is a particle in a box we should choose $\psi_T(\boldsymbol{r})$ to be zero at the boundaries.

It is important to note that the only thing variational methods predict is an upper bound of the ground state energy. The resulting trial state, $|\psi_T\rangle$, may have a close value for the expectation value of the energy but there is no reason why it should be close to the ground state. Similarly if we choose our parameters, $\{\alpha_i\}$, to have some physical meaning we shouldn't read too much into whatever values they have.

27.2   Hydrogen Ground State

Suppose we want to use variational methods to estimate the ground state energy of hydrogen. After various symmetry considerations we conclude that the ground state will have $\ell = 0$ which means the radial part of the wave function, Y_ℓ^m, will be spherically symmetric. Thus we choose our trial solution to be the radial component of the wave function. Further $\ell = 0$ is the only value where the radial component doesn't need to vanish at the origin. The only requirement is that our function be square integrable on $[0, \infty)$, i.e. $\psi_T \in \mathrm{L}^2([0, \infty))$. One such function is

$$\psi_T(r) = Ce^{-\alpha r}$$

where C is a normalisation constant and α is a variational parameter (not the fine structure coefficient). Notice that this is actually the form of the true wave function for the ground state.

First we will normalise $|\psi_T\rangle$ to avoid constant factors of $1/\langle\psi_T|\psi_T\rangle$.

$$\begin{aligned}
\langle\psi_T|\psi_T\rangle &= \int_V \psi_T^*(r)\psi_T(r)\,\mathrm{d}V \\
&= |C|^2 \int_0^{2\pi} \mathrm{d}\varphi \int_0^\pi \mathrm{d}\vartheta \int_0^\infty \mathrm{d}r\, e^{-2\alpha r} r^2 \sin\vartheta \\
&= |C|^2 \int_0^{2\pi} \mathrm{d}\varphi \int_0^\pi \sin\vartheta\,\mathrm{d}\vartheta \int_0^\infty r^2 e^{-2\alpha r}\,\mathrm{d}r \\
&= |C|^2 \cdot 2\pi \cdot 2 \cdot \left(\left[-\frac{1}{2\alpha} r^2 e^{-2\alpha r} \right]_0^\infty + \frac{1}{\alpha} \int_0^\infty r e^{-2\alpha r}\,\mathrm{d}r \right) \\
&= 4\pi|C|^2 \frac{1}{\alpha} \int_0^\infty r e^{-2\alpha r}\,\mathrm{d}r \\
&= 4\pi|C|^2 \frac{1}{\alpha} \left(\left[-\frac{1}{2\alpha} r e^{-2\alpha r} \right]_0^\infty + \frac{1}{2\alpha} \int_0^\infty e^{-2\alpha r}\,\mathrm{d}r \right) \\
&= 4\pi|C|^2 \frac{1}{2\alpha^2} \int_0^\infty e^{-2\alpha r}\,\mathrm{d}r \\
&= 4\pi|C|^2 \frac{1}{2\alpha^2} \left[-\frac{1}{2\alpha} e^{-\alpha r} \right]_0^\infty \\
&= \pi|C|^2 \frac{1}{\alpha^3}
\end{aligned}$$

For this to equal 1 we must have $C = \sqrt{\alpha^3/\pi}$.

The expectation value is linear so we can find $\langle\hat{H}\rangle$ using

$$\langle\hat{H}\rangle = \langle\hat{T} + \hat{V}\rangle = \langle\hat{T}\rangle + \langle\hat{V}\rangle.$$

We will start with $\langle\hat{V}\rangle$.

$$\langle\hat{V}\rangle = \langle\psi_T|\hat{V}|\psi_T\rangle$$

$$= -\frac{|C|^2 e^2}{4\pi\varepsilon_0} \int_V \frac{1}{r} e^{-2\alpha r} \, \mathrm{d}V$$

$$= -\frac{|C|^2 e^2}{4\pi\varepsilon_0} \int_0^{2\pi} \mathrm{d}\varphi \int_0^\pi \mathrm{d}\vartheta \int_0^\infty \mathrm{d}r \, \frac{1}{r} e^{-2\alpha r} r^2 \sin\vartheta$$

$$= -\frac{|C|^2 e^2}{\varepsilon_0} \int_0^\infty r e^{-2\alpha r} \, \mathrm{d}r$$

$$= -\frac{|C|^2 e^2}{\varepsilon_0} \left(\left[-\frac{1}{2\alpha} r e^{-2\alpha r} \right]_0^\infty + \frac{1}{2\alpha} \int_0^\infty e^{-2\alpha r} \, \mathrm{d}r \right)$$

$$= -\frac{|C|^2 e^2}{\varepsilon_0} \frac{1}{2\alpha} \int_0^\infty e^{-2\alpha r} \, \mathrm{d}r$$

$$= -\frac{|C|^2 e^2}{\varepsilon_0} \frac{1}{2\alpha} \left[-\frac{1}{2\alpha} e^{-2\alpha r} \right]_0^\infty$$

$$= -\frac{1}{4\alpha^2} \frac{|C|^2 e^2}{\varepsilon_0}$$

$$= -\frac{\alpha e^2}{4\pi\varepsilon_0}$$

Next we find the expectation value of the energy of the kinetic energy. To do this we will use the following identity:

$$\nabla \cdot (f\boldsymbol{v}) = f(\nabla \cdot \boldsymbol{v}) + (\nabla f) \cdot \boldsymbol{v}.$$

Considering the particular case $f = \varphi^*$ and $\boldsymbol{v} = \nabla\varphi$ this becomes

$$\nabla \cdot (\varphi^* \nabla\varphi) = \varphi^*(\nabla \cdot (\nabla\varphi)) + (\nabla\varphi^*) \cdot (\nabla\varphi).$$

Integrating over all space this becomes

$$\int_V \nabla \cdot (\varphi^* \nabla\varphi) \, \mathrm{d}V = \int_V \varphi^* \nabla^2 \varphi \, \mathrm{d}V + \int_V (\nabla\varphi^*) \cdot (\nabla\varphi) \, \mathrm{d}V \,.$$

Using the divergence theorem the left hand side becomes

$$\int_V \nabla \cdot (\varphi^* \nabla\varphi) \, \mathrm{d}V = \int_S (\varphi^* \nabla\varphi) \cdot \mathrm{d}\boldsymbol{S} \,.$$

If we take V to be all of space then the boundary, S, is at infinity. Since we are assuming $\varphi \in \mathrm{L}^2(\mathbb{R}^3)$ it will be zero when evaluated at infinity so the left hand side of the identity disappears and we are left with

$$\int_V \varphi^* \nabla^2 \varphi \, \mathrm{d}V = -\int_V |\nabla\varphi|^2 \, \mathrm{d}V \,.$$

The expectation value of the the kinetic energy is then

$$\langle \hat{T} \rangle = \frac{1}{2m} \langle \psi_T | \hat{P}^2 | \psi_T \rangle$$

$$= -\frac{\hbar^2}{2m} \int_V \psi_T^* \nabla^2 \psi_T \, \mathrm{d}^3 r$$

$$= \frac{\hbar^2}{2m} \int_V |\nabla\psi_T|^2 \, \mathrm{d}^3 r$$

The trial function, ψ_T, depends only on r so

$$\nabla\psi_T = \frac{\partial \psi_T}{\partial r} \boldsymbol{e_r}$$

and hence

$$|\nabla\psi_T|^2 = \left| \frac{\partial \psi_T}{\partial r} \right| = \alpha^2 |\psi_T|^2.$$

So

$$\langle \hat{T} \rangle = \frac{\hbar^2 \alpha^2}{2m}.$$

So the expectation value of the Hamiltonian is

$$E(\alpha) = \langle \hat{H} \rangle = \langle \hat{T} \rangle + \langle \hat{V} \rangle = \frac{\hbar^2 \alpha^2}{2m} - \frac{e^2 \alpha}{4\pi\varepsilon_0}.$$

Differentiating with respect to α we have

$$\frac{\partial E}{\partial \alpha} = \frac{\hbar^2 \alpha}{m} - \frac{e^2}{4\pi\varepsilon_0}.$$

Demanding that this is zero so E is minimised we have

$$\alpha = \frac{me^2}{r\pi\varepsilon_0 \hbar^2} = \frac{1}{a_0}$$

where a_0 is the Bohr radius. Thus an upper bound on the ground state energy is

$$E = -\frac{e^2}{2(4\pi\varepsilon_0)a_0} = -1\,\mathrm{Ry} = -13.6\,\mathrm{eV}.$$

This is actually the ground state energy because our trial function with $\alpha = 1/a_0$ is the ground state eigenfunction so $\langle \hat{H} \rangle$ is simply the expectation value of the energy of a particle in a ground state, which is the ground state energy.

27.3   Helium Ground State

Suppose we want to use variational methods to calculate the ground state energy of helium. As a trial wave function we take a product of 1s orbital wave functions with an effective nuclear charge $\tilde{Z}$ which accounts for shielding of the nucleus by the other electron.

$$\psi_T(r_1, r_2) = \frac{\tilde{Z}^3}{\pi a_0^3} \exp\left[-\frac{\tilde{Z}}{a_0}(r_1 + r_2)\right].$$

Setting $\tilde{Z} = Z$ would be the equivalent of no electron-electron interaction. The variational parameter in this wave function is $\tilde{Z}$.

For the kinetic energy we can use the same result as for hydrogen:

$$\langle \hat{T} \rangle = -\frac{\hbar^2}{2m}|\psi_T\rangle(\hat{P}_1^2 + \hat{P}_2^2)|\psi_T\rangle = \frac{\hbar^2 \tilde{Z}^2}{ma_0^2} = 2\tilde{Z}^2\,\mathrm{Ry}.$$

Similarly for the potential energy

$$\langle \hat{V} \rangle = -\frac{2Ze^2}{4\pi\varepsilon_0 a_0}\tilde{Z} = -4Z\tilde{Z}\,\mathrm{Ry}.$$

Finally the electron-electron Coulomb term has an expectation value of $5/4$Ry. The total energy is then

$$E(\tilde{Z}) = 2\left[\tilde{Z}^2 - \left(2Z - \frac{5}{8}\right)\tilde{Z}\right]\,\mathrm{Ry}.$$

Differentiating with respect to $\tilde{Z}$, setting equal to zero, and then rearranging gives us

$$\tilde{Z} = Z - \frac{5}{16}.$$

The energy corresponding to this is then

$$E(\tilde{Z}) = -5.695\,\mathrm{Ry} = -77.46\,\mathrm{eV}.$$

This is within $2\,\%$ of the measured value of $-78.975\,\mathrm{eV}$ and is better than the first order perturbation result of $-74.8\,\mathrm{eV}$.

27.4   Excited States

It is sometimes possible to use variational methods to find upper bounds on excited states. Suppose we know enough about the ground state to choose a trial state that is orthogonal to the ground state such that $\langle n = 1 | \psi_T \rangle = 0$. Then the expansion of $|\psi_T\rangle$ in $\{|n\rangle\}$ doesn't contain $|n = 1\rangle$ and so

$$\frac{\langle \psi_T | \hat{H} | \psi_T \rangle}{\langle \psi_T | \psi_T \rangle} = \frac{\sum_{n=2} |c_n|^2 E_n}{\sum_{n=2} |c_n|^2} \geq E_2.$$

Using variational methods as before we can then find an upper bound on E_2.

The problem with this is we don't typically know the eigenstates so we can't construct an orthogonal state. We can avoid this problem if we are interested in an excited state with symmetries lower energy states don't posses. If we choose a trial state that does have these symmetries then it is guaranteed to be orthogonal to the lower energy states and we can use variational methods.

Consider, for example, helium. It's ground state is a ^{1}S state. The first four excited states are ^{3}S, ^{1}S, ^{3}P, and ^{1}P states/ Apart from the ^{1}S state these all have symmetries different from the ground state. For example P states have $L = 1$ and are therefore orthogonal to the ground state which has $L = 0$. A trial function that may be considered when trying to find an upper bound on the state 1s2p is

$$\psi_T^{\pm}(\boldsymbol{r_1}, \boldsymbol{r_2}) = C_{\pm}[f(r_1)g(r_2)\cos\vartheta_1 \pm f(r_2)g(r_2)\cos\vartheta_2]$$

where

$$f(r) = e^{-\alpha r/a_0}, \qquad \text{and} \qquad g(r) = re^{-\beta r/2a_0}.$$

This has two variational parameters α and β.

Part VII

Applications of Quantum Theory

# 28   Hidden Variables, EPR Paradox and Bell's Inequality

28.1   Hidden Variables

For most of the time up until this point we haven't thought too hard about what anything that we are doing actually means. We started by stating that the state of a system is specified by its wave function. We then went on to say that observables don't have definite values until we measure them. The most we can predict in general is the probability of a given result. This assumes (correctly) that there is no pre-existing value of the observable. This upset a lot of important physicists in early quantum theory, including Einstein who famously said 'God does not play dice'.

Einstein believed that there was an underlying reality in which observables have definite values and we are simply ignorant of this deeper reality. This idea, called a hidden variables theory, is similar to the assumption that, given infinite computing power, we could deterministically calculate the future of any system and wouldn't have to resort to statistical mechanics. While it turns out that hidden variables theories are wrong in showing so we develop several key ideas.

28.2   EPR Paradox

28.2.1   Background

Recall that for a spin 1/2 system when measuring spin along the z-axis we can get either $\pm 1/2$. This idea is the basis for the Stern–Gerlach experiment. The details of how aren't important here but a Stern–Gerlach experiment splits a beam of particles into two beams according to the spin along a given axis. For example, figure 28.1 shows a Stern–Gerlach magnet splitting a beam according to its spin along the z-axis. The two beams leaving this machine correspond to the two eigenstates of $\hat{S}_z$. Recall that $\hat{S}_z$
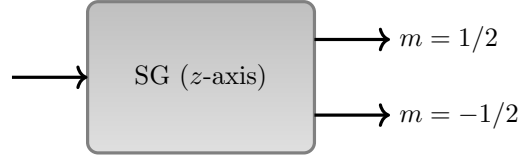
Figure 28.1: Stern–Gerlach experiment splits a beam according to the spin along the z-axis.

can be represented as
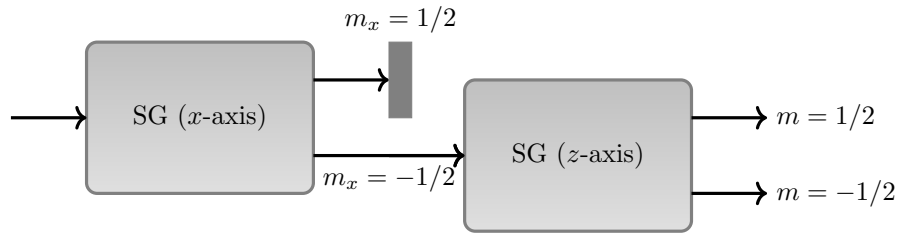
$$\hat{S}_z \to \frac{1}{2}\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

This has two eigenvectors which have the following representations

$$|m = 1/2\rangle \to \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \text{and} \qquad |m = -1/2\rangle \to \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

It is equally valid to have a Stern–Gerlach experiment measure along the x-axis. The corresponding operator, $\hat{S}_x$, has the representation

$$\hat{S}_x \to \frac{1}{2}\hbar \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

This has two eigenvectors which have the following representations

$$|m_x = 1/2\rangle \to \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \text{and} \qquad |m_x = -1/2\rangle \to \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

These can be written as a linear combination of eigenstates of $\hat{S}_z$:

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Using a Stern–Gerlach magnet oriented in the x direction we can select the beam containing only particles in the $|m_x = -1/2\rangle$ state. If we then use a second Stern–Gerlach magnet to measure the z component of the spin then we find that the probability of measuring $m = 1/2$ is $\left|1/\sqrt{2}\right|^2 = 1/2$. This is shown in figure 28.2 We can more generally measure the spin along some unit vector in the (x, z)-plane, $\hat{n} =$

Figure 28.2: Stern–Gerlach experiment splits a beam according to the spin along the z-axis.

$(\sin\vartheta, 0, \cos\vartheta)$. The corresponding operator for this measurement is

$$\hat{\boldsymbol{S}} \cdot \hat{\boldsymbol{n}} = \hat{S}_x \sin\vartheta + \hat{S}_z \cos\vartheta \to \frac{1}{2} \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ \sin\vartheta & -\cos\vartheta \end{pmatrix}.$$

This has eigenvectors

$$|m_\vartheta = 1/2\rangle \to \begin{pmatrix} \cos\frac{\vartheta}{2} \\ \sin\frac{\vartheta}{2} \end{pmatrix}, \qquad \text{and} \qquad |m_\vartheta = -1/2\rangle \to \begin{pmatrix} \sin\frac{\vartheta}{2} \\ \cos\frac{\vartheta}{2} \end{pmatrix}.$$

28.2.2   The EPR Thought Experiment

In 1935 Einstein, Podolsky, and Rosen published a paper discussing a thought experiment that they thought demonstrated something was missing from quantum mechanics. We will discuss this thought experiment, known as the EPR paradox here.

Consider a process where a particle with no spin decays into a pair of spin $1/2$ particles, for example a photon could produce an electron-positron pair. To conserve momentum these particles will go off in opposite directions from the point where they were created. Suppose that we analyse the two particles with Stern–Gerlach magnets, A and B, which can be orientated in an arbitrary direction. We will place A and B sufficiently far apart such that within the time taken for the experiment to occur even light cannot get between from A to B. This means that no information can be shared during the experiment. This is shown in figure 28.3. Suppose that both experiments are aligned to measure the spin along the
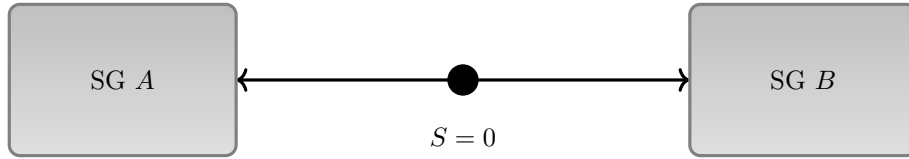
Figure 28.3: The setup for the EPR paradox.

z-axis. Conservation of angular momentum means that since we started with $S = 0$ we must have $S = 0$ at all times. The only way for this to happen is if the two particles have equal and opposite spins along the z-axis. Therefore if we measure the spin at A and get $\hbar/2$ we know for a fact that at B they must measure $-\hbar/2$. At first this seems to be the same as classical mechanics which also has conservation of angular momentum. However in quantum mechanics the spin isn't defined until we take a measurement. This means that when we make the measurement at A before the measurement the spin is undefined at both A and B. After the measurement the spin is defined at both A and B. Therefore by making a measurement at A we necessarily change something at B. This raises two questions:

1. Is the fact that the outcome at B is entirely predictable a consequence of the measurement at A?

2. If so how does the information about the measurement at A get to B?

Suppose instead we had A measuring along the x-axis and B measuring along the z-axis. Then the measurement at A doesn't affect the outcome of the experiment at B as both of $|m_x = \pm 1/2\rangle$ are a superposition of equal parts of $|m = \pm 1/2\rangle$. Suppose instead that B measures at an angle ϑ to the z-axis in the (x, z)-plane. If A measured $-\hbar/2$ as the x spin then we know that the particle at B must have an x component of $\hbar/2$. We can then decompose the $\hbar/2$ state along the z-axis in terms of the spin along the ϑ-axis:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \cos\frac{\vartheta}{2} \begin{pmatrix} \cos\frac{\vartheta}{2} \\ \sin\frac{\vartheta}{2} \end{pmatrix} + \sin\frac{\vartheta}{2} \begin{pmatrix} \sin\frac{\vartheta}{2} \\ -\cos\frac{\vartheta}{2} \end{pmatrix}$$
$$= \cos\frac{\vartheta}{2}|m_\vartheta = 1/2\rangle + \sin\frac{\vartheta}{2}|m_\vartheta = -1/2\rangle.$$

So measuring along the ϑ-axis B measures spin-up with probability $\cos^2(\vartheta/2)$ and spin down with probability $\sin^2(\vartheta/2)$. The measurement at B is still affected by the measurement at A.

The fact that measurement at A seems to affect measurements at B regardless of how far apart the two experiments are seems to imply action at a distance or faster than light information transmission. The solution that EPR suggested was that the particles knew before hand (or rather some hidden variables fixed before hand) the spin they would have when measured along any axis. However this was shown to be false by John Bell.

28.3   Bell's Inequality

Bell sought to the following question:

1. Is quantum mechanics a realistic local theory?

2. If not can we make a prediction or observation that can demonstrate this.

To answer these he came up with the Bell inequality.

We start with the same setup as with the EPR paradox. We define three co-planar directions, $\{\hat{\boldsymbol{n}}_i\}$, along which we can measure the spin.

28.3.1   Realistic Theory

In a realistic local theory the spin $1/2$ particles must individually have definite values, $\pm\hbar/2$, for each direction we can measure. Notice that we don't require measurement of all three directions simultaneously, this can't be done as generally spin operators along different axes don't commute. We will denote the two possible outcomes along a given axis by $+$ and $-$. The spin state of a particle can then be denoted $(\pm, \pm, \pm)$ where, for example, $(+, -, +)$ denotes a particle that will have spin $\hbar/2$ if measured along $\hat{\boldsymbol{n}}_1$ or $\hat{\boldsymbol{n}}_3$ and $-\hbar/2$ if measured along $\hat{\boldsymbol{n}}_2$. We can denote the spin state of the whole system as $(\pm, \pm, \pm, \pm, \pm, \pm)$ which is simply the state of one particle written immediately after the state of the other particle. A generic state can then be written as $(\sigma_1, \sigma_2, \sigma_3, \tau_1, \tau_2, \tau_3)$ where σ_i and τ_i represent the spin of the particle at A and B respectively when measured along $\hat{\boldsymbol{n}}_i$. We denote by $f(\sigma_1, \sigma_2, \sigma_3, \tau_1, \tau_2, \tau_3)$ the fraction of particle pairs that are in the state $(\sigma_1, \sigma_2, \sigma_3, \tau_1, \tau_2, \tau_3)$. Since ew assume that the net spin is 0 we have $f(\sigma_1, \sigma_2, \sigma_3, \tau_1, \tau_2, \tau_3) = 0$ unless $\sigma_i = -\tau_i$ for all i. This satisfies locality as we can set σ_i and τ_i when the particles are first produced and close and then nothing changes when we make a measurement.

28.3.2   Spin Correlations

Denote by $P_{++}(\hat{\boldsymbol{n}}_i, \hat{\boldsymbol{n}}_j)$ the probability that when measuring along $\hat{\boldsymbol{n}}_i$ $(\hat{\boldsymbol{n}}_j)$ at A (B) we get $+$ $(+)$. Then, for example

$$P_{++}(\hat{\boldsymbol{n}}_1, \hat{\boldsymbol{n}}_2) = \sum_{\sigma_2, \sigma_3} \sum_{\tau_1, \tau_2} f(+, \sigma_2, \sigma_3, \tau_1, +, \tau_3)$$
$$= f(+, -, +, -, +, -) + f(+, -, -, -, +, +)$$

where we have used the fact that f is zero unless $\sigma_i = -\tau_i$, Similarly we have

$$P_{++}(\hat{\boldsymbol{n}}_3, \hat{\boldsymbol{n}}_2) = f(+, -, +, -, +, -) + f(-, -, +, +, +, -),$$
$$P_{++}(\hat{\boldsymbol{n}}_1, \hat{\boldsymbol{n}}_3) = f(+, +, -, -, -, +) + f(+, -, -, -, +, +).$$

Since f are all non-negative we can add these last two results and we get

$$P_{++}(\hat{\boldsymbol{n}}_1, \hat{\boldsymbol{n}}_2) \leq P_{++}(\hat{\boldsymbol{n}}_3, \hat{\boldsymbol{n}}_2) + P_{++}(\hat{\boldsymbol{n}}_1, \hat{\boldsymbol{n}}_3).$$

28.3.3   Non-Realistic Theory

We can now compute the spin correlations that we get from quantum mechanics. Suppose $\hat{\boldsymbol{n}}_1$ corresponds to the z-axis and at B we choose to measure the spin along the $\hat{\boldsymbol{n}}_2$ direction which is at an angle ϑ to the z-axis. If ϑ_{ij} is the angle between $\hat{\boldsymbol{n}}_i$ and $\hat{\boldsymbol{n}}_j$ then

$$P_{++}(\hat{\boldsymbol{n}}_i \hat{\boldsymbol{n}}_j) = \frac{1}{2}\sin^2\frac{\vartheta_{ij}}{2}, \qquad \text{and} \qquad P_{-+}(\hat{\boldsymbol{n}}_i, \hat{\boldsymbol{n}}_j) = \frac{1}{2}\cos^2\frac{\vartheta_{ij}}{2}.$$

If Bell's inequality is to be satisfied then we must have

$$\sin^2\frac{\vartheta_{12}}{2} \leq \sin^2\frac{\vartheta_{23}}{2} + \sin^2\frac{\vartheta_{13}}{2}.$$

Consider the case that

$$\vartheta_{13} = \vartheta_{23} = \frac{\vartheta_{12}}{2}.$$

Then Bell's inequality becomes

$$\sin^2\vartheta_{13} \leq 2\sin^2\frac{\vartheta_{13}}{2}.$$

This simplifies to

$$\cos^2\frac{\vartheta_{13}}{2} \leq \frac{1}{2}.$$

This is not satisfied if $0 \leq \vartheta_{13} \leq \pi/2$. So if quantum mechanics is correct we would expect Bell's inequality to be violated showing it is not a realistic local theory.

28.3.4 Experimental Tests of Bell's Inequality

Now that we know Bell's inequality can differentiate between a realistic local theory and a non-realistic local theory we need to test to see if Bell's theory is ever violated. This was first done not with spins but with linearly polarised photons. Photons where polarised at A and B in directions $\hat{\boldsymbol{n}}_i$ and $\hat{\boldsymbol{n}}_j$ respectively. We define a result to be $+$ when the polarisation is parallel to $\hat{\boldsymbol{n}}_i$ and $-$ if it is perpendicular. The correlation coefficient is defined as

$$E(\hat{\boldsymbol{n}}_i, \hat{\boldsymbol{n}}_j) = P_{++}(\hat{\boldsymbol{n}}_i, \hat{\boldsymbol{n}}_j) + P_{--}(\hat{\boldsymbol{n}}_i, \hat{\boldsymbol{n}}_j) - P_{-+}(\hat{\boldsymbol{n}}_i, \hat{\boldsymbol{n}}_j) - P_{+-}(\hat{\boldsymbol{n}}_i, \hat{\boldsymbol{n}}_j).$$

This is the expectation value in the singlet state:

$$\langle (\boldsymbol{\sigma} \cdot \hat{\boldsymbol{n}}_i)(\boldsymbol{\sigma} \cdot \hat{\boldsymbol{n}}_j) \rangle = \cos \vartheta_{ij}.$$

In the original experiment the combination

$$S = E(\hat{\boldsymbol{n}}_i, \hat{\boldsymbol{n}}_j) - E(\hat{\boldsymbol{n}}_i, \hat{\boldsymbol{n}}_j') + E(\hat{\boldsymbol{n}}_i', \hat{\boldsymbol{n}}_j) + E(\hat{\boldsymbol{n}}_i', \hat{\boldsymbol{n}}_j').$$

Where $\hat{\boldsymbol{n}}_i$, $\hat{\boldsymbol{n}}_i'$ and $\hat{\boldsymbol{n}}_j$, $\hat{\boldsymbol{n}}_j'$ represent two different orientations of the polarises at A and B respectively.

In a realistic local theory a Bell inequality for this quantity is

$$-2 \leq S \leq 2.$$

Quantum mechanics on the other hand predicts

$$-2\sqrt{2} \leq S \leq 2\sqrt{2}.$$

Orientations where chosen such that

$$\hat{\boldsymbol{n}}_i \cdot \hat{\boldsymbol{n}}_j = \hat{\boldsymbol{n}}_i' \cdot \hat{\boldsymbol{n}}_j = \hat{\boldsymbol{n}}_i' \cdot \hat{\boldsymbol{n}}_j' = \cos \varphi, \qquad \text{and} \qquad \hat{\boldsymbol{n}}_i \cdot \hat{\boldsymbol{n}}_j' = \cos 3\varphi.$$

The value of S was then determined for various values of φ between $0°$ and $90°$. The results violated Bell's inequality. The maximum violation was at $\varphi = 22.5°$ where they measured $S = 2.697 \pm 0.015$.

29 Quantum Teleportation

In the last section we introduced the idea of entangled particles. These can be used in many different ways to do things that can't be done classically. The first use we will consider is quantum teleportation. This allows for information about one quantum state to be transferred to another location. No matter is transferred, only information. This process is called teleportation because if a second particle is placed into the same state that is sent to the remote location then it as if the particle teleported. The process requires both quantum and classical communication and the restriction of the classical communication to be slower than the speed of light means that information cannot be transferred faster than the speed of light.

29.1 Entanglement

In classical computing the most basic unit of information is the bit which can be either 0 or 1. In quantum computing the most basic unit of information is the **qubit** which is a two state system, which can be described by the Hilbert space $\mathcal{H}_{\text{qubit}}$. Often we identify one state with 0 and the other with 1. As a matter of notation in this section we will refer to two states, $|+\rangle$ and $|-\rangle$ although later on we will refer to the same two states as $|0\rangle$ and $|1\rangle$ respectively in a different context. We will often think of the states as electron spins, for example we may identify spin up with $|+\rangle$ and spin down with $|-\rangle$. We choose electron spins simply because we are familiar with the mathematics behind them but in practice any two state system could be used. In this case it is natural to make the following identification with spinor states:

$$|+\rangle \to \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \text{and} \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The main difference from a classical bit is that a qubit is in a quantum superposition of the two pure states. A generic qubit state, $|\psi\rangle \in \mathcal{H}_{\text{qubit}}$, is then

$$|\psi\rangle = a_-|-\rangle + a_+|+\rangle$$

for some $a_-, a_+ \in \mathbb{C}$ such that $|a_-|^2 + |a_+|^2 = 1$. We can parametrise this as

$$|\psi\rangle = \cos\frac{\vartheta}{2}|+\rangle + e^{i\varphi}\sin\frac{\vartheta}{2}|-\rangle$$

where $\vartheta \in [0, \pi]$ and $\varphi \in [0, 2\pi]$. We are free to choose the phase of $|\psi\rangle$ to make a_- real. When we do this we often represent the state as a point on the Bloch sphere as shown in figure 29.1. The next step
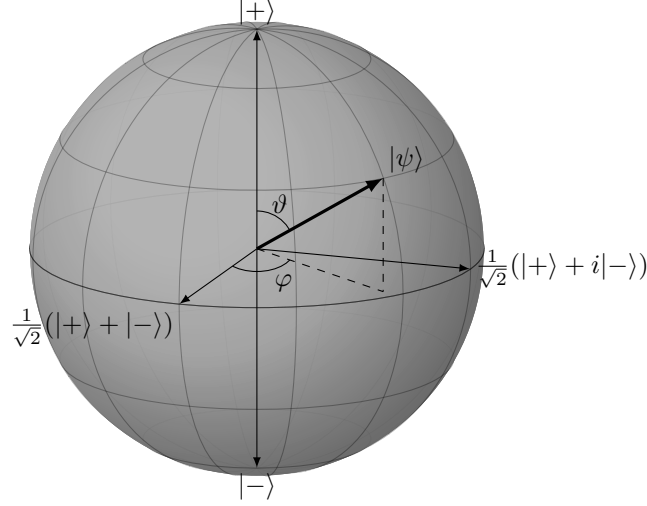

Figure 29.1: A Bloch sphere parametrised in terms of ϑ and φ.

up from a single bit is a pair of bits. Two bits can represent $2^2 = 4$ different states, 00, 01, 10, and 11. The quantum analogue of a two bit system is a two qubit system which is two two state states described by $\mathcal{H} = \mathcal{H}_{\text{qubit}} \otimes \mathcal{H}_{\text{qubit}}$. A generic state, $|\psi\rangle \in \mathcal{H}$, is then given by

$$|\psi\rangle = a_{--}|--\rangle + a_{-+}|-+\rangle + a_{+-}|+-\rangle + a_{++}|++\rangle$$

where $a_{ij \in \mathbb{C}}$ such that $\sum_{ij}|a_{ij}|^2 = 1$. Each state is simply $|ij\rangle = |i\rangle \otimes |j\rangle$.

Now that we have two systems together we have the possibility for entanglement. Call the two systems A and B. An **entangled state** between systems A and B is any state that can be written as $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$ where $|\psi_A\rangle$ and $|\psi_B\rangle$ are the states of A and B respectively. Note that not every state in $\mathcal{H}$ can be written this way even though $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. When discussing entangled states it is often best to work in a basis of entangled states called the **Bell states**:

$$|\Phi^\pm\rangle = \frac{\sqrt{2}}{2}(|++\rangle \pm |--\rangle), \qquad \text{and} \qquad |\Psi^\pm\rangle = \frac{\sqrt{2}}{2}(|+-\rangle \pm |-+\rangle).$$

29.2   No Cloning Theorem

> **Theorem 9: No Cloning Theorem**
>
> It is impossible to create an identical copy of an arbitrary *unknown* quantum state.

Let $|\chi\rangle \in \mathcal{H}$ be an unknown quantum state that we wish to duplicate. The duplicate exists with the original so the entire system is described by states in $\mathcal{H} \otimes \mathcal{H}$. Suppose we start with some arbitrary normalised state $|j\rangle$. The system then starts in the state

$$|j\rangle \otimes |\chi\rangle = |j\rangle|\chi\rangle = |j\chi\rangle \in \mathcal{H} \otimes \mathcal{H}.$$

The question is can we somehow act on this state to reach the state $|\chi\rangle \otimes |\chi\rangle$? There are two possible things that we can do to change states. First we can observe them but this collapses the state and we lose information and, in general, alter $|\chi\rangle$, so this option isn't useful here. The second thing we can do is control the Hamiltonian of the system so that the system evolves with some unitary time evolution operator, $\hat{U}(t)$, which evolves the system into the desired state. This allows for an alternate framing of the no cloning theorem:

> **Theorem 9: No Cloning Theorem (Alternate Formulation)**
>
> There is no unitary operator, $\hat{U}$, on $\mathcal{H} \otimes \mathcal{H}$ with the action
>
> $$\hat{U}(|j\rangle \otimes |\chi\rangle) = e^{i\varphi}|\chi\rangle \otimes |\chi\rangle.$$
>
> Where $|\chi\rangle \in \mathcal{H}$ is an *unknown* state and $e^{i\varphi}$ is some arbitrary phase.

Proof. Suppose that there exists a unitary operator, $\hat{U}$, on $\mathcal{H} \otimes \mathcal{H}$ such that

$$\hat{U}(|j\rangle \otimes |\chi\rangle) = \exp[-i\Theta(j, \chi)](|\chi\rangle \otimes |\chi\rangle),$$

and

$$\hat{U}(|j\rangle \otimes |\xi\rangle) = \exp[-i\Theta(j, \xi)](|\xi\rangle \otimes |\xi\rangle),$$

where $|\chi\rangle, |\xi\rangle \in \mathcal{H}$ are arbitrary states, and $\Theta(j, \chi)$ and $\Theta(j, \xi)$ are phases. Using $\hat{U}^\dagger \hat{U} = \mathbb{1}$ we have

$$\langle j|\hat{U}^\dagger \hat{U}|j\rangle = \langle j|j\rangle = 1.$$

Thus

$$\begin{aligned}
\langle \chi|\xi\rangle &= \langle \chi| \otimes \langle j|\hat{U}^\dagger \hat{U}|j\rangle|\xi\rangle \\
&= \exp[i(\Theta(j, \chi) - \Theta(j, \xi))](\langle \chi| \otimes \langle \chi|)(|\xi\rangle \otimes |\xi\rangle) \\
&= \exp[i(\Theta(j, \chi) - \Theta(j, \xi))](\langle \chi|\xi\rangle)^2.
\end{aligned}$$

Consider now $|\langle \chi|\xi\rangle|$. Since the exponential term is unitary we have $|\langle \chi|\xi\rangle| = |\langle \chi|\xi\rangle|^2$ which means either $|\langle \chi|\xi\rangle| = 1$ or $|\langle \chi|\xi\rangle| = 0$. Therefore $|\chi\rangle$ and $|\xi\rangle$ are either parallel or orthogonal. This contradicts the assumption that $|\chi\rangle$ and $|\xi\rangle$ are arbitrary states so we conclude that no such operator can exist. $\qquad\square$

The no cloning theorem means that quantum teleportation is one of the only way to get information about an unknown state from A to B. We could simply move the particle from A to B of course but doing this without disturbing the state becomes increasingly difficult as distances and complexities increase. One could imagine taking all pertinent information about a state and sending it classically to B and recreating the state there but this violates the no cloning theorem. There is a corollary theorem, called the **no teleportation theorem**, which states that an arbitrary quantum system cannot be converted into a system of classical bits. If this were possible then we could violate the no cloning theorem. Note that the word arbitrary is important here. We will see in the next section that quantum teleportation doesn't work with arbitrary states.

The no cloning theorem can be seen, like weird things in quantum mechanics, as a corollary of the uncertainty principle. If we could clone an arbitrary state then we could make a large number of copies and perform multiple measurements and measure conjugate variables to arbitrary precision violating the uncertainty relationship between them.

29.3   Quantum Teleportation

Suppose we have two observers, Alice[13] or A, and Bob or B. Alice has a particle in quantum state $|\chi\rangle$. This may be an electron with a particular spin or a photon with a given polarisation, the details aren't important. Alice may not know the precise state of this system. Bob, at another location, has an identical particle. The goal is for Alice to send complete information about her state, $|\chi\rangle$, to Bob to input into their identical particle so that Bob ends up with a particle in an identical state. This is the process of **quantum teleportation**. Information is the only thing transferred in quantum teleportation. No matter is transferred.

To understand the process of quantum teleportation we will take the simplest case where $|\chi\rangle$ is a single qubit which we will imagine to be a spin $1/2$ electron. Quantum teleportation two entangled particles. Suppose that they are in the Bell state

$$|\Psi_{AB}^-\rangle = \frac{\sqrt{2}}{2}[|+_A -_B\rangle - |-_A +_B\rangle].$$

---

[13]`https://en.wikipedia.org/wiki/Alice_and_Bob`

Here subscripts denote which observer starts with the relevant particle. The details of how this pair of particles are created are not important. It could be that Alice or Bob produced them at the start of the experiment or some third party, Charlie, could have provided them. The important thing is that the particles are entangled and the entangled state is known. Sharing this entangled state between Alice and Bob is called establishing a **quantum channel** between them. There is also a classical channel such as a phone line, messaging service or bicycle messenger between them, again the details aren't important as long as Alice can somehow send at least two classical bits to Bob.

To recap Alice now has two particles. The original particle she wishes to send information about to Bob, which is in state $|\chi\rangle$, and her entangled particle in state $|\chi_A\rangle$. The most general state that $|\chi\rangle$ can be in is

$$|\chi\rangle = c|+\rangle + d|-\rangle$$

where $|c|^2 + |d|^2 = 1$. The state of the entire, three particle system is

$$\begin{aligned}
|\Xi\rangle &= |\chi\rangle \otimes |\Psi_{AB}^-\rangle \\
&= [c|+\rangle + d|-\rangle] \otimes \frac{\sqrt{2}}{2}[|+_A-_B\rangle - |-_A+_B\rangle] \\
&= [c|+\rangle + d|-\rangle] \otimes \frac{\sqrt{2}}{2}[|+_A\rangle \otimes |-_B\rangle - |-_A\rangle \otimes |+_B\rangle] \\
&= \frac{c\sqrt{2}}{2}[|+\rangle \otimes |+_A\rangle \otimes |-_B\rangle + |+\rangle \otimes |-_A\rangle \otimes |+_B\rangle] \\
&\quad + \frac{d\sqrt{2}}{2}[|-\rangle \otimes |+_A\rangle \otimes |-_B\rangle + |-\rangle \otimes |-_A\rangle \otimes |+_B\rangle].
\end{aligned}$$

We can rewrite the state of the two particles that Alice has in terms of Bell states. For example

$$|+\rangle \otimes |+_A\rangle = \frac{\sqrt{2}}{2}[|\Phi_A^+\rangle + |\Phi_A^-\rangle].$$

If we do this for all other pairs of states then we have

$$\begin{aligned}
|\Xi\rangle = \frac{1}{2} \big[&|\Psi_A^-\rangle \otimes (-c|+_B\rangle - d|-_B\rangle) + |\Psi_A^+\rangle \otimes (-c|+_B\rangle + d|-_B\rangle) \\
&+ |\Phi_A^-\rangle \otimes (c|-_B\rangle + d|+_B\rangle) + |\Phi_A^+\rangle \otimes (c|-_B\rangle - d|+_B\rangle) \big].
\end{aligned} \tag{29.1}$$

Both Alice and Bob have enough information to write this expansion. Notice that the states of B are now similar in form to the state $|\chi\rangle$, in fact the first state is identical to the state of $|\chi\rangle$ up to total phase of -1. The other states are related to $|\chi\rangle$ in a simple way. Simple here meaning that there is a unitary transformation that will take any given state to $|\chi\rangle$.

At this point Alice makes an observation of her particles and this collapses the wave function. Alice will observe one of the four Bell states, $|\Psi_A^\pm\rangle$ and $|\Phi_A^\pm\rangle$. Whichever she observes she will send this information to Bob via the classical channel. This requires two classical bits of information as there are four Bell states. Once Bob receives this information they will know exactly what state their particle is in because they too know the above expansion. Bob can then perform the appropriate unitary transformation and move their state to coincide with the state $|\chi\rangle$. So we have sufficiently teleported $|\chi\rangle$ from Alice to Bob.

Notice that it was necessary for Alice to make an observation which disturbs the state of the particle she started with and so the particle hasn't been cloned as we don't now have two copies. Another important feature is that it is not necessary for Alice to know anything about $|\chi\rangle$. If this was required then Alice would have to make measurements which would collapse $|\chi\rangle$ before it could be teleported. If Alice does know the state of her particle then there is really no reason to use quantum teleportation as she can simply send this information classically to Bob. However, we will see later that quantum teleportation can be used as the basis for secure communication of information so it may still be useful even if Alice knows the state of her particle.

30 Secure Communication and Superdense Coding

30.1 Secure Communication

One important application of quantum teleportation is securely sending information from one place to another without a third party being able to listen in. There are many procedures that have been

developed with this goal in mind. In this context a procedure is called a **protocol**. The simplest protocol uses quantum teleportation to send a message that can only be understood if you have one of the entangled particles.

We will consider the protocol called quantum secure direct communication. Suppose Alice wants to send a binary message to Bob. If she used a classical channel then she risks a third party, Eve, listening in. Instead Alice and Bob set up N pairs of entangled particles for a message of N bits. Alice then sets up N particles in states $|\chi_i\rangle$ being either $|+\rangle$ or $|-\rangle$ to match the ith bit of the message. Alice and Bob then follow the normal quantum teleportation procedure to send information about $|\chi_i\rangle$ to Bob. This protocol is secure because while Eve could listen in to the classical channel and find out what state Alice measures without access to the quantum channel, that is the entangled pair, that knowledge does not provide enough information to know what state $|\chi_i\rangle$ is. This protocol can actually be made more efficient as Bob only cares if $|\chi_i\rangle$ represents $+$ or $-$ and therefore Alice only needs to send one classical bit as the overall phase is not important.

For example, suppose Alice wants to send a $+$ bit of data to Bob. She starts by preparing her state,

$$|\chi\rangle = c|+\rangle + d|-\rangle,$$

with $c = 1$ and $d = 0$. The state of the overall system is given in equation 29.1. If Alice measures either of $|\Psi_A^\pm\rangle$ and tells Bob she measured one of the Ψ states then Bob's measurement of their state will give $+$ and they will know that $c = 1$. Similarly if Alice measures either of $|\Phi_A^\pm\rangle$ and tells Bob she measured one of the Φ states then Bob's measurement of their state will give $-$ and again they will know that $c = 1$.

Problems that arise when this protocol is used mostly stem from the fact that it is only as secure as the quantum channel is. For example if Alice is producing the entangled pairs and sending Bob their entangled particle and Eve can intercept the particle then she can listen in on the classical channel and use the entangled particle to decode the message in the same way Bob would. Eve can then send the entangled particle on to Bob to avoid suspicion. However there is no way that Eve can do this without making measurements of the particles and interfering with the entanglement. Therefore Alice and Bob should test the entangled particles they use to ensure that the correlation still exists to know that no one has interfered.

Another protocol, called remote state preparation, works as follows. Suppose Alice wishes to send Bob the state

$$|\chi\rangle = c|+\rangle + d|-\rangle$$

and she knows the values of c and d. As usual Alice and Bob must start of sharing a pair of entangled particles. Suppose they start in state

$$|\Psi_{AB}^-\rangle = \frac{\sqrt{2}}{2}(|+_A-_B\rangle - |-_A+_B\rangle).$$

This state can be expressed in a different basis made of $|\chi\rangle$ and a state $|\chi^\perp\rangle$, which is perpendicular to $|\chi\rangle$. For example

$$|\chi^\perp\rangle = d^*|+\rangle - c^*|-\rangle.$$

In this basis

$$|\Psi_{AB}^-\rangle = \frac{\sqrt{2}}{2}(-|\chi_A\chi_B^\perp\rangle + |\chi_A^\perp\chi_B\rangle).$$

Alice measures her state in this basis. If she measures $\chi_A^\perp$ then Bob will have a particle in the desired state χ_B. If she measures χ_A then Bob will have a particle in the state $\chi_B^\perp$. Alice sends one bit along a classical channel to tell Bob which state she measured and then Bob knows which state they have. This protocol is also secure in the sense that Eve needs access to both the classical and quantum channels to eavesdrop and even if she does Alice and Bob can tell that their communication has been interfered with.

30.2   Superdense Coding

Superdense coding is a protocol that uses only a quantum channel to communicate classical information. In this section we will consider two states, $|0\rangle$ and $|1\rangle$, and suppose Alice has a two bit message that she

wishes to send Bob. Alice and Bob start with an entangled state, for example

$$|\Phi^+\rangle = \frac{\sqrt{2}}{2}(|00\rangle + |11\rangle),$$

where the first bit corresponds to the state of Alice's particles and the second to the state of Bob's particle. If the message that Alice wants to send is the two bits (m, n) then the protocol is as follows:

1. If $m = 1$ Alice applies the unitary transformation $\hat{\sigma}_z \otimes \mathbb{1}$ to the state.

2. If $n - 1$ Alice applies the unitary transformation $\hat{\sigma}_z \otimes \mathbb{1}$ to the state.

3. Alice sends her particle to Bob without allowing it to change state.

4. Bob applies the control NOT, or CNOT, operation, $\hat{C}$, to the state.

5. Bob applies the Hadamard operation, $\hat{H}_d \otimes \mathbb{1}$, to the state.

6. The system will now be in the state $|mn\rangle$ and Bob simply makes measurements to find out the message.

The operators mentioned above have matrix representations. There are two representations that we use in this protocol. The first is the two dimensional tensor product representation where each state is simply a superposition of tensor products of states. Three operators above, $\hat{\sigma}_z$, $\hat{\sigma}_x$ and $\hat{H}_d$ are given in the two dimensional representation:

$$\hat{\sigma}_z \to \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \hat{\sigma}_x \to \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \text{and} \qquad \hat{H}_d \to \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Note that σ_i are simply the Pauli matrices. The third operator, $\hat{C}$, is best described in the four dimensional representation where

$$\lambda|00\rangle + \mu|01\rangle + \nu|10\rangle + \xi|11\rangle \to \begin{pmatrix} \lambda \\ \mu \\ \nu \\ \xi \end{pmatrix}.$$

In this representation

$$\hat{C} \to \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Suppose Alice wants to send the message $(m, n) = (1, 1)$ to Bob. In the tensor product representation the starting state of the system is

$$|\Phi^+\rangle \to \frac{\sqrt{2}}{2} \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right].$$

Since $m = 1$ Alice starts by applying $\hat{\sigma}_z \otimes \mathbb{1}$:

$$(\hat{\sigma}_z \otimes \mathbb{1})|\Phi^+\rangle \to \left[\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \mathbb{1} \right] \frac{\sqrt{2}}{2} \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]$$

$$= \frac{\sqrt{2}}{2} \left[\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \mathbb{1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \mathbb{1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]$$

$$= \frac{\sqrt{2}}{2} \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right].$$

Next since $n = 1$ Alice applies $\hat{\sigma}_x \otimes \mathbb{1}$:

$$(\hat{\sigma}_x \otimes \mathbb{1})(\hat{\sigma}_z \otimes \mathbb{1})|\Phi^+\rangle = (\hat{\sigma}_x \hat{\sigma}_z \otimes \mathbb{1})|\Phi^+\rangle$$

$$\to \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \mathbb{1} \right] \frac{\sqrt{2}}{2} \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]$$

$$= \frac{\sqrt{2}}{2}\left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \mathbb{1}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ -1 \end{pmatrix} \otimes \mathbb{1}\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right]$$

$$= \frac{\sqrt{2}}{2}\left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right]$$

Alice then sends this particle to Bob. First Bob applies the CNOT operator which is best expressed in the four dimensional representation. In this representation the state of the system is

$$(\hat{\sigma}_x\hat{\sigma}_z \otimes \mathbb{1})|\Phi^+\rangle \to \frac{\sqrt{2}}{2}\begin{pmatrix} 0 \\ -1 \\ 1 \\ 1 \end{pmatrix}.$$

So Bob applying the CNOT gives

$$\hat{C}(\hat{\sigma}_x\hat{\sigma}_z \otimes \mathbb{1})|\Phi^+\rangle \to \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}$$

$$= \frac{\sqrt{2}}{2}\begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}.$$

Moving back to the two dimensional tensor product representation we have

$$\hat{C}(\hat{\sigma}_x\hat{\sigma}_z \otimes \mathbb{1})|\Phi^+\rangle \to \frac{\sqrt{2}}{2}\left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right].$$

The last step is for Bob to apply $\hat{H}_d \otimes \mathbb{1}$:

$$(\hat{H}_d \otimes \mathbb{1})\hat{C}(\hat{\sigma}_x\hat{\sigma}_z \otimes \mathbb{1})|\Phi^+\rangle \to \left[\frac{\sqrt{2}}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \otimes \mathbb{1}\right]\frac{\sqrt{2}}{2}\left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right]$$

$$= \frac{1}{2}\left[\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \mathbb{1}\begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} -1 \\ 0 \end{pmatrix} \otimes \mathbb{1}\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right]$$

$$= \frac{1}{2}\left[\begin{pmatrix} 1 \\ -1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ -1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right]$$

$$= \frac{1}{2}\begin{pmatrix} 0 \\ -2 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ -1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\leftarrow -|11\rangle.$$

So when Bob measures the state they will find it is, up to a phase, $|11\rangle$ which corresponds to $(m,n) = (1,1)$.

31   Quantum Computing

31.1   Classical Computing

A classical computer deals with binary information. Everything it does can be boiled down to ones and zeros. These can be combined and manipulated in many ways. This gives rise to Boolean algebra. Any given operation can be expressed in a truth table which simply lists all possible inputs and the corresponding outputs. For example the addition of two bits is given by

a	b	s	c	Equation
0	0	0	0	$0 + 0 = 0$
0	1	1	0	$0 + 1 = 1$
1	0	1	0	$1 + 0 = 1$
1	1	0	1	$1 + 1 = 10_2 = 2_{10}$

Here a and b are the two inputs, s is the sum and c is a bit that carries over to the next position. Many operations can be in classical computing can be implemented through the logic gates NOT, AND, OR, XOR, NAND, NOR, and XNOR. The truth tables for these gates and the algebraic notation are shown in table 31.1. A computer is then simply a large collection of gates arranged to do more complicated

a	$\neg a$
0	1
1	0

(a) NOT

a	b	$a \wedge b$
0	0	0
0	1	0
1	0	0
1	1	1

(b) AND

a	b	$a \vee b$
0	0	0
0	1	1
1	0	1
1	1	1

(c) OR

a	b	$a \oplus b$
0	0	0
0	1	1
1	0	1
1	1	0

(d) XOR

a	b	$\neg(a \wedge b)$
0	0	1
0	1	1
1	0	1
1	1	0

(e) NAND

a	b	$\neg(a \vee b)$
0	0	1
0	1	0
1	0	0
1	1	0

(f) NOR

a	b	$\neg(a \oplus b)$
0	0	1
0	1	0
1	0	0
1	1	1

(g) XNOR

Table 31.1: The truth tables for the logic gates NOT, AND, OR, XOR, NAND, NOR, and XNOR.

things. These can be represented with circuit diagrams where each gate has a unique symbol, shown in figure 31.1. For example the addition operation defined above, called a half adder, can be constructed
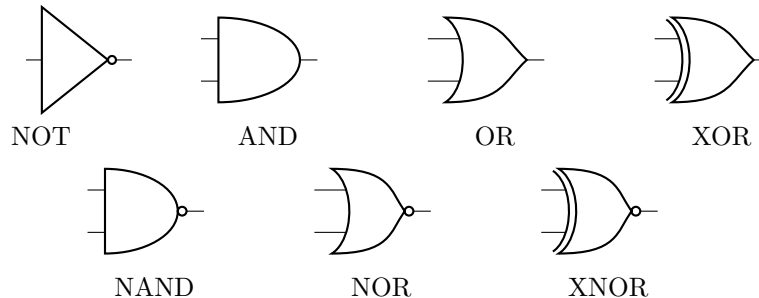

Figure 31.1: Circuit diagram symbols for NOT, AND, OR, XOR, NAND, NOR, and XNOR.

from an AND gate and an XOR gate as shown in figure 31.2. Comparing the truth tables XOR and
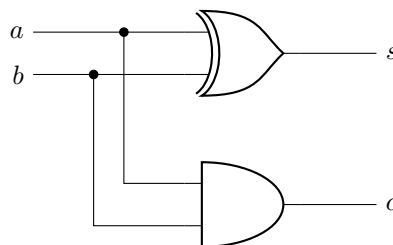

Figure 31.2: Half adder

AND we see that the outputs correspond to s and c respectively from the half adder.

31.2   Quantum Computing

The field of quantum computing essentially follows the ideas of classical computing but instead of electric currents representing bits we work with state vectors. Instead of gates and switches we evolve the state vector to the desired output.

Quantum computing first arose from a practical issue. Suppose we have a binary gate with two inputs and one output. Because this gate is not reversible some information is necessarily lost when it is used. The most efficient possible case would be to have one electron go into each input and then have one electron come out and the other dissipate away. The loss of information here leads to an entropy increase of $\ln 2$. If the system is at temperature T then this leads to generation of heat $k_{\mathrm{B}} T \ln 2$. In a real computer far more than one electron is used and this simply increases the heat released. While better designs can reduce the heat output of a computer there is still an intrinsic heat release due to information loss that simply cannot be removed. Another concern with classical computing is the breakdown of Moore's law. Moore's law, in its simplest form, states that the density of transistors on a chip doubles approximately every two years. The problem is that there is a quantum limit to this. At some point transistors will becomes so small that quantum effects such as tunnelling cause real problems and transistors simply can't be made smaller. Since almost all systems, including logic gates, are implemented with transistors this causes some real problems for the future of computing.

One solution to these problems was the suggestion that computers should be designed so that every operation is reversible. It was then suggested that quantum mechanics could be exploited to do this and so quantum computing was birthed. Another motivation was that quantum computers could be used to simulate quantum mechanical systems realistically in a way that simply isn't possible with classical computers. Many algorithms have been designed to work with quantum computers exploiting quantum mechanics to perform certain tasks orders of magnitude faster than is possible with classical computers.

The basic quantum computer works with a set of n-qubits. A full description then uses states in a Hilbert space of dimension 2^n. The n-qubits are formed into tensor product states and these sates are acted on by unitary operators leading to various superpositions of the 2^n basis states. The idea is then to use these superpositions to cancel out certain states and leave behind only the desired output. The goal is to find a combination of basic operations that leads to something computationally useful. This is then called a quantum algorithm.

31.3   Quantum Registers and Logic Gates

In a classical computer the bits of data are held in a register and then the data is accessed and combined using logic gates. The language used in quantum computing is similar to this however the implementation is not. A quantum register is a state vector. For example a one qubit register is simply the state of the qubit, $|\psi\rangle = a_0|0\rangle + a_1|1\rangle$. A two qubit register is state of the two qubit system,

$$|\psi\rangle = a_{00}|00\rangle + a_{01}|01\rangle + a_{10}|10\rangle + a_{11}|11\rangle.$$

Higher qubit registers are defined similarly. Note that a quantum register of multiple qubits is more complicated than simply multiple copies of a single qubit register as it allows for entanglement.

A standard way to think about quantum computing algorithms is to write the quantum register in the form

$$\sum_{c,t} a_{ct}|c\rangle_n \otimes |t\rangle_m.$$

Here the first terms, c, are called the control or input register and the second terms, t, are called the target or output register. These names don't necessarily match with the actual implementation of the algorithm but we aren't concerned about that here. The subscripts m and n are simply the number of qubits which is also the length of the binary number needed to specify a state. For example $|c\rangle_4$ is an input register that can take one of $2^4 = 16$ states labelled from 0 to 15.

A quantum logic gate is an operator that acts on the quantum register. Since we are interested in reversible computations most of the operators we consider are unitary and therefore preserve the norm of the quantum register. Since the time evolution operator is unitary we can think of these gates as simply controlling the Hamiltonian so that the state evolves in a desirable way. Non-unitary operators are also sometimes considered. For example making a measurement is generally non-unitary. It is also possible that interactions with the environment will cause decoherence in a real life application. It is

possible that non-unitary operators could be used to reverse these effects. Since we are only going to look at introductory quantum computing we will consider all operators to be unitary.

One operator that we have already seen is the Hadamard operator, $\hat{H}_d$. This acts on a single qubit. In the two dimensional spinor representation this gate is given by

$$\hat{H}_d \to \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

This can be written more abstractly as

$$\hat{H}_d = \frac{\sqrt{2}}{2}(|0\rangle\langle0| + |0\rangle\langle1| + |1\rangle\langle0| - |1\rangle\langle1|).$$

Gates can also be represented diagrammaticality as a labelled box:

$$-\boxed{\hat{H}_d}- \;.$$

The action of this operator on, for example, the state $|0\rangle$, is then given by

$$|0\rangle -\boxed{\hat{H}_d}- \tfrac{\sqrt{2}}{2}(|0\rangle + |1\rangle).$$

Other gates that we have met are the Pauli gates which in the two dimensional spinor representation are represented by the Pauli spin matrices,

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

These are written more abstractly as

$$\hat{\sigma}_x = \hat{X} = |1\rangle\langle0| + |0\rangle\langle1|, \qquad \hat{\sigma}_y = \hat{Y} = i|1\rangle\langle0| - i|0\rangle\langle1|, \qquad \text{and} \qquad \hat{\sigma}_z = \hat{Z} = |0\rangle\langle0| - |1\rangle\langle1|.$$

Or diagrammatically:

$$-\boxed{\hat{X}}- \;, \qquad -\boxed{\hat{Y}}- \;, \qquad -\boxed{\hat{Z}}- \;.$$

These can be generalised further. Fore example a generalisation of the Pauli-Z gate is the R_φ gate which in the two dimensional spinor representation is

$$\hat{R}_\varphi \to \begin{pmatrix} 1 & 0 \\ 0 & e^{i\varphi} \end{pmatrix},$$

or more abstractly

$$\hat{R}_\varphi = |0\rangle\langle0| + e^{i\varphi}|1\rangle\langle1|,$$

or diagrammatically

$$-\boxed{\hat{R}_\varphi}- \;.$$

There are also gates that act on two or more qubits at once. For example the CNOT gate represented by

$$\hat{C} \to \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

or more abstractly

$$\hat{C} = |00\rangle\langle00| + |01\rangle\langle01| + |10\rangle\langle11| + |11\rangle\langle10|,$$

or diagrammatically

$$\hat{C}$$

.

This gate can also be written in a more detailed form as

$$|c\rangle \quad\quad\quad |c\rangle$$
$$|t\rangle \quad\quad \hat{C} \quad\quad |t \oplus c\rangle.$$

Here $\oplus$ is simply addition modulo two which is the same as XOR. We can think of CNOT as the reversible version of XOR as the output is c and $t \oplus c$ which gives us enough information to determine c and t. This is also sometimes written as

$$|c\rangle \quad\quad\quad |c\rangle$$
$$|t\rangle \quad\quad \hat{X} \quad\quad |t \oplus c\rangle.$$

as we can view $\hat{C}$ as a block matrix:

$$\hat{C} \to \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \hat{X} \end{pmatrix}.$$

Note that the order of operators in a diagram is left to right in the order we would think of a circuit passing through gates whereas the order of operators in a product is right to left as with matrix multiplication. Hence

$$|\psi\rangle \quad\quad \hat{A} \quad\quad \hat{B} \quad\quad \hat{B}\hat{A}|\psi\rangle.$$

32   Deutsch Algorithm

The Deutsch algorithm is one of the simplest quantum computing algorithms and provides useful insight into the key features of many quantum computing algorithms. The Deutsch algorithm aims to answer the following question: Given a function, $f\colon \{1,0\} \to \{1,0\}$, that is a function from a bit to a bit, is $f(0) = f(1)$, that is is f constant? To answer this question classically we would need to compute $f(0)$ and $f(1)$ and compare the results. This corresponds to two measurements. The Deutsch algorithm gives us a way to make one measurement and it will tell us if f is constant or not. It should be noted that the actual value is not known after this algorithm, just if it is the same for both inputs or different.

A variable that we wish to investigate in quantum computing is often encoded in an operator, $\hat{O}$, usually called the **oracle**. The exact specifics of how this is done depend on the object we want to study. In this case we wish to know about f and we define the oracle as an operator on the register with the action

$$\hat{O}_f(|c\rangle \otimes |t\rangle) = |c\rangle \otimes |c \oplus f(c)\rangle,$$

where $\oplus$ is addition modulo 2. This can be represented diagrammatically as

$$|c\rangle \quad\quad\quad |c\rangle$$
$$|t\rangle \quad\quad \hat{O}_f \quad\quad |t \oplus f(c)\rangle.$$

Since the domain and codomain of f are small we can exactly specify all four possible functions:

f_i	$f_i(0)$	$f_i(1)$
f_1	0	0
f_2	0	1
f_3	1	0
f_4	1	1

We can encode each of these into an oracle, $\hat{O}_{f_i}$. For example the action of $\hat{O}_{f_2}$ is

$$\hat{O}_{f_2}|0\rangle \otimes |0\rangle = |0\rangle \otimes |0 \oplus f_2(0)\rangle = |0\rangle \otimes |0 \oplus 0\rangle = |0\rangle \otimes |0\rangle,$$
$$\hat{O}_{f_2}|0\rangle \otimes |1\rangle = |0\rangle \otimes |1 \oplus f_2(0)\rangle = |0\rangle \otimes |1 \oplus 0\rangle = |0\rangle \otimes |1\rangle,$$
$$\hat{O}_{f_2}|1\rangle \otimes |0\rangle = |1\rangle \otimes |0 \oplus f_2(1)\rangle = |1\rangle \otimes |0 \oplus 1\rangle = |1\rangle \otimes |1\rangle,$$
$$\hat{O}_{f_2}|1\rangle \otimes |1\rangle = |1\rangle \otimes |1 \oplus f_2(1)\rangle = |1\rangle \otimes |1 \oplus 1\rangle = |1\rangle \otimes |0\rangle.$$

Notice that this is the same action as $\hat{C}$.

The next step is the actual algorithm which is just a list of operators to apply and a final measurement. The process for finding the appropriate operators is mostly trial and error and not very illuminating so we will just give the final algorithm. Start with the register $|0\rangle \otimes |0\rangle$, it is common for quantum algorithms to start with a state of the form $|0\rangle \otimes \cdots \otimes |0\rangle$. Compute

$$(\hat{H}_d \otimes \mathbb{1})\hat{O}_f(\hat{H}_d \otimes \hat{H}_d)(\hat{X} \otimes \hat{X})(|0\rangle \otimes |0\rangle).$$

Measure the state. If you get $|c\rangle = |0\rangle$ then the function is constant. If $|c\rangle = |1\rangle$ then the function is not constant. We will now carry out this algorithm.

First we need to compute
$$(\hat{X} \otimes \hat{X})(|0\rangle \otimes |0\rangle).$$

We do this in the two dimensional tensor product representation:

$$\hat{X}|0\rangle \rightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
$$= \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
$$\leftarrow |1\rangle$$

So
$$(\hat{X} \otimes \hat{X})(|0\rangle \otimes |0\rangle) = (\hat{X}|0\rangle) \otimes (\hat{X}|0\rangle) = |1\rangle \otimes |1\rangle.$$

Next we compute
$$(\hat{H}_d \otimes \hat{H}_d)(|1\rangle \otimes |1\rangle)$$

again in the two dimensional tensor product representation we have

$$\hat{H}_d|1\rangle \rightarrow \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
$$= \frac{\sqrt{2}}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$
$$= \frac{\sqrt{2}}{2} \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]$$
$$\leftarrow \frac{\sqrt{2}}{2}(|0\rangle - |1\rangle).$$

So

$$(\hat{H}_d \otimes \hat{H}_d)(|1\rangle \otimes |1\rangle) = (\hat{H}_d|1\rangle) \otimes (\hat{H}_d \otimes |1\rangle)$$
$$= \frac{1}{2}(|0\rangle - |1\rangle) \otimes (|0\rangle - |1\rangle)$$
$$= \frac{1}{2}(|0\rangle \otimes |0\rangle - |0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle).$$

Next we act with the oracle:

$$|\psi\rangle = \hat{O}_f \frac{1}{2}(|0\rangle \otimes |0\rangle - |0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle)$$
$$= \frac{1}{2}(|0\rangle \otimes |0 \oplus f(0)\rangle - |0\rangle \otimes |1 \oplus f(0)\rangle - |1\rangle \otimes |0 \oplus f(1)\rangle + |1\rangle \otimes |1 \oplus f(1)\rangle)$$

$$= \frac{1}{2}[|0\rangle \otimes (|0 \oplus f(0)\rangle - |1 \oplus f(0)\rangle) + |1\rangle \otimes (-|0 \oplus f(1)\rangle + |1 \oplus f(1)\rangle)].$$

Finally we apply the Hadamard operator to only the control state. In the two dimensional tensor product representation we have

$$\hat{H}_d|0\rangle \to \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \leftarrow \frac{\sqrt{2}}{2}(|0\rangle + |1\rangle)$$

$$\hat{H}_d|1\rangle \to \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} \leftarrow \frac{\sqrt{2}}{2}(|0\rangle - |1\rangle)$$

and so

$$(\hat{H}_d \otimes \mathbb{1})|\psi\rangle = \frac{\sqrt{2}}{4}[(|0\rangle + |1\rangle) \otimes (|0 \oplus f(0)\rangle - |1 \oplus f(0)\rangle) + (|0\rangle - |1\rangle) \otimes (-|0 \oplus f(1)\rangle + |1 \oplus f(1)\rangle)]$$

$$= \frac{\sqrt{2}}{4}|0\rangle \otimes (|0 \oplus f(0)\rangle - |1 \oplus f(0)\rangle - |0 \oplus f(1)\rangle + |1 \oplus f(1)\rangle)$$

$$+ \frac{\sqrt{2}}{4}|1\rangle \otimes (|0 \oplus f(0)\rangle - |1 \oplus f(0)\rangle + |0 \oplus f(1)\rangle - |1 \oplus f(1)\rangle)$$

$$= \frac{\sqrt{2}}{4}|0\rangle \otimes (|f(0)\rangle - |1 \oplus f(0)\rangle - |f(1)\rangle + |1 \oplus f(1)\rangle)$$

$$+ \frac{\sqrt{2}}{4}|1\rangle \otimes (|f(0)\rangle - |1 \oplus f(0)\rangle + |f(1)\rangle - |1 \oplus f(1)\rangle)$$

Now consider the case when f is constant so $f(0) = f(1)$. We then have

$$(\hat{H}_d \otimes \mathbb{1})|\psi\rangle = \frac{\sqrt{2}}{4}|0\rangle \otimes (|f(0)\rangle - |1 \oplus f(0)\rangle - |f(0)\rangle + |1 \oplus f(0)\rangle)$$

$$+ \frac{\sqrt{2}}{4}|1\rangle \otimes (|f(0)\rangle - |1 \oplus f(0)\rangle + |f(0)\rangle - |1 \oplus f(0)\rangle)$$

$$= \frac{\sqrt{2}}{2}|1\rangle \otimes (|f(0)\rangle - |1 \oplus f(0)\rangle).$$

So if we measure this state we will find $|c\rangle = |1\rangle$. If instead $f(0) \neq f(1)$ then $1 \oplus f(0) = 0 \oplus f(1)$ and $0 \oplus f(0) = 1 \oplus f(1)$. Also $f(0) = 1 \oplus f(1)$ and $f(1) = 1 \oplus f(0)$. Thus

$$(\hat{H}_d \otimes \mathbb{1})|\psi\rangle = \frac{\sqrt{2}}{4}|0\rangle \otimes (|f(0)\rangle - |1 \oplus f(0)\rangle - |1 \oplus f(0)\rangle + |f(0)\rangle)$$

$$+ \frac{\sqrt{2}}{4}|1\rangle \otimes (|f(0)\rangle - |1 \oplus f(0)\rangle + |1 \oplus f(0)\rangle - |f(0)\rangle)$$

$$= \frac{\sqrt{2}}{2}|0\rangle \otimes (|f(0)\rangle - |1 \oplus f(0)\rangle).$$

So if upon measuring the register we get $|c\rangle = |0\rangle$ we know that $f(0) \neq f(1)$.

32.1   Constructing The Oracle

When defining a quantum algorithm one should show that all operators are unitary and, if possible, that they can be constructed from elementary gates such as $\hat{C}$, $\hat{X}$, etc. This is useful as it would allow a general quantum computer to use the algorithm without needing special hardware. The only new operator in the Deutsch algorithm is the oracle. It is fairly simple to show directly that the oracles can be constructed as follows:

$$\hat{O}_{f_1} = \mathbb{1},$$
$$\hat{O}_{f_2} = \hat{C},$$
$$\hat{O}_{f_3} = \hat{C}(\mathbb{1} \otimes \hat{X}),$$

$$\hat{O}_{f_4} = (\hat{I} \otimes \hat{X}).$$

We can construct 4×4 matrices representing these. To do so we need to define the tensor product of two 2×2 matrices. Let A and B be 2×2 matrices with components a_{ij} and b_{ij}. Then

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{21} & a_{12}b_{21} & a_{12}b_{21} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{21} & a_{22}b_{21} & a_{22}b_{21} \end{pmatrix}.$$

This generalises to higher dimensional matrices as well. This allows us to write terms like $\mathbb{1} \otimes \hat{X}$ as 4×4 matrices. Thus the four oracles in this representation and as quantum circuits are given by

$$\hat{O}_{f_1} = \mathbb{1}$$

$$\rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{array}{c} |c\rangle \;\rule{1.5em}{0.4pt}\; |c\rangle \\[1em] |t\rangle \;\rule{1.5em}{0.4pt}\; |t \oplus f_1(c)\rangle \end{array}$$

$$\hat{O}_{f_2} = \hat{C}$$

$$\rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$= \begin{array}{c} |c\rangle \;\rule{1.5em}{0.4pt}\bullet\rule{1.5em}{0.4pt}\; |c\rangle \\[1em] |t\rangle \;\boxed{\hat{X}}\; |t \oplus f_2(c)\rangle \end{array}$$

$$\hat{O}_{f_3} = \hat{C}(\mathbb{1}\hat{X})$$

$$\rightarrow \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{array}{c} |c\rangle \;\rule{1.5em}{0.4pt}\bullet\rule{1.5em}{0.4pt}\; |c\rangle \\[1em] |t\rangle \;\boxed{\hat{X}}\;\boxed{\hat{X}}\; |t \oplus f_3(c)\rangle \end{array}$$

$$\hat{O}_{f_4} = \mathbb{1} \otimes \hat{X}$$

$$\rightarrow \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$= \begin{array}{c} |c\rangle \;\rule{1.5em}{0.4pt}\; |c\rangle \\[1em] |t\rangle \;\boxed{\hat{X}}\; |t \oplus f_4(c)\rangle \end{array}$$

33   Grover's Algorithm

Grover's algorithm can be stated in many ways. It can be thought of as a search algorithm which finds a given state in a quantum register or as a way to probe an unknown oracle and find the state that it

picks out. The algorithm starts with the state $|0\rangle^{\otimes N}$ where

$$|0\rangle^{\otimes N} = \underbrace{|0\rangle \otimes |0\rangle \otimes \cdots \otimes |0\rangle}_{N \text{ times}},$$

and N is the number of things we have to search through. Then the Hadamard operator is applied to each term which results in the state $\hat{H}_d^{\otimes N}|0\rangle^{\otimes N}$ which is a linear combination of all basis states such that all states have equal probability. We will consider for an example here the case $n = 3$. The starting state is then $|000\rangle$ and the first step gives us

$$|\psi\rangle = (\hat{H}_d \otimes \hat{H}_d \otimes \hat{H}_d)|000\rangle = \frac{\sqrt{2}}{4}(|000\rangle + |001\rangle + |010\rangle + |100\rangle + |011\rangle + |101\rangle + |110\rangle + |111\rangle).$$

Suppose we have an oracle that picks out the state $|s\rangle$. The Grover algorithm gives us a way to move the system towards this state in a unitary manner. This is important as it means that operations are reversible and we can continue to use the state for further computations after the algorithm. Compare this to the normal way we pick out a state by simply taking the inner product with $\langle s|$.

The next step of the Grover algorithm is to apply the oracle, $\hat{O}$, which flips the sign of the state, $|s\rangle$, which it is singling out. This oracle is given by

$$\hat{O} = \mathbb{1} - 2|s\rangle\langle s|.$$

This has a diagonal matrix representation of ones down the diagonal except for the entry associated with $|s\rangle$ which has a -1.

Suppose for our example that $|s\rangle = |110\rangle$. Then the action of the oracle is

$$|\psi_-\rangle = \hat{O}|\psi\rangle = \frac{\sqrt{2}}{4}(|000\rangle + |001\rangle + |010\rangle + |100\rangle + |011\rangle + |101\rangle - |110\rangle + |111\rangle).$$

The only change is that the coefficient of $|s\rangle = |110\rangle$ is now negative.

The next step of the algorithm is to apply the Grover operation,

$$\hat{G} = 2|\psi\rangle\langle\psi| - \mathbb{1}$$

which for our example gives us

$$|\psi_d\rangle = \frac{\sqrt{2}}{8}(|000\rangle + |001\rangle + |010\rangle + |100\rangle + |011\rangle + |101\rangle + 5|110\rangle + |111\rangle).$$

Notice that the net effect of $\hat{G}\hat{O}$ is to increase the probability of measuring the register in the state $|s\rangle$. It turns out that applying $\hat{G}\hat{O}$ iteratively $\sqrt{N}$ times will cause $|s\rangle$ to have its maximum amplitude. So if we do this then measure the register it is likely that we will measure the state $|s\rangle$.

We can better understand this algorithm by considering the action of $\hat{O}$ and $\hat{G}$ on the desired state $|s\rangle$ and the initial state $|\psi\rangle = \hat{H}_d^{\otimes N}|0\rangle^{\otimes N}$:

$$\hat{O}|s\rangle = -|s\rangle,$$
$$\hat{O}|\psi\rangle = |\psi\rangle - \frac{2}{\sqrt{N}}|s\rangle,$$
$$\hat{G}|s\rangle = \frac{2}{\sqrt{N}}|\psi\rangle - |s\rangle,$$
$$\hat{G}|\psi\rangle = |\psi\rangle.$$

We see that the action of both operators always returns a linear combination of $|s\rangle$ and $|\psi\rangle$. This means that the action of $\hat{G}\hat{O}$ on $|\psi\rangle$ will be another state in the two dimensional subspace spanned by $|\psi\rangle$ and $|s\rangle$.

We can write an arbitrary state, $|u\rangle \in \mathcal{H}$, as

$$|u\rangle = a|s_\perp\rangle + b|s\rangle$$

Page 134

where $|s_\perp\rangle$ is some state perpendicular to $|s\rangle$ so $\langle s|s_\perp\rangle = 0$. Note that this is not a basis for the whole space as another state, $|v\rangle \in \mathcal{H}$ may require a different $|s_\perp\rangle$ but all states can be decomposed this way. Similarly we can write an arbitrary state, $|u'\rangle \in \mathcal{H}$ as

$$|u'\rangle = a|\psi\rangle + b|\psi_\perp\rangle$$

where $|\psi_\perp\rangle$ is perpendicular to $|\psi\rangle$ so $\langle\psi|\psi_\perp\rangle = 0$.

The first step of Grover's algorithm is to act on $|\psi\rangle$ with $\hat{O}$. We can decompose $|\psi\rangle$ in terms of $|s\rangle$ and $|s_\perp\rangle$. First note that since each coefficient in $|\psi\rangle$ is equal for a properly normalised state we must have $\langle s|\psi\rangle = N^{-1/2}$ and so

$$\hat{O}|\psi\rangle = \hat{O}\left(\sqrt{\frac{N-1}{N}}|s_\perp\rangle + \frac{1}{\sqrt{N}}|s\rangle\right)$$

$$= \sqrt{\frac{N-1}{N}}|s_\perp\rangle - \frac{1}{\sqrt{N}}|s\rangle.$$

We can consider this geometrically as a reflection of $|\psi\rangle$ in the $|s\rangle$ axis, see figure 33.1. Notice that if $|\psi\rangle$
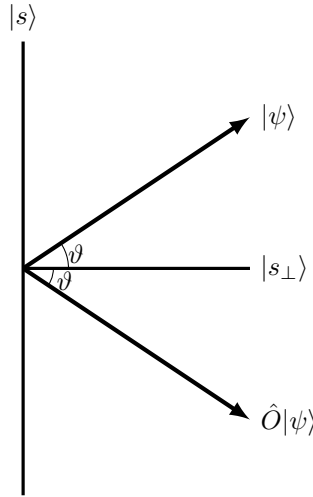
Figure 33.1: The action of $\hat{O}$ on $|\psi\rangle$ can be viewed as a reflection in the $|s\rangle$ axis.

makes an angle[14] of ϑ to $|s_\perp\rangle$ then the action of $\hat{O}$ is a rotation through -2ϑ.

The action of $\hat{G}$ can be similarly interpreted as a reflection in the line defined by the original state $|\psi\rangle$. This can also be viewed as a rotation through 4ϑ and so $\hat{G}\hat{O}$ has a net result of a rotation of 2ϑ towards $|s\rangle$.

Each iterative application of $\hat{G}\hat{O}$ rotates the state 2ϑ towards $|s\rangle$. At some point we will overshoot $|s\rangle$ and start moving away again. We need to know how many times to apply $\hat{G}\hat{O}$ to maximise the probability of measuring the register to be in state $|s\rangle$.

The initial angle, ϑ, is related to the original state as shown in figure 33.2. We see that

Figure 33.2: $|\psi\rangle$ decomposed into $|s_\perp\rangle$ and $|s\rangle$ components.

$$\sin\vartheta = \frac{1}{\sqrt{N}}.$$

---

[14]for some definition of angles in a non-Euclidean complex vector space

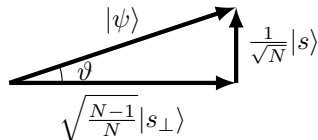For large N ϑ must be small so we have $\vartheta \approx N^{-1/2}$. Similarly $|\psi\rangle$ is almost entirely aligned with $|s_\perp\rangle$ so we need to rotate it through an angle of $\pi/2$ to be aligned in the $|s\rangle$ direction. Since each iteration moves us an angle $2\vartheta \approx 2/\sqrt{N}$ we need $\pi/(4\vartheta) = \pi\sqrt{N}/4$ (or the closest integer to this) iterations to reach maximal probability of measuring $|s\rangle$.

To implement this algorithm one needs to be able to construct $\hat{O}$ and $\hat{G}$ from basic gates. This can be done but is not very illuminating and quite involved so we won't do it here.

We mentioned at the start of this section that the Grover algorithm can be seen as a search algorithm. The context for this is as follows: suppose we have a quantum register, $|\varphi\rangle$, and we wish to know if a given state, $|s\rangle$, is in this register. Then we use the Grover algorithm as described above and after $\sqrt{N}$ iterations we measure the register and we expect to measure $|s\rangle$ with high probability[15] if $|s\rangle$ is in $|\varphi\rangle$. From this we conclude that the Grover algorithm, when thought of as a search algorithm searching N qubits, is $\mathcal{O}(\sqrt{N})$ since we need to apply $\hat{G}\hat{O}$ $\mathcal{O}(\sqrt{N})$ times. This is much faster than similar classical search algorithms which are $\mathcal{O}(N)$.

Quantum computing is still a developing field. One of the key features is the ability to explore multiple solutions at once by studying superpositions of states. For example a system of n spin $1/2$ particles has $N = 2^n$ possible states so by studying 10 particles we can study $2^{10} = 1024$ different states. The number of different states grows exponentially with the size of the system. If each state corresponds with a possible solution then we can test many solutions at once.

One hope that is held for quantum computing is the ability to solve problems that fall in a class called NP problems. NP stands for non-deterministic polynomial time which, roughly speaking, means that once a solution is found we can check that this is a solution in polynomial time. That is for a given NP problem there is an algorithm that will return true or false for a given solution and that algorithm is $\mathcal{O}(N^n)$ for some integer n and input size N. While a solution for a NP problem can be found in polynomial time whether a solution can be found in polynomial time for all NP problems is an open question[16].

For an example of an NP problem consider the task of finding a particular number in a set of all n-digit numbers which are randomly ordered. If we were to simply check every number then the time take grows exponentially with n (as there are $\mathcal{O}(10^n)$ n-digit numbers). On the other hand if someone has already found the number in the list then we simply need to check each digit to check it is the correct number and the number of digits grows linearly with n (as the number of digits *is* n).

It is possible that a quantum computing algorithm exists which allows us to solve NP problems faster. Suppose we have a system of n spin $1/2$ particles. If we can somehow find a correspondence between the 2^n states with set of solutions and apply a quantum algorithm to the system which enhances the probability amplitude of the correct answer it is possible that this would lead to an answer faster than any classical algorithm can find a solution. Grover's algorithm is the most simple example where this is possible. We manipulate the probability amplitudes to find the desired state in $\mathcal{O}(\sqrt{N})$ as opposed to the classical $\mathcal{O}(N)$ algorithms.

34   Information Theory

34.1   Classical Information Theory

Information theory is the study of messages and communications with the goal of developing efficient communication. The formal meaning of this is as follows. A (formal) **language** is a set of N symbols, $\Sigma = \{s_1, \ldots, s_N\}$. These can be arranged into a **message**, X, which is an ordered set of symbols of length M_X, $X = s_i^1 s_j^2 \cdots s_k^{M_X}$ where superscripts denote position in the message and subscripts take values from 1 to N such that $s_i \in \Sigma$. At this point there is an important decision to make. We need to decide what we care about. If we only care about a few possible messages then we can be very efficient by simply assigning each a number and sending that number. If we want to be able to send arbitrary messages then we need to consider the entire language. In the following this decision reveals itself whenever there is a sum or we count something we need to choose to sum over or count the part of the language that we care about.

---

[15]with high probability is used here in the technical sense that an event with probability on N occurs with high probability iff that probability goes to one as $N \to \infty$.

[16]see P vs. NP

If we denote by M_i^X then number of times the symbol s_i appears in the message X then the frequency, or probability, of s_i is $f_i = p_i = M_i^X/M_X$. We say that the **information** of the symbol s_i with probability p_i is

$$I(p_i) = -\log_2 p_i.$$

Notice that this logarithm is taken in base 2. This is because we want to measure information in bits. The information of a symbol is a useful quantity when we are studying a large number of symbols. Roughly the information characterises how much we learn when we are sent a particular symbol. In the extreme case of a language with only one symbol or where we restrict messages to be only one message the information is zero as we know what every message is before it is sent. In a language with two symbols with equal probability the information of either symbol is 1 as we learn one thing when the message is sent (we learn which symbol was sent).

The **Shannon entropy** of a message is the average information of all symbols:

$$S_{\text{Sh}} = -\sum_i p_i \log_2 p_i.$$

This has units of bits per symbol. The Shannon source coding theorem states that the Shannon entropy is the minimum number of bits needed to encode the message.

For a first example suppose we have a language of two symbols, $\{A, B\}$. Suppose that the probability of A is p_A. Then the Shannon entropy is

$$S_{\text{Sh}} = -p_A \log_2 p_A - (1 - p_A) \log_2 p_A.$$

This is zero in the limits when $p_A = 0, 1$ and peaks at $p_A = 0.5$ with $S_{\text{Sh}} = 1\,\text{bits/symbol}$[17]. The case of $p_A = 0.5$ is equivalent to choosing randomly between A and B and the Shannon source coding theorem simply states that it requires one bit of information to tell someone which one we chose.

For a more realistic (but only slightly) example we need more symbols. Consider a language of four symbols, $\{A, B, C, D\}$, with probabilities 0.5, 0.25, 0.125, and 0.125 respectively. The Shannon entropy of this language is

$$S_{\text{Sh}} = -0.5 \log_2 0.5 - 0.25 \log_2 0.25 - 0.125 \log_2 0.125 - 0.125 \log_2 0.125 = 1.75\,\text{bits/symbol}.$$

So we need at least $1.75\,\text{bits/symbol}$ to encode a character in this language. For example consider the following encoding:

Symbol	Encoding
A	1
B	01
C	001
D	000

This encoding seems a bit odd at first, why not just use 0 for A, 1 for B, 10 for C and 11 for D? Because then there is ambiguity, for example 10 could be either AB or C. Computers don't have this issue as they work with fixed blocks of bits or use an extra symbol to denote the end of a symbol but this is not the most efficient way to encode. Notice that the information of each of the symbols above is equal to the number of bits in the encoding. Suppose we want to encode the message $AABABACD$. Then our encoded message is 11011011001000. This is 14 bits and as the message is 8 symbols long the information per symbol is $14\,\text{bits}/8\,\text{symbol} = 1.75\,\text{bits/symbol}$. So we have achieved maximum efficiency in our encoding. Note that this limit is for the *average* message. If we want to send $AAAA$ the this can be done in four bits so the Shannon entropy is $1\,\text{bits/symbol}$. Similarly if we want to send $CDCD$ then this requires 12 bits so the Shannon entropy is $3\,\text{bits/symbol}$.

The idea when encoding a language is that we want the highest frequency bit to have the shortest possible encoding, hence why we chose A to have a single bit encoding in the example above. There is one way that we can improve on this information rate. We can encode words. Suppose that the string we encoded before, $AABABACD$, is representative of all messages in the language. Notice that AA appears once, BA appears twice, and CD appears once. The Shannon entropy of $\{AA, BA, CD\}$ is $1.5\,\text{bits/symbol}$. If we encode AB as 1, AA as 01 and CD as 001 then we can encode our message as 0111001 which is

---

[17]This function may be familiar from the entropy in statistical mechanics

7 bits for 14 symbols[18], so one bit better than before. However there are now messages we can't encode, for example $BACCDB$.

34.2   Quantum Information Theory

34.2.1   Density Matrix

The stats that we have dealt with so far are what are called **pure states**. This means that they are a linear combination of basis states. The other type of state is a **mixed state**. These are collections of pure states with some classical probability associated with them. These most commonly arise when we aren't certain what pure state a system is in but we do know what the probability of a given state is. The probabilities here are classical in that they aren't inherent to the system but are due to our lack of knowledge of the system.

To account for both possibilities we define the **density matrix**:

$$\hat{\rho} = \sum_i p_i |\chi_i\rangle\langle\chi_i|$$

where p_i is the classical probability of the state $|\chi_i\rangle$ from the mixed state. The density matrix for a pure state, $|\chi\rangle$, is simply the projection operator, $|\chi\rangle\langle\chi|$. Some useful properties of the density matrix are:

- *Hermitian* – This follows since $\dagger$ distributes over addition and $(|\chi\rangle\langle\chi|)^\dagger = |\chi\rangle\langle\chi|$.

- *Trace property* – The trace of the density matrix is $\text{Tr}(\hat{\rho}) = 1$. To see this let $\{|\varphi_i\rangle\}$ be a complete set of states. Then

$$\begin{aligned}
\text{Tr}(\hat{\rho}) &= \sum_j \rho_{jj} \\
&= \sum_j \langle\varphi_j|\hat{\rho}|\varphi_j\rangle \\
&= \sum_{ij} \langle\varphi_j|p_i|\chi_i\rangle\langle\chi_i|\varphi_j\rangle \\
&= \sum_i p_i \sum_j \langle\varphi_j|\chi_i\rangle\langle\chi_i|\varphi_j\rangle \\
&= \sum_i p_i \sum_j \langle\chi_i|\varphi_j\rangle\langle\varphi_j|\chi_i\rangle \\
&= \sum_i p_i \langle\chi_i|\chi_i\rangle \\
&= \sum_i p_i \\
&= 1.
\end{aligned}$$

- *Positive-semidefinite* – For all $|\varphi\rangle \in \mathcal{H}$

$$\langle\varphi|\hat{\rho}|\varphi\rangle = \sum_i p_i\langle\varphi|\chi_i\rangle\langle\chi_i|\varphi\rangle = \sum_i p_i|\langle\varphi|\chi_i\rangle|^2 \geq 0.$$

- *Eigenvalues lie in* $[0,1]$ – Let α_i be the eigenvalues of $\hat{\rho}$. Since $\hat{\rho}$ is Hermitian it only has real eigenvalues. Since $\hat{\rho}$ is positive-semidefinite it's eigenvalues are non-negative. In the eigenbasis of $\hat{\rho}$ it is diagonal with it's eigenvalues on the diagonal. The trace is invariant under change of basis so the trace is 1 meaning $\sum_i \alpha_i = 1$. The only way for positive real numbers to sum to one is if all of those numbers are less than 1.

- *Trace property for* $\hat{\rho}^2$ – $\text{Tr}(\hat{\rho}^2) \leq 1$. In the eigenbasis of $\hat{\rho}$ we have $\text{Tr}(\hat{\rho}^2) = \sum_i \alpha_i^2$. By the triangle inequality

$$1 = \left[\sum_i \alpha_i\right] \geq \sum_i \alpha_i$$

---

[18] the hit, quantum physics, sequel to seven brides for seven brothers featuring hit songs such as barn raise(ing and lowering operators) and bless your beautiful hide(rogen atom)

Consider now the special case of a pure state. Then $\hat{\rho}^2 = \sum_i \alpha_i^2 |\alpha_i\rangle\langle\alpha_i|$ so $\hat{\rho}^2 = \hat{\rho}$. Hence if $\text{Tr}(\hat{\rho}^2) = 1$ or $\hat{\rho}^2 = \hat{\rho}$ then $\hat{\rho}$ represents a pure state.

34.2.2   Von Neumann Entropy

The **von Neummann entropy** is defined as

$$S_{\text{vN}} = -\text{Tr}(\hat{\rho}\log_2\hat{\rho}),$$

where $\log_2$ is applied element wise. It is, in many ways, the quantum analogue of the Shannon entropy. Consider, for example, the case of a mixed state where each state, $|\varphi_i\rangle$, appears with probability p_i, and $\{|\varphi_i\rangle\}$ are mutually orthogonal. In this case we can express the density matrix as a diagonal matrix in the $\{|\varphi_i\rangle\}$ basis. In this case

$$
\begin{aligned}
S_{\text{vN}} &= -\text{Tr}[\hat{\rho}\log_2\hat{\rho}] \\
&= -\text{Tr}\left[\begin{pmatrix} p_1 & 0 & \dots & 0 \\ 0 & p_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & p_N \end{pmatrix}\begin{pmatrix} \log_2 p_1 & 0 & \dots & 0 \\ 0 & \log_2 p_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \log_2 p_N \end{pmatrix}\right] \\
&= -\text{Tr}\begin{pmatrix} p_1\log_2 p_1 & 0 & \dots & 0 \\ 0 & p_2\log_2 p_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & p_N\log_2 p_N \end{pmatrix} \\
&= \sum_i p_i\log_2 p_i \\
&= S_{\text{Sh}}.
\end{aligned}
$$

So for orthogonal mixed states the von Neumann entropy reduces to the Shannon entropy. Consider instead the case of a pure state, $|\varphi\rangle$. The density matrix is the $\hat{\rho} = |\varphi\rangle\langle\varphi|$. In this case $\hat{\rho}$ can be expressed as a matrix that is zero everywhere apart from the diagonal element corresponding to $|\varphi\rangle$ in some basis including $|\varphi\rangle$ as a basis vector. In this slot it will be 1. Since $\log_2 1 = 0$ the von Neumann entropy will be zero. This is analogous to the case of a language of a single symbol.

The von Neumann entropy starts to differ from the Shannon entropy when we consider a mixture of non-orthogonal quantum states. For example consider the set of two spin states, one spin up along the z-axis, $|+\rangle$, and one spin up along the y axis, $|y\rangle = \sqrt{2}(|-\rangle + i|+\rangle)/2$. Suppose both states occur with equal probability in the mixed state. The density matrix in the $\{|+\rangle, |-\rangle\}$ basis is then

$$
\begin{aligned}
\hat{\rho} &= 0.5|+\rangle\langle+| + 0.5\left[\frac{\sqrt{2}}{2}(i|+\rangle + |-\rangle)\right]\left[\frac{\sqrt{2}}{2}(-i|+\rangle + |-\rangle)\right] \\
&= \left(0.5 + i\frac{1}{2}\right)|+\rangle\langle+| + 0.5i\frac{1}{2}|+\rangle\langle-| - 0.5i\frac{1}{2}|-\rangle\langle+| + 0.5\frac{1}{2}|-\rangle\langle-| \\
&= \begin{pmatrix} 0.75 & 0.25i \\ -0.25i & 0.25 \end{pmatrix}.
\end{aligned}
$$

The eigenvalues of this matrix are $\alpha_1 = 0.146$ and $\alpha_2 = 0.854$. The von Neumann entropy is then

$$S_{\text{vN}} = -0.146\log_2 0.146 - 0.854\log_2 0.854 = 0.601\,\text{bits}.$$

This differs from the Shannon entropy for two symbols which occur with equal probability in which case $S_{\text{Sh}} = 1\,\text{bits/symbol}$. In general $S_{\text{vN}} \leq S_{\text{Sh}}$. Equality holds only when all states in the mixed state are orthogonal.

34.2.3   Quantum Information Theory

In quantum mechanics the information is encoded into the quantum states. We have seen this already with quantum computing. Quantum information theory seeks to find out how much information a state can possibly hold. In the quantum case states replace the language so suppose we have a collection of

states, $\{|i\rangle\}$, which may or may not be orthogonal. Consider the ensemble where each of these states exists with probability p_i. The density matrix for this ensemble is

$$\hat{\rho} = \sum_i p_i |i\rangle\langle i|.$$

The collection of all possible messages of n states is then given by the tensor product density matrix:

$$\hat{\rho}^{\otimes n} = \underbrace{\hat{\rho} \otimes \hat{\rho} \otimes \cdots \otimes \hat{\rho}}.$$

The naive way to send n quantum states is to use a Hilbert space of dimension n. However the quantum analogue of the the Shannon source coding theorem, which is called the Schumacher coding theorem, tells us that the smallest dimension of the Hilbert space needed to send n quantum states satisfies

$$\log_2(\dim \mathcal{H}) = n S_{\text{vN}}(\hat{\rho}).$$

For example suppose we have the two orthogonal states $|+\rangle$ and $|-\rangle$ and they appear with equal probability. Then

$$\hat{\rho} = 0.5|+\rangle\langle+| + 0.5|-\rangle\langle-|,$$

and the von Neumann entropy is

$$S_{\text{vN}} = 2(-0.5\log_2 0.5) = 1.$$

So each state we wish to send requires the Hilbert space to have an extra dimension. This is not that different from the classical case. We can simply assign one state in the Hilbert space for each state. We can extend this to states including many orthogonal states such as $|+--+\rangle$.

Again the difference arises when we have states which aren't orthogonal. In this case the von Neumann entropy is less than the Shannon entropy. This means that we can compress the data more than is allowed classically. This is because by virtue of the states being non-orthogonal they have some overlap in information and this has no classical counterpart.

One way to view this overlap is explained in the following example. Consider the case of the two, non-orthogonal states $|+\rangle$ and $|y\rangle$ defined in the previous section. Suppose that we wish to send $|+\rangle$ but instead $|y\rangle$ is sent. All is not yet lost as $|y\rangle = 2(i|+\rangle + |-\rangle)/2$ still contains $|+\rangle$ with probability $1/2$ of measuring $|+\rangle$. The amount in which two states, $|\varphi\rangle$ and $|\psi\rangle$, overlap is measured by the **fidelity**, F, defined to be

$$F(|\varphi\rangle, |\psi\rangle) = |\langle\varphi|\psi\rangle|.$$

More generally we can define the fidelity of a pure state, $|\varphi\rangle$, and a density matrix, $\hat{\rho}$, as

$$F(|\varphi\rangle, \hat{\rho}) = \sqrt{\langle\varphi|\hat{\rho}|\varphi\rangle}.$$

If $\hat{\rho}$ represents the pure state $|\psi\rangle$ then this reduces to the case of two pure states.

The fidelity can be seen as a measure of how close the information is to the information intended. For example with our $|+\rangle$ and $|y\rangle$ density matrix the eigenvalues are

$$|+'\rangle = \cos\frac{\pi}{8}|+\rangle - i\sin\frac{\pi}{8}|-\rangle,$$
$$|-'\rangle = \sin\frac{\pi}{8}|+\rangle + i\cos\frac{\pi}{8}|-\rangle.$$

The corresponding eigenvalues are $\cos^2(\pi/8)$ and $\sin^2(\pi/2)$ respectively. The fidelities between the primed and un-primed states are

$$F(|+\rangle, |+'\rangle) = F(|y\rangle, |+'\rangle) = \cos\frac{\pi}{8} = 0.924,$$
$$F(|+\rangle, |-'\rangle) = F(|y\rangle, |-'\rangle) = \sin\frac{\pi}{8} = 0.0.383.$$

Suppose we don't know which of the two un-primed states was sent. If we simply guess $|+'\rangle$ then this still has considerable overlap with the actual state, whether it is $|+\rangle$ or $|y\rangle$. This means that $|+'\rangle$ may be a good enough estimate of the state if we don't need to be 100% accurate. Classically we cannot make this sort of approximation. The symbol sent is either correct or incorrect. There is no overlap.

34.2.4   Entanglement Entropy

We can use the entropy to measure how much a system is entangled. Recall that a two particle system is entangled if it is in a state which *cannot* be written as a tensor product of state vectors of the initial single particle Hilbert spaces. For example if the state is $|+-\rangle$ then we can write this as $|+\rangle \otimes |-\rangle$ so the state is not entangled. Any two particle state that *cannot* be written as

$$(a_-|-\rangle + a_+|+\rangle) \otimes (b_-|-\rangle + b_+|+\rangle)$$

for $a_-, a_+, b_-, b_+ \in \mathbb{C}$ is entangled.

Suppose our system can be split into two subsystems, A and B. These may be individual particles but they could also be more complicated systems. The Hilbert space describing the system is then $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ where $\mathcal{H}_A$ and $\mathcal{H}_B$ are the Hilbert spaces describing subsystems A and B respectively. Let $|\Psi\rangle \in \mathcal{H}$ be a pure state of the combined system. We wish to know if this state is entangled between the two subsystems, and if so to what degree. Note that the existence of the two subsystems is critical here. It makes no sense to ask how entangled a given state is, only how entangled a subsystem is when it is in this state. As mentioned at the start of the section this is measured by the entropy:

$$S = -\operatorname{Tr}[\hat{\rho}_A \log_2 \hat{\rho}_A]$$

which we call the **entanglement entropy of subsystem A**. Here $\hat{\rho}_A$ is the reduced density matrix of subsystem A which is defined by tracing out the states of subsystem B:

$$\hat{\rho}_A = \operatorname{Tr}_B[\hat{\rho}_{AB}]$$

where $\hat{\rho}_{AB}$ is the density matrix of the entire system, for a pure state $\hat{\rho}_{AB} = |\Psi\rangle\langle\Psi|$. Also $\operatorname{Tr}_B[\hat{\rho}_{AB}]$ is the **partial trace** with respect to subsystem B. It is defined as

$$\operatorname{Tr}_B(\hat{P} \otimes \hat{Q}) = \operatorname{Tr}\{\}[\hat{Q}]\hat{P}$$

where $\hat{P}$ is some operator on $\mathcal{H}_A$ and $\hat{Q}$ is some operator on $\mathcal{H}_B$ and Tr is the normal trace defined on a single operator.

Example 34.1. Suppose that $|\Psi\rangle = |\Psi_A\rangle \otimes |\Psi_B\rangle$, i.e. the system is *not* entangled. Then

$$\hat{\rho}_{AB} = |\Psi\rangle\langle\Psi| = (|\Psi_A\rangle \otimes |\Psi_B\rangle)(\langle\Psi_A| \otimes \langle\Psi_B|) = |\Psi_A\rangle\langle\Psi_A| \otimes |\Psi_B\rangle\langle\Psi_B|.$$

Hence

$$\hat{\rho}_A = \operatorname{Tr}_B[|\Psi\rangle\langle\Psi|] = \operatorname{Tr}[|\Psi_B\rangle\langle\Psi_B|]|\Psi_A\rangle\langle\Psi_A| = |\Psi_A\rangle\langle\Psi_A|.$$

Since $|\Psi_B\rangle\langle\Psi_B|$ has a matrix representation with zeros everywhere apart from the diagonal entry corresponding to the state $|\Psi_B\rangle$ where it is one and so it has unit trace. The entanglement entropy is then

$$S = -\operatorname{Tr}[\hat{\rho}_A \log_2 \hat{\rho}_A] = 0.$$

As with $|\Psi_B\rangle\langle\Psi_B|$ the entries in $\hat{\rho}_A$ are either one or zero[19]. Since $\log_2 1 = 0$ this means that $\log_2 \hat{\rho}_A$ is the zero matrix and so the entanglement entropy is zero, as it should be for an state that isn't entangled.

Example 34.2. Consider a two state system with two spin states, $|+\rangle$ and $|-\rangle$. Suppose the system is in the pure state[20]

$$|\Psi\rangle = \frac{\sqrt{2}}{2}(|+-\rangle \pm |-+\rangle).$$

The density matrix for the entire system is $\hat{\rho}_{AB} = |\Psi\rangle\langle\Psi|$. We can write this as a matrix in a basis where the rows and columns correspond to $|++\rangle$, $|-+\rangle$, $|+-\rangle$, and $|--\rangle$, in that order. This is the same four-dimensional basis we have used when talking about quantum computing. In this basis

$$\hat{\rho}_{AB} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/2 & \pm 1/2 & 0 \\ 0 & \pm 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

---

[19]while $\log_2 0$ is undefined we are multiplying by zero and so seemingly this doesn't matter.
[20]we use $\pm$ for a general result but this is not a mixed state, the state is $+$ or $-$ and we choose which.

The density matrix for particle A is then

$$\begin{aligned}
\hat{\rho}_A &= \mathrm{Tr}_B[\hat{\rho}_{AB}] \\
&= \langle +B|\hat{\rho}_{AB}|+_B\rangle + \langle -_B|\hat{\rho}_{AB}|-_B\rangle \\
&= \frac{1}{2}(|-_A\rangle\langle -_A| + |+_A\rangle\langle +_A|).
\end{aligned}$$

As a matrix with rows and columns corresponding to $|+_A\rangle$ and $|-_A\rangle$ in that order this can be written as

$$\hat{\rho}_A = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$$

Hence

$$S = -\mathrm{Tr}[\hat{\rho}_A] = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = 1.$$

We say that the system is in a maximally entangled state as 1 is the highest entropy a two particle state can have.

34.2.5   Purification

The notion of **purification** is that for a given density matrix we can find a Hilbert space such that the density matrix represents a pure state in that Hilbert space. That is given a system A with density matrix $\hat{\rho}_A$ we can introduce another system, R, and find a pure state, $|AR\rangle$, in the joint system such that $\hat{\rho}_A = \mathrm{Tr}_R[|AR\rangle\langle AR|]$.

To show this is true recall that $\hat{\rho}_A$ is Hermitian and therefore can be diagonalised. Let $\{|i_A\rangle\}$ be a basis such that $\hat{\rho}_A$ is diagonal. Then

$$\hat{\rho}_A = \sum_i p_i|i_A\rangle\langle i_A|$$

where p_i is the classical probability that the system is in the state $|i_A\rangle$. Suppose that $\{|i_R\rangle\}$ is an orthonormal basis for system R. Let

$$|AR\rangle = \sum_i \sqrt{p_i}|i_A i_R\rangle.$$

Then the reduced density matrix for the system, which we get by taking the partial trace with respect to R, is

$$\begin{aligned}
\mathrm{Tr}_R[|AR\rangle\langle AR|] &= \sum_{ij}\sqrt{p_i p_j}|i_A\rangle\langle j_A|\,\mathrm{Tr}\{\}[|i_R\rangle\langle j_R|] \\
&= \sum_{ij}\sqrt{p_i p_j}|i_A\rangle\langle j_A|\delta_{ij} \\
&= \sum_i p_i|i_A\rangle\langle i_A| \\
&= \hat{\rho}_A.
\end{aligned}$$

The state $|AR\rangle$ is called the purification of $\hat{\rho}_A$.

34.2.6   Klein's Inequality

Let $\hat{\rho}$ and $\hat{\sigma}$ be density matrices which are diagonal in the bases $\{|i\rangle\}$ and $\{|j\rangle\}$ respectively. Then

$$\hat{\rho} = \sum_i p_i|i\rangle\langle i|, \qquad \text{and} \qquad \hat{\sigma} = \sum_j q_j|j\rangle\langle j|.$$

The **relative entropy** of $\hat{\rho}$ with respect to $\hat{\sigma}$ is a measure of how close $\hat{\rho}$ and $\hat{\sigma}$ are. It is defined as

$$S_{\mathrm{vN}}(\hat{\rho}\|\hat{\sigma}) = \mathrm{Tr}[\hat{\rho}\log_2\hat{\rho} - \hat{\rho}\log_2\hat{\sigma}] = \sum_i p_i\log_2 p_i - \sum_i\langle i|\hat{\rho}\log_2\hat{\sigma}|i\rangle.$$

The Klein inequality simply states that this is non-negative, that is

$$S_{\text{vN}}(\hat{\rho}\|\hat{\sigma}) \geq 0,$$

or equivalently,

$$S_{\text{vN}}(\hat{\rho}) \leq -\operatorname{Tr}[\hat{\rho}\log_2 \hat{\sigma}].$$

To see why this should be true first note that

$$\langle i|\log_2 \hat{\sigma}|i\rangle = \langle i|\sum_j \log_2 q_j |j\rangle\langle j|i\rangle = \sum_j P_{ij}\log_2 q_j$$

where $P_{ij} = \langle i|j\rangle\langle j|i\rangle = |\langle i|j\rangle|^2 \geq 0$. For complete, orthogonal basis sets P_{ij} satisfies

$$\begin{aligned}
\sum_i P_{ij} &= \sum_i \langle i|j\rangle\langle j|i\rangle \\
&= \sum_i \langle j|i\rangle\langle i|j\rangle \\
&= \langle j|j\rangle \\
&= 1,
\end{aligned}$$

and similarly $\sum_j P_{ij} = 1$. Substituting this into the definition of the relative entropy we have

$$S_{\text{vN}}(\hat{\rho}\|\hat{\sigma}) = \sum_i p_i \left[p_i \log_2 p_i - \sum_j P_{ij}\log_2 q_j \right].$$

It can be shown that $\log_2$ is a concave function and therefore satisfies

$$\log_2[(1-a)x + ay] \geq (1-a)\log_2 x + a\log_2 y$$

for $a \in [0,1]$. Hence

$$\sum_j P_{ij}\log_2 q_j \leq \log_2 r_i < 0$$

where

$$r_i = \sum_j P_{ij}q_j.$$

Equality occurs if for some value of j we have $P_{ij} = 1$. In this case all other $P_{ij} = 0$. Using this we have

$$S_{\text{vN}}(\hat{\rho}\|\hat{\sigma}) \geq \sum_i p_i \log_2 \frac{p_i}{r_i} = -\sum_i p_i \log_2 \frac{r_i}{p_i}.$$

Now we use the following identity for transforming between bases:

$$-\log_2(x)\ln 2 = -\ln x \geq 1 - x$$

which is valid for $x > 0$ with equality only when $x = 1$. Finally

$$S_{\text{vN}}(\hat{\rho}\|\hat{\sigma}) \geq -\frac{1}{\ln 2}\sum_i p_i \left[1 - \frac{r_i}{p_i} \right] = \frac{1}{\ln 2}\sum_i (r_i - p_i) = 0.$$

34.2.7   Mathematical Properties of von Neumann Entropy

The von Neumann entropy has many nice properties that aid in calculation or provide some insight. We list some here:

1. *Invariance under unitary transformation of the density matrix.*

   To show this we first need a result about logarithms of matrices. The log of a matrix is defined to be the function such that if $\exp(A) = B$ then $A = \log(B)$, i.e. the inverse of exponentiation. If B is Hermitian we can diagonalise it; meaning that we can find a diagonal matrix D such that

$UDU^\dagger = B$ for some unitary transformation U. Hence $\exp(A) = UDU^\dagger$ and $A = \log(UDU^\dagger)$. Recall that $\exp(A)$ is defined through a power series,

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

Hence

$$U^\dagger \exp(A) U = U^\dagger \left[\sum_{n=0}^{\infty} \frac{A^n}{n!} \right] U$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} U^\dagger A^n U.$$

Now we use

$$(U^\dagger A U)^n = \underbrace{U^\dagger A U U^\dagger A U \cdots U^\dagger A U}_{n \text{ times}} = U^\dagger \underbrace{A \mathbb{1} A \mathbb{1} \cdots \mathbb{1} A}_{n \text{ times}} U = U^\dagger A^n U$$

and so

$$U^\dagger \exp(A) U = \sum_{n=0}^{\infty} \frac{1}{n!} U^\dagger A^n U = \sum_{n=0}^{\infty} \frac{1}{n!} (U^\dagger A U)^n = \exp(U A U^\dagger).$$

Hence

$$U^\dagger \exp(A) U = \exp(U^\dagger A U) = D \implies U^\dagger A U = \log(D)$$

and so

$$U U^\dagger A U U^\dagger = A = U \log(D) U^\dagger = \log(B) = \log(UDU^\dagger).$$

Now let $\hat{U}$ be an arbitrary unitary operator. Then

$$S_{\text{vN}}(\hat{U} \hat{\rho} \hat{U}^\dagger) = - \operatorname{Tr}\left[\hat{U} \hat{\rho} \hat{U}^\dagger \log_2(\hat{U} \hat{\rho}^\dagger) \right]$$

$$= - \operatorname{Tr}\left[\hat{U} \hat{\rho} \hat{U}^\dagger \hat{U} \log_2(\hat{\rho})^{\hat{\dagger}} \right]$$

$$= - \operatorname{Tr}\left[\hat{U} \hat{\rho} \log_2(\hat{\rho})^{\hat{\dagger}} \right]$$

$$= - \operatorname{Tr}\left[\hat{U}^\dagger \hat{U} \hat{\rho} \log_2 \hat{\rho} \right]$$

$$= - \operatorname{Tr}[\hat{\rho} \log_2 \hat{\rho}]$$

$$= S_{\text{vN}}(\hat{\rho}).$$

Here we have also used the cyclic property of the trace:

$$\operatorname{Tr}[ABC] = \operatorname{Tr}[CAB].$$

2. *Positivity:* $S_{\text{vN}}(\hat{\rho}) \geq 0$.

   Since $\hat{\rho}$ is Hermitian it can be diagonalised. Recall from section 34.2.1 that the eigenvalues, α_i, of $\hat{\rho}$ all lie in the interval $[0, 1]$. Hence $\log_2 \alpha_i \leq 0$ with equality only when $\alpha_i = 1$. In the eigenbasis of $\hat{\rho}$ the trace is simply the sum over the eigenvalues and hence

   $$S_{\text{vN}}(\hat{\rho}) = - \sum_i \alpha_i \log_2 \alpha_i \geq 0$$

   with equality only when $\alpha_i = 1$ which means we are in a pure state.

3. *Maximum:* $S_{\text{vN}} \leq S_{\text{vN}}^{\max} = \log_2 D$ *where D is the number of non-zero eigenvalues of the density matrix.*

   In the eigenbasis of $\hat{\rho}$ the entropy is

   $$- \sum_i \alpha_i \log_2 \alpha_i.$$

We know from section 34.2.1 that

$$\sum_i \alpha_i = 1.$$

It can then be shown that $S_{\text{vN}}(\hat{\rho})$ is maximised when all eigenvalues have the same value. If D is the dimensionality of $\hat{\rho}$ then this occurs when $\alpha_i = 1/S$ and hence

$$S_{\text{vN}}(\hat{\rho}) = -\sum_{i=1}^{D} \alpha_i \log_2 \alpha_i = \log_2 D.$$

4. *Subadditivity: $S_{\text{vN}}(\hat{\rho}_{AB}) \leq S_{\text{vN}}(\hat{\rho}_A) + S_{\text{vN}}(\hat{\rho}_B)$.*

   That is the von Neumann entropy of the combined system of A and B is always at most the von Neumann entropy of the two individual systems. We will show this for the simplest case when $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ where $\mathcal{H}_{AB}$ is the Hilbert space for the whole system and $\mathcal{H}_A$ and $\mathcal{H}_B$ are the Hilbert spaces for the two subsystems. In this case $\hat{\rho}_{AB} = \hat{\rho}_A \otimes \hat{\rho}_B$. Let $\hat{U}_i$ be the unitary transformation that diagonalises $\hat{\rho}_i$. Then $\hat{U}_A \otimes \hat{U}_B$ diagonalises $\hat{\rho}_A \otimes \hat{\rho}_B$. In this diagonal basis

$$S_{\text{vN}}(\hat{\rho}_{AB}) = -\sum_{ij} p_i^{(A)} p_j^{(B)} \log_2(p_i^{(A)} p_j^{(B)})$$
$$= -\sum_{ij} p_i^{(A)} p_j^{(B)} [\log_2 p_i^{(A)} + \log_2 p_j^{(B)}]$$
$$= -\sum_{ij} p_i^{(A)} p_j^{(B)} \log_2 p_i^{(A)} - \sum_{ij} p_i^{(A)} p_j^{(B)} \log_2 p_j^{(B)}$$
$$= -\sum_i p_i^{(A)} \log_2 p_i^{(A)} - \sum_{ij} p_j^{(B)} \log_2 p_j^{(B)}$$

   where we have used the fact that $p_i^{(j)}$ are probabilities so sum to 1. In the case where $\mathcal{H}_A$ and $\mathcal{H}_B$ are more correlated then this can only decrease the number of states available and hence the entropy. For example if there is a potential between the two particles then the particles will be limited to certain states and won't be able to access states that they could when they were free. Hence if $\mathcal{H}_{AB} \subset \mathcal{H}_A \otimes \mathcal{H}_B$ then we find that the joint entropy is less than the sum of the individual entropies.

   We can prove this more generally using Klein's inequality (section 34.2.6). Let $\hat{\rho} = \hat{\rho}_{AB}$ and $\hat{\sigma} = \hat{\rho}_A \otimes \hat{\rho}_B$. Then by Klein's inequality

$$S_{\text{vN}}(\hat{\rho}_{AB}) \leq -\text{Tr}[\hat{\rho} \log_2 \hat{\sigma}]$$
$$= -\text{Tr}[\hat{\rho}_{AB}(\log_2 \hat{\rho}_A + \log_2 \hat{\rho}_B)]$$
$$= -\text{Tr}[\hat{\rho}_A \log_2 \hat{\rho}_A] - \text{Tr}[\hat{\rho}_B \log_2 \hat{\rho}_B]$$
$$= S_{\text{vN}}(\hat{\rho}_A) + S_{\text{vN}}(\hat{\rho}_B).$$

   Equality holds when $\hat{\rho}_{AB} = \hat{\rho}_A \otimes \hat{\rho}_B$.

5. *Concavity: For $\beta_i \geq 0$ such that $\sum_i \beta_{i=1}^n = 1$ we have*

$$S_{\text{vN}}\left(\sum_{i=1}^n \beta_i \hat{\rho}_i\right) = S_{\text{vN}}(\beta_1 \hat{\rho}_1 + \beta_2 \hat{\rho}_2 + \cdots + \beta_n \hat{\rho}_n)$$

$$\geq \beta_1 S_{\text{vN}}(\hat{\rho}_1) + \beta_2 S_{\text{vN}}(\hat{\rho}_2) + \cdots + \beta_n S_{\text{vN}}(\hat{\rho}_n) = \sum_{i=1}^n \beta_i S_{\text{vN}}(\hat{\rho}_i).$$

   Sums of density matrices occur when we have an ensemble of ensembles. Intuitively the entropy is larger for the ensemble of density matrices as we know less about the system. On the other hand we may know a lot about each individual system and so the sum of entropies is lower.

   To show this mathematically let $\hat{\rho}_A = \sum_j \beta_j \hat{\rho}_j$ and $\hat{\rho}_B = \sum_j \beta_j |j\rangle\langle j|$ for some orthonormal basis, $\{|j\rangle\}$. The joint state is then $\hat{\rho}_{AB} = \sum_j \beta_j \hat{\rho}_j \otimes |j\rangle\langle j|$. Notice then that

$$S_{\text{vN}}(\hat{\rho}_{AB}) = S_{\text{Sh}}(\{\beta_j\}) + \sum_j \beta_j S_{\text{vN}}(\hat{\rho}_j)$$

and concavity then follows from subadditivity.

6. *Triangle inequality:* $S_{\text{vN}}(\hat{\rho}_{AB}) \geq |S_{\text{vN}}(\hat{\rho}_A|) - S_{\text{vN}}(\hat{\rho}_B)$.

In the case $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ equality holds which follows from the discussion in the section on subadditivity. In general let R be a system that purifies A and B. Recall that this means that we create a pure state $|AR\rangle$ such that $\hat{\rho}_A = \text{Tr}_R[|AR\rangle\langle AR|]$. This can be done for any system. By construction the system ABR is in a pure state which means that

$$S_{\text{vN}}(\hat{\rho}_{AR}) = S_{\text{vN}}(\hat{\rho}_B)$$

and

$$S_{\text{vN}}(\hat{\rho}_R) = S_{\text{vN}}(\hat{\rho}_{AB}).$$

Applying subadditivity to systems A and R we have

$$S_{\text{vN}}(\hat{\rho}_R) + S_{\text{vN}}(\hat{\rho}_A) \geq S_{\text{vN}}(\hat{\rho}_{AR}).$$

Combining these statements we have

$$S_{\text{vN}}(\hat{\rho}_{AB}) \geq S_{\text{vN}}(\hat{\rho}_B) - S_{\text{vN}}(\hat{\rho}_A).$$

Since all of this is symmetric in A and B we also have

$$S_{\text{vN}}(\hat{\rho}_{AB}) \geq S_{\text{vN}}(\hat{\rho}_A) - S_{\text{vN}}(\hat{\rho}_B).$$

These can then be combined into the triangle inequality.

7. *Strong subadditivity:* $S_{\text{vN}}(\hat{\rho}_{ABC}) + S_{\text{vN}}(\hat{\rho}_B) \leq S_{\text{vN}}(\hat{\rho}_{AB}) + S_{\text{vN}}(\hat{\rho}_{BC})$.

The proof of is not given.

34.3   Other Measures of Information

Suppose we have two classical languages, A and B. For a language $J = A, B$ denote by $p_i^{(J)}$ the probability of the ith symbol in the language. The Shannon entropy for each language is given by

$$S_{\text{Sh}}(J) = -\sum_i p_i^{(J)} \log_2 p_i^{(J)}.$$

We can imagine joining both languages together to get a new language AB with entropy

$$S_{\text{Sh}}(A, B) = -\sum p_i^{(AB)} \log_2 p_i^{(AB)}$$

where $p_i^{(AB)}$ is the probability of the ith symbol in the joint language. If A and B are completely independent then i would simply index over all possible product states, $j_A k_B$, of the two sub-languages. However if the two languages are not independent then this is not the case. We define the **mutual information** of the joint system to be

$$I_{\text{Sh}}(A, B) = S_{\text{Sh}}(A) + S_{\text{Sh}}(B) - S_{\text{Sh}}(A, B).$$

This measures the contribution to the information of both systems and then removes the excess information that is over counted since it is provided by both systems[21]. For completely independent systems $I_{\text{Sh}} = 0$.

Another measure that we could use is the **conditional entropy** of a system which is defined as

$$S_{\text{Sh}}(A|B) = \sum_{i_B} p_{i_B}^{(B)} S_{\text{Sh}}(A|B = i_B)$$

where

$$S_{\text{Sh}}(A|B = i_B) = -\sum_{i_A} p(A_{i_A}|B_{i_B}) \log_2 p(A_{i_A}|B_{i_B}).$$

---

[21]cf. $|A \cup B| = |A| + |B| - |A \cap B|$ and $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$.

Here $p(X|Y)$ has the usual meaning in probability of 'probability of X given that Y is true'. The conditional entropy of A as defined above measures the information content of A when B is in all possible states. In classical probability theory $S_{\text{Sh}}(A|B) = S_{\text{Sh}}(A, B) - S_{\text{Sh}}(B)$. This allows for an equivalent definition of the mutual information in classical probability theory:

$$J_{\text{Sh}}(A : B) = S_{\text{Sh}}(A) - S_{\text{Sh}}(A|B).$$

In quantum information theory probability distributions are replaced with density matrices and Shannon entropies are replaced with von Neumann entropies. The joint entropy is then $S_{\text{vN}}(\hat{\rho}_{AB})$ where $\hat{\rho}_{AB}$ is the density matrix for the combined system of A and B. The **quantum mutual information** then has an analogous definition to its classical counterpart:

$$I_{\text{vN}}(A, B) = S_{\text{vN}}(\hat{\rho}_A) + S_{\text{vN}}(\hat{\rho}_B) - S_{\text{vN}}(\hat{\rho}_{AB}).$$

If $I_{\text{vN}}(A, B) \neq 0$ then we say that A and B are **correlated**.

Knowing the information necessarily requires measurements to be made and therefore in quantum information theory we have to consider the collapse of the wave function. To find the conditional entropy in quantum mechanics we need a well defined notion of a measurement. The most commonly used are called the von Neumann measurements or local projective measurements. Suppose that we measure subsystem B and find it in the state $|j_B\rangle$. The probability of this is $p_{j_B} = \text{Tr}\big[\hat{\rho}_{AB|j_B}\big]$ where

$$\hat{\rho}_{AB|j_B} = \frac{1}{p_{j_B}}(\mathbb{1}_A \otimes \hat{\Pi}_{j_B})\hat{\rho}_{AB}(\mathbb{1}_A \otimes \hat{\Pi}_{j_B})$$

where

$$\hat{\Pi}_{j_B} = |j_B\rangle\langle j_B|.$$

The conditional entropy for A when B is measured in all possible states is then

$$S_{\text{vN}}(\hat{\rho}_{AB}|\{\hat{\Pi}_{j_B}\}) = \sum_{j_B} p_{j_B} S_{\text{vN}}(\hat{\rho}_{A|j_B})$$

where

$$\hat{\rho}_{A|j_B} = \text{Tr}_B[\hat{\rho}_{AB|j_B}].$$

Clearly this is dependent on which set of projection operators, $\{\hat{\Pi}_{j_B}\}$, we choose. This means that what we measure B as will effect the conditional entropy of A.

We can now define the other form of mutual information:

$$J_{\text{vN}}(\hat{\rho}_{AB}) = S_{\text{vN}}(\hat{\rho}_A) - S_{\text{vN}}(\hat{\rho}_{AB}|\{\hat{\Pi}_{j_B}\}).$$

The **quantum discord** is defined as

$$D_{AB}(\hat{\rho}_{AB}) = I_{\text{vN}}(\rho_{\hat{A}B}) - \max_{\{\hat{\Pi}_{j_B}\}} [J_{\text{vN}}(\hat{\rho}_{AB})].$$

That is it measures the difference between the two types of mutual information when $\{\hat{\Pi}_{j_B}\}$ are chosen to maximise J_{vN}. Classically both measures of information are equivalent and the quantum discord is zero. This is no longer true in the quantum case. Roughly we can think of the quantum discord as a measure of how 'quantum' a system is. Another way to think of it is as the purely quantum correction to the definition of the mutual information as I_{vN} accounts for both quantum and classical effects whereas J_{vN} accounts only for the classical effects and so there difference is simply the quantum effect.

Much discussion about the philosophy of quantum mechanics centres around these measures of information and what we interpret them to mean. However we won't discuss them further here.

Appendix

A $\quad$ Algebraic Structures

In section 1.2.1 we discussed vector spaces. We defined a vector space, $\mathcal{H}$, over a field, $\mathbb{F}$. But what is a field? Intuitively a field is a set with addition, multiplication, subtraction, and division defined as we would expect. To be more rigorous we need a couple of definitions first.

A.1 $\quad$ Groups

A **group**, $(G, \circ)$, is a set, G, along with a binary operation

$$\circ\colon G \times G \to G$$
$$x, y \mapsto x \circ y = z \in G.$$

This binary operation satisfies the following:

- Associativity – For all $x, y, z \in G$ we have

$$x \circ (y \circ z) = (x \circ y) \circ z.$$

- Identity – There exists $e \in G$ such that for all $x \in G$ we have

$$e \circ x = x \circ e = x.$$

  e is called the identity.

- Inverse – For all $x \in G$ there exists $x^{-1} \in G$ such that

$$x \circ x^{-1} = x^{-1} \circ x = e$$

  where e is the identity.

An example of a group is the group of all permutations, $\sigma\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ such that σ is a bijection. Here the group operation is function composition.

If $(G, \circ)$ is a group and for all $x, y \in G$ we have

$$x \circ y = y \circ x$$

then we say $(G, \circ)$ is an **abelian group**. An example of an abelian group is $\mathbb{Z}_2 = \{0, 1\}$ under addition modulo 2.

A.2 $\quad$ Rings

A **ring**, $(R, +, \cdot)$, is a set, R, along with two binary operations:

$$+\colon R \times R \to R$$
$$x, y \mapsto x + y = z \in R,$$

which we call addition, and

$$\cdot\colon R \times R \to R$$
$$x, y \mapsto xy = z' \in R,$$

which we call multiplication. These two operations satisfy the following:

- $(R, +)$ is an abelian group.
- Associativity of multiplication – For all $x, y, z \in R$

$$x(yz) = (xy)z.$$

- Distributivity of multiplication over addition – For all $x, y, z \in R$

$$x(y + z) = xy + xz, \qquad \text{and} \qquad (x + y)z = xz + yz.$$

We denote the identity of $(R, +)$ by 0 and inverses of addition as $-x$ (as opposed to x^{-1}). An example of a ring is $M_2(2\mathbb{Z})$, which is the set of 2×2 matrices with entries from $2\mathbb{Z} = \{0, \pm 2, \pm 4, \pm 6, \dots\}$. Here the operations are matrix addition and matrix multiplication.

If R is a ring (we drop the $+$ and $\cdot$ for brevity) and there exists $1 \in R$ such that $1x = x1 = x$ for all $x \in R$ then R is a **ring with unity**. An example of a ring with unity is $M_2(\mathbb{Z})$ which is the set of all 2×2 matrices with entries from $\mathbb{Z}$, in this case the identity matrix plays the role of the multiplicative identity.

If R is a ring with unity and for all $x \in R$ such that $x \neq 0$ there exists $x^{-1} \in R$ such that $xx^{-1} = x^{-1}x = 1$ then R is a **division ring**. An example of a ring with unity is the quaternions, $\mathbb{H}$, with quaternion addition and multiplication.

If R is a ring and for all $x, y \in R$ we have $xy = yx$ then R is a **commutative ring**. An example of a commutative ring is $2\mathbb{Z}$ with normal integer addition and multiplication.

If R is a commutative ring and for all $a, b \in R$ we have $ab = 0$ if and only if $a = 0$ and/or $b = 0$ then R is an **integral domain**. An example of an integral domain is $\mathbb{Z}$ with normal integer addition and multiplication.

If R is a commutative division ring, and $1 \neq 0$, then it is a **field**. Examples of fields includes $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$, with their respective addition and multiplication operations. As well as these $\mathbb{Z}_p = \{0, \dots, p - 1\}$ for prime p under addition and multiplication modulo p is a field.

B   Sum of k from 0 to n proof

Theorem 10

For $n \in \mathbb{N}$

$$\sum_{r=0}^{n} r = \frac{1}{2}n(n + 1).$$

Proof. We will prove this by induction. First note that

$$\sum_{r=0}^{0} r$$

is the empty sum which is defined to be zero by convention. Now suppose that this result holds for some $k \in \mathbb{N}$, that is

$$\sum_{r=0}^{k} r = \frac{1}{2}k(k + 1).$$

Then

$$\begin{aligned}
\sum_{r=0}^{k+1} r &= \sum_{r=0}^{k} r + (k + 1) \\
&= \frac{1}{2}k(k + 1) + (k + 1) \\
&= (k + 1)\left[\frac{1}{2}k + 1\right] \\
&= \frac{1}{2}(k + 1)(k + 2)
\end{aligned}$$

m_1, m_2 \ j	3/2
1, 1/2	1

(a) $m = 3/2$

m_1, m_2 \ j	3/2	1/2
1, 1/2	$\sqrt{1/3}$	$\sqrt{2/3}$
0, 1/2	$\sqrt{2/3}$	$-\sqrt{1/3}$

(b) $m = 1/2$

Table C.2: Clebsch–Gordan coefficients for $j_1 = 1$ and $j_2 = 1/2$

m_1, m_2 \ j	2
1, 1	1

(a) $m = 2$

m_1, m_2 \ j	2	1
1, 0	$\sqrt{1/2}$	$\sqrt{1/2}$
0, 1	$\sqrt{1/2}$	$-\sqrt{1/2}$

(b) $m = 1$

m_1, m_2 \ j	2	1	0
1, -1	$\sqrt{1/6}$	$\sqrt{1/2}$	$\sqrt{1/3}$
0, 0	$\sqrt{2/3}$	0	$-\sqrt{1/3}$
-1, 1	$\sqrt{1/6}$	$-\sqrt{1/2}$	$\sqrt{1/3}$

(c) $m = 0$

Table C.3: Clebsch–Gordan coefficients for $j_1 = 1$ and $j_2 = 1$

m_1, m_2 \ j	2
3/2, 1/2	1

(a) $m = 2$

m_1, m_2 \ j	2	1
3/2, -1/2	1/2	$\sqrt{3/4}$
1/2, 1/2	$\sqrt{3/4}$	$-1/2$

(b) $m = 1$

m_1, m_2 \ j	2	1
1/2, -1/2	$\sqrt{1/2}$	$\sqrt{1/2}$
-1/2, 1/2	$\sqrt{1/2}$	$-\sqrt{1/2}$

(c) $m = 0$

Table C.4: Clebsch–Gordan coefficients for $j_1 = 3/2$ and $j_2 = 1/2$

m_1, m_2 \ j	5/2
3/2, 1	1

(a) $m = 5/2$

m_1, m_2 \ j	5/2	3/2
3/2, 0	$\sqrt{2/5}$	$\sqrt{3/5}$
1/2, 1	$\sqrt{3/5}$	$-\sqrt{2/5}$

(b) $m = 3/2$

m_1, m_2 \ j	5/2	3/2	1/2
3/2, -1	$\sqrt{1/10}$	$\sqrt{2/5}$	$\sqrt{1/2}$
1/2, 0	$\sqrt{3/5}$	$\sqrt{1/15}$	$-\sqrt{1/3}$
-1/2, 1	$\sqrt{3/10}$	$-\sqrt{8/15}$	$\sqrt{1/6}$

(c) $m = 0$

Table C.5: Clebsch–Gordan coefficients for $j_1 = 3/2$ and $j_2 = 1$

which is exactly what we would expect if the hypothesis held for $k + 1$. Hence by induction

$$\sum_{r=0}^{n} r = \frac{1}{2}n(n+1)$$

D   Deutsch Algorithm

I wrote a Mathematica note book to apply the Deutsch algorithm.

Define the operators:

id = DiagonalMatrix[{1, 1}]

$\{\{1, 0\}, \{0, 1\}\}$

X = PauliMatrix[1]

$\{\{0, 1\}, \{1, 0\}\}$

CNOT = {{1, 0, 0, 0}, {0, 1, 0, 0}, {0, 0, 0, 1}, {0, 0, 1, 0}}

$\{\{1, 0, 0, 0\}, \{0, 1, 0, 0\}, \{0, 0, 0, 1\}, \{0, 0, 1, 0\}\}$

Hd = $\frac{1}{\text{Sqrt[2]}}$ $*$ {{1, 1}, {1, −1}}

$\left\{ \left\{ \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\}, \left\{ \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right\} \right\}$

Of1 = DiagonalMatrix[{1, 1, 1, 1}]

$\{\{1, 0, 0, 0\}, \{0, 1, 0, 0\}, \{0, 0, 1, 0\}, \{0, 0, 0, 1\}\}$

Of2 = CNOT

$\{\{1, 0, 0, 0\}, \{0, 1, 0, 0\}, \{0, 0, 0, 1\}, \{0, 0, 1, 0\}\}$

Of3 = CNOT.KroneckerProduct[id, X]

$\{\{0, 1, 0, 0\}, \{1, 0, 0, 0\}, \{0, 0, 1, 0\}, \{0, 0, 0, 1\}\}$

Of4 = KroneckerProduct[id, X]

$\{\{0, 1, 0, 0\}, \{1, 0, 0, 0\}, \{0, 0, 0, 1\}, \{0, 0, 1, 0\}\}$

operators = {{id, "id"}, {X, X}, {CNOT, "CNOT"}, {Hd, "Hd"},

{Of1, "Of1"}, {Of2, "Of2"}, {Of3, "Of3"}, {Of4, "Of4"}};

Do[Print[operators[[i]][[2]], " = ", MatrixForm[operators[[i]][[1]]]], {i, 8}]

$$\text{id} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\text{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\text{Hd} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\text{Of1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\text{Of2} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\text{Of3} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\text{Of4} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

The first step of the algorithm is to apply $(\hat{H}_d \otimes \hat{H}_d)(\hat{X} \otimes \hat{X})$ to the state $|0\rangle|0\rangle$, so define this operator and this state and compute the product:

initOp = KroneckerProduct[Hd, Hd].KroneckerProduct[X, X]

$$\left\{ \left\{ \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2} \right\}, \left\{ -\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2} \right\}, \left\{ -\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2} \right\}, \left\{ \tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2} \right\} \right\}$$

initState = {1, 0, 0, 0}

$\{1, 0, 0, 0\}$

step1 = initOp.initState

$\left\{ \tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2} \right\}$

Next we apply the oracle:

step2 = {Of1.step1, Of2.step1, Of3.step1, Of4.step1}

$$\left\{\left\{\tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right\}, \left\{\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}\right\}, \left\{-\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right\}, \left\{-\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}\right\}\right\}$$

Then we apply $(\hat{H}_d \otimes \hat{I})$:

finalOp = KroneckerProduct[Hd, id]

$$\left\{\left\{\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}, 0\right\}, \left\{0, \tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\right\}, \left\{\tfrac{1}{\sqrt{2}}, 0, -\tfrac{1}{\sqrt{2}}, 0\right\}, \left\{0, \tfrac{1}{\sqrt{2}}, 0, -\tfrac{1}{\sqrt{2}}\right\}\right\}$$

finalStates = {finalOp.step2[[1]], finalOp.step2[[2]], finalOp.step2[[3]], finalOp.step2[[4]]}

$$\left\{\left\{0, 0, \tfrac{1}{\sqrt{2}}, -\tfrac{1}{\sqrt{2}}\right\}, \left\{\tfrac{1}{\sqrt{2}}, -\tfrac{1}{\sqrt{2}}, 0, 0\right\}, \left\{-\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0, 0\right\}, \left\{0, 0, -\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}\right\}\right\}$$

Do[Print["Result with oracle ", i, " = ", MatrixForm[finalStates[[i]]]], {i, 4}]

$$\text{Result with oracle } 1 = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\text{Result with oracle } 2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{pmatrix}$$

$$\text{Result with oracle } 3 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{pmatrix}$$

$$\text{Result with oracle } 4 = \begin{pmatrix} 0 \\ 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

If we were to measure the register of these states we would get 0 for the second and third oracles (meaning $f(0) \neq f(1)$) and we would get 1 for the first and fourth oracles (meaning $f(0) = f(1)$).