

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 Debbio¹ and Professor Arjun Berera². 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 L^AT_EX³. Graphs were plotted using Python⁴, Matplotlib⁵, NumPy⁶, and SciPy⁷. Diagrams were drawn with tikz⁸.

This is version 1.0 of these notes, which is up to date as of 23/12/2020.

Willoughby Seago

s1824487@ed.ac.uk

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

Contents	ii
List of Figures	v
List of Tables	v
Acronyms	v
I States	1
1 Quantum States	1
1.1 Classical States	1
1.2 Quantum States	1
1.2.1 Vector Spaces	1
1.2.2 Scalar Products	3
1.2.3 Linear Functionals	3
1.2.4 Norms	4
1.2.5 Orthogonality	4
1.2.6 Operators	5
1.2.7 Bases	6
2 Two-State Systems	7
2.1 Calculations Within a Basis	7
2.1.1 Vectors in Two Bases	7
2.1.2 Scalar Products in a Basis	8
2.2 Operators in a Basis	8
2.3 Tensor Product	9
2.3.1 What is a Tensor Product?	9
2.3.2 Why are Tensor Products Useful?	9
2.3.3 Operators acting on Tensor Product Spaces	9
2.4 Two-State Systems	10
2.4.1 Two Two-State Systems	11
3 One-Dimensional System	13
3.1 Discretised System	13
3.1.1 Basis	13
3.1.2 Scalar Product and Norm	14
3.2 Continuous System	14
3.2.1 Basis	14
3.2.2 Scalar Product and Norm	15
3.2.3 Scalar Product of Basis Vectors	15
II Observables	16
4 Observables	16
4.1 Some Statistics	16
4.2 Position Operator	17
4.3 Generic Observables	18
4.3.1 Observing Observables	18
4.4 Eigenvalues and Eigenstates	18
4.5 Hermitian Operators	20
4.5.1 Hermitian Conjugate	20
4.5.2 Hermitian Operator	21
4.5.3 Properties of Hermitian Operators	21
4.5.4 Spectral Decomposition	23

5	Collapse of the State Vector	23
5.1	Compatible Observables	24
5.2	Complete Sets of Compatible observables	24
6	Continuous Spectra	25
6.1	Orthonormality	26
6.2	Momentum Operator	27
6.3	The Canonical Commutation Relation	28
III	Dynamics	28
7	Dynamics	29
7.1	Dynamics in Classical Mechanics	29
7.2	Dynamics in Quantum Mechanics	29
7.3	Time Evolution of a Wave Function	30
7.4	Justification of the Schrödinger Equation	30
7.5	Eigenstates of the Hamiltonian	31
7.5.1	Time Evolution of the Energy Eigenstates	31
7.5.2	Time Evolution of Matrix Elements	31
7.6	Time Evolution of a Generic State	32
7.7	The Eigenbasis of the Hamiltonian	33
7.8	Time Evolution of Expected Values	34
8	One-Dimensional Systems	35
8.1	Recap	35
8.2	Time Independent Schrödinger Equation	36
8.3	Free Particle	36
8.4	Time Dependent Schrödinger Equation	38
8.4.1	General Properties of a Wave Function	38
8.5	Physical Properties	38
8.5.1	Minimum Energy	38
8.5.2	Ehrenfest's Theorems	39
8.6	Degeneracy of the Hamiltonian	41
9	Potentials	42
9.1	Potential Step	42
9.1.1	The Potential	42
9.1.2	Case 1: $E > V$	43
9.1.3	Case 2: $E < V$	44
9.2	Potential Barrier	45
9.2.1	Tunnelling	46
9.2.2	Scanning Tunnelling Microscope	46
9.3	Infinite Potential Well	46
9.3.1	Zero Point Energy	47
9.3.2	Symmetry of the Solutions	48
9.4	Finite Potential Well	49
10	Harmonic Oscillators	50
10.1	The Classical Harmonic Oscillators	50
10.2	The Quantum Harmonic Oscillator	51
10.3	The Solutions to the Harmonic Oscillator	51
10.4	Solving the Harmonic Oscillator (The Fancy Way)	52
10.4.1	Factorising the Hamiltonian	52
10.4.2	Creation and Annihilation	53
10.4.3	The Wave Function	55
11	Systems in Three Spatial Dimensions	56
11.1	States	56

11.2	Observables	57
11.2.1	Position Operator	57
11.2.2	Momentum Operator	58
11.2.3	The Canonical Commutation Relation	58
11.3	Dynamics	58
11.3.1	Three-Dimensional Harmonic Oscillator	59
IV	Angular Momentum and Spin	60
12	Angular Momentum	60
12.1	Angular Momentum in Classical Mechanics	60
12.2	Angular Momentum in Quantum Mechanics	61
12.2.1	Angular Momentum Operators	61
12.2.2	Angular Momentum Commutators	61
12.2.3	Angular Momentum Differential Operators	62
12.2.4	Square of the Angular Momentum	62
12.2.5	Eigenfunctions of \hat{L}^2 and \hat{L}_z	63
12.3	Algebraic Solution to the Eigenvalue Problem	64
12.3.1	Raising and Lowering Operators	64
12.3.2	Bounds on m	65
12.3.3	Normalisation Factors	66
13	Central Potentials	67
13.1	What is a Central Potential	67
13.2	Stationary States	67
13.3	Normalisation	68
13.4	Behaviour Near the Origin	69
14	Spin	70
14.1	Matrix Elements	72
14.2	Eigenvectors of \hat{S}_z and \hat{S}^2	73
14.3	Scalar Products	74
14.4	Spin Along the x Direction	74
15	Addition of Angular Momenta	75
15.1	Vector Spaces	75
15.2	Total Angular Momentum	76
16	Identical Particles	78
16.1	Helium Atom	78
16.2	Two Electron State	79
16.3	Pauli Exclusion Principle	80
17	The Hydrogen Atom	80
17.1	Stationary States	80
17.1.1	Solution to the Radial Equation	82
17.1.2	Physical Interpretation	83
18	Non-degenerate Time-Independent Perturbation Theory	84
18.1	Zeroth Order	84
18.2	First Order	85
18.2.1	Energy Perturbation	85
18.2.2	Eigenstate Perturbation	85
18.3	Second Order	86
18.4	Perturbation Examples	86
18.4.1	Infinite Well With a Cosine Floor	86
18.4.2	Helium	87

List of Figures

2.1	The Bloch Sphere, the two pure states, $ 1\rangle$ and $ 2\rangle$, are at the poles.	11
3.1	A one dimensional discrete system of $N = 6$ points, x_i , spaced evenly ε apart.	13
9.1	The step potential	43
9.2	The potential barrier	45
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.	46
9.4	The first four wave functions, u_n , for the infinite square well and the the associated probability densities.	48
13.1	$ \chi_\ell(r) ^2$ compared for $\ell = 0, 1, 2$ and $r < 1$	69
17.1	The potential for the hydrogen atom.	81

List of Tables

15.1	Clebsch–Gordan coefficients for two spin 1/2 particles.	78
16.1	The symmetry of various parts of a state under $\hat{\mathcal{P}}_{12}$. S denotes symmetric and A denotes antisymmetric. The symmetry of the total system, $ \Psi\rangle$, is set by the spin statistics theorem. The symmetry of the spin part, $ s, m\rangle$, follows from the expression in the uncoupled basis. The symmetry of the spatial part, $ \psi_{12}\rangle$, is then fixed to make the symmetries of the other parts match.	80

Acronyms

CSCO	complete set of commuting/compatible observables.
HUP	Heisenberg’s uncertainty principle.
PDF	probability density function.
QM	quantum mechanics.
STM	scanning tunnelling microscope.
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, \mathbf{x} , and momentum, \mathbf{p} . These are 3-dimensional real vectors living in \mathbb{R}^3 . We can instead view the object (\mathbf{x}, \mathbf{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 (\mathbf{x}, \mathbf{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, $+: \mathcal{H} \times \mathcal{H} \rightarrow \mathcal{H}$, 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, $\cdot: \mathbb{F} \times \mathcal{H} \rightarrow \mathcal{H}$, 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 $\alpha_i \in \mathbb{C}$ 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:

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

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:

$$\begin{aligned} \cdot: \mathcal{H} \times \mathcal{H} &\rightarrow \mathbb{C} \\ |u\rangle, |v\rangle &\mapsto \langle u|v\rangle = z \in \mathbb{C}. \end{aligned}$$

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 $\alpha, \beta \in \mathbb{C}$

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

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

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

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

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

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,

$$\begin{aligned} f: \mathcal{H} &\rightarrow \mathbb{C} \\ |v\rangle &\mapsto f(|v\rangle) = z \in \mathbb{C}, \end{aligned}$$

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}^* .

⁹Technically 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.

1.2.4 Norms

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

$$\begin{aligned}\|\cdot\|: \mathcal{H} &\rightarrow \mathbb{R} \\ |v\rangle &\mapsto \|v\|\end{aligned}$$

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 $|v\rangle \in \mathcal{H}$, $\lambda \in \mathbb{C}$:

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

- 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

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

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:

$$\begin{aligned}\hat{O}: \mathcal{H} &\rightarrow \mathcal{H} \\ |v\rangle &\mapsto \hat{O}|v\rangle = |w\rangle \in \mathcal{H}.\end{aligned}$$

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

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

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

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

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

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

where 1 is the identity operator defined by

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

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

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

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 set of linearly independent vectors, $\mathcal{B} = \{|e_k\rangle \mid k = 1, \dots, n\} \subset \mathcal{H}$ 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

$$\mathbf{v} = v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2 + v_3 \mathbf{e}_3 = \sum_{i=1}^3 v_i \mathbf{e}_i.$$

Here $v_i \in \mathbb{R}$ are the components of \mathbf{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

$$|v\rangle = \sum_k v_k |e_k\rangle = \sum_k \langle e_k | v \rangle |e_k\rangle = \sum_k |e_k\rangle \langle e_k | v \rangle \quad (1.1)$$

using this we get

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

and then a direct computation gives us

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

$$= \sum_k |v_k|^2.$$

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

$$\sum_k \hat{\mathcal{P}}_k = \sum_k |e_k\rangle \langle e_k| = 1$$

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 :

$$\begin{aligned} v'_k &= \langle e'_k|v\rangle \\ &= \langle e'_k|1|v\rangle \\ &= \langle e'_k|\sum_l |e_l\rangle \langle e_l|v\rangle \\ &= \sum_l \langle e'_k|e_l\rangle \langle e_l|v\rangle \\ &= \sum_l \langle e'_k|e_l\rangle v_l \\ &= \sum_l U_{kl} v_l \end{aligned}$$

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

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

$$\begin{aligned}
&= \langle e'_k | 1 | e'_k \rangle \\
&= \langle e'_k | e'_k \rangle \\
&= 1
\end{aligned}$$

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

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

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 :

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

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:

$$\begin{aligned}\otimes: \mathcal{H}_1 \times \mathcal{H}_2 &\rightarrow \mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \\ |v^{(1)}\rangle, |v^{(2)}\rangle &\mapsto |v^{(1)}\rangle \otimes |v^{(2)}\rangle = |v^{(1)}v^{(2)}\rangle \in \mathcal{H}.\end{aligned}$$

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

$$\begin{aligned}|v^{(1)}\rangle |v^{(2)}\rangle &= [\alpha |w_1^{(1)}\rangle + \beta |w_2^{(1)}\rangle] |v^{(2)}\rangle \\ &= \alpha |w_1^{(1)}\rangle |v^{(2)}\rangle + \beta |w_2^{(1)}\rangle |v^{(2)}\rangle.\end{aligned}$$

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 \mathcal{H}_1 and $\mathcal{B}_2 = \{|e_l^{(2)}\rangle\}$ 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

Suppose we have an operator, $\hat{O}^{(1)}: \mathcal{H}_1 \rightarrow \mathcal{H}_1$, 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, $\hat{O}^{(1)}: \mathcal{H} \rightarrow \mathcal{H}$, 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, $\hat{O}^{(2)}: \mathcal{H}_2 \rightarrow \mathcal{H}_2$, 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 $\hat{O}^{(2)}: \mathcal{H} \rightarrow \mathcal{H}$ as

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

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

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

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

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

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 $\vartheta = \pi$.

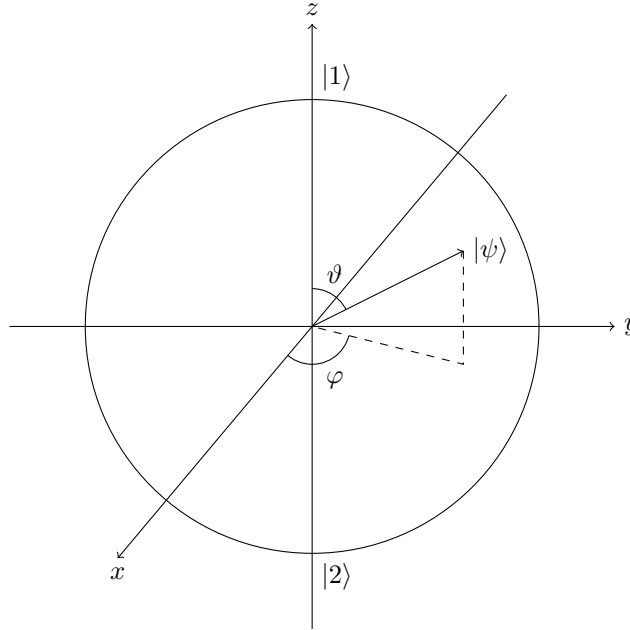


Figure 2.1: The Bloch Sphere, the two pure states, $|1\rangle$ and $|2\rangle$, are at the poles.

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 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

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

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

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

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

$$\begin{aligned} |\varphi^{(1)}\varphi^{(2)}\rangle &= [c_1^{(1)}|1^{(1)}\rangle + c_2^{(1)}|2^{(1)}\rangle] [c_1^{(2)}|1^{(2)}\rangle + c_2^{(2)}|2^{(2)}\rangle] \\ &= c_1^{(1)}c_1^{(2)}|11\rangle + c_1^{(1)}c_2^{(2)}|12\rangle + c_2^{(1)}c_1^{(2)}|21\rangle + c_2^{(1)}c_2^{(2)}|22\rangle \end{aligned}$$

The norm of this vector is

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

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**, $\psi: \mathbb{R} \rightarrow \mathbb{C}$. 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

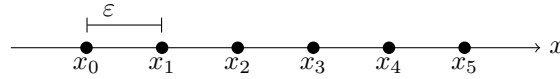


Figure 3.1: A one dimensional discrete system of $N = 6$ points, x_i , spaced evenly ε apart.

wave function to $\{x_i \mid i = 0, \dots, 5\}$. 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 \mathbf{v}_x , it's just wrong.

3.1.1 Basis

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

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

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, $p(x) dx$, here ε plays the role of dx .

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 $N \rightarrow \infty$ and $\varepsilon \rightarrow 0$. After doing this we have that $|\psi\rangle \in \mathcal{H}_\infty$ where \mathcal{H}_∞ 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 $N \rightarrow \infty$ and $\varepsilon \rightarrow 0$ we can write $|\psi\rangle \in \mathcal{H}_\infty$ as

$$|\psi\rangle = \int dx \psi(x) |x\rangle$$

where

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

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

$$|\psi\rangle = \int dx \langle x | \psi \rangle |x\rangle = \int dx |x\rangle \langle x | \psi \rangle.$$

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

$$\int dx |x\rangle\langle x| = 1.$$

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 $N \rightarrow \infty$ and $\varepsilon \rightarrow 0$ we see that for $|\varphi\rangle, |\psi\rangle \in \mathcal{H}_\infty$ we have

$$\langle \varphi | \psi \rangle = \int dx \varphi^*(x) \psi(x).$$

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

$$\|\psi\|^2 = \langle \psi | \psi \rangle = \int dx \psi^*(x) \psi(x) = \int dx |\psi(x)|^2.$$

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

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:

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

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

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

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

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

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

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

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

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

So we see that

$$\int dx' f(x') \langle x | x' \rangle = f(x)$$

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

$$\int dx' f(x') \delta(x - x') = f(x).$$

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

$$p: S \rightarrow [0, 1]$$

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

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

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

$$P(x \in S) = \int_S dx p(x) = 1.$$

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

$$\mu_n = \int dx p(x) x^n.$$

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

$$\mu_1 = \langle x \rangle = \int dx p(x) x.$$

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

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

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

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

We introduce an operator, $\hat{X}: \mathcal{H} \rightarrow \mathcal{H}$, 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

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

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 n th moment fairly easily:

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

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

Here we have used the following theorem.

Theorem 4.1. *If $\hat{O}: \mathcal{H} \rightarrow \mathcal{H}$ 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 $\hat{O}^0 = 1$ so

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

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 $n \in \mathbb{N}$. \square

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, $\hat{O}: \mathcal{H} \rightarrow \mathcal{H}$.

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 $\{O^{(1)}, O^{(2)}, \dots, O^{(n)}\}$. 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

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

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 $\hat{O}: \mathcal{H} \rightarrow \mathcal{H}$ 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 $n = 1, \dots, g$. 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:

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

where we have used the result of theorem 4.1, 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 $\hat{O}: \mathcal{H} \rightarrow \mathcal{H}$. Then we define the **hermitian conjugate** of \hat{O} , denoted \hat{O}^\dagger , to be another operator, $\hat{O}^\dagger: \mathcal{H} \rightarrow \mathcal{H}$, 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 $\hat{O} = \frac{d}{dx}$ defined by

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

Compute \hat{O}^\dagger .

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

inserting completeness gives

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

We can now integrate by parts:

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

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

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

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

So we can identify

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

4.5.2 Hermitian Operator

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

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

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

see equation 4.1 for how to do this integral

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

So we can identify that

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

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 $\hat{O}: \mathcal{H} \rightarrow \mathcal{H}$ 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 $\hat{O}: \mathcal{H} \rightarrow \mathcal{H}$,

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

where $O_k \neq O_l$ then

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

3. If $\hat{O}: \mathcal{H} \rightarrow \mathcal{H}$ 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 4.2. *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

$$\langle\psi_k|\hat{O}|\psi_k\rangle = \langle\psi_k|O_k|\psi_k\rangle = O_k\langle\psi_k|\psi_k\rangle = O_k \quad (4.2)$$

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 $O_k \in \mathbb{R}$. □

Theorem 4.3. *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

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

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

$$\langle\psi_2|\hat{O}|\psi_1\rangle = O_1\langle\psi_2|\psi_1\rangle. \quad (4.3)$$

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. □

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, $\hat{O}: \mathcal{H} \rightarrow \mathcal{H}$, which has a nondegenerate 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 eigenvalue 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 $n = 1, \dots, g_k$. The state after measurement is then

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

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

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

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 $|v_k\rangle = |u_j\rangle$ 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 5.1 (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 $n = 1, \dots, g_k$. Suppose that

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

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 $\{A, B, C, \dots\}$ 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

$$\int df |f\rangle\langle f| = 1.$$

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

$$|\psi\rangle = \int df c(f)|f\rangle.$$

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:

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

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

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

To be properly normalised we demand that

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

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:

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

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

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

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:

$$|\psi\rangle = \int df c(f)|f\rangle.$$

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

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

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

$$\int df \langle f|f\rangle = 1.$$

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

$$|\psi\rangle = \int dx \psi(x)|x\rangle$$

where

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

By definition

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

Thus

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

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.

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

Requiring a normalised state we then have

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

The expected position is given by

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

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, $\hat{P}: \mathcal{H} \rightarrow \mathcal{H}$, is defined by

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

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

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

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

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

However d/dx 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

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

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:

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

This holds for all ψ therefore we conclude that

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

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:

$$\mathbf{F} = m\mathbf{a} = m \frac{d^2 \mathbf{r}}{dt^2}.$$

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**:

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

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

$$\begin{aligned} T = \frac{p^2}{2m} &\rightarrow \hat{T} = \frac{\hat{P}^2}{2m}, \\ V = V(x) &\rightarrow \hat{V} = V(\hat{X}). \end{aligned}$$

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

$$\begin{aligned} |\Psi(t + \varepsilon)\rangle &= |\Psi(t)\rangle + \varepsilon \frac{d}{dt} |\Psi(t)\rangle \\ &= |\Psi(t)\rangle - \frac{i\varepsilon}{\hbar} \hat{H} |\Psi(t)\rangle \end{aligned}$$

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

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

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

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

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:

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

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

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

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 :

$$\|\Psi(t)\|^2 = \int dx \Psi^*(x, t)\Psi(x, t) = 1.$$

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

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

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

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

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

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

We can easily check that this satisfies the Schrödinger equation:

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

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

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

This gives us

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

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:

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

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:

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

We can show that

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

is a solution to the Schrödinger equation:

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

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

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

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

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

Let $k' = k - 1$:

$$\begin{aligned}
&= \hat{H} \sum_{k'=0}^{\infty} \frac{1}{k'!} \left(-\frac{i}{\hbar}\right)^{k'} \hat{H}^{k'} |\Psi(0)\rangle t^{k'} \\
&= \hat{H} \exp\left(-\frac{i}{\hbar} \hat{H} t\right) |\Psi(0)\rangle. \\
&= \hat{H} \hat{U}(t) |\Psi(0)\rangle.
\end{aligned}$$

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

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

From now on we will assume this.

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

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

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

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

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

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:

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

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

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

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

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

Then using the definition of the hermitian conjugate we have

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

noting that \hat{O} is hermitian we have

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

Hence

$$\begin{aligned}
\frac{d}{dt} O(t) &= \frac{i}{\hbar} \langle \Psi(t) | \hat{H} \hat{O} | \Psi(t) \rangle - \frac{i}{\hbar} \langle \Psi(t) | \hat{O} \hat{H} | \Psi(t) \rangle \\
&= \frac{i}{\hbar} \langle \Psi(t) | (\hat{H} \hat{O} - \hat{O} \hat{H}) | \Psi(t) \rangle
\end{aligned}$$

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

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:

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

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

$$\begin{aligned}
&= \frac{i}{\hbar} \langle [\hat{H}, \hat{O}] \rangle + \int dx \Psi^*(x, t) \frac{\partial \hat{O}}{\partial t} \Psi(x, t) \\
\frac{d}{dt} \langle O \rangle &= \frac{i}{\hbar} \langle [\hat{H}, \hat{O}] \rangle + \left\langle \frac{\partial \hat{O}}{\partial t} \right\rangle.
\end{aligned}$$

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

$$\int dx |x\rangle \langle x| = 1.$$

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

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

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

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

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

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

Thus the TISE can be written as

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

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

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

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

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

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

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

The solutions to this are known to be

$$\begin{aligned} \psi_{k,+}(x) &= e^{ikx} \\ \psi_{k,-}(x) &= e^{-ikx} \end{aligned}$$

We can perform a brief dimensions check here.

$$[k]^2 = \frac{[\text{mass}][\text{energy}]}{[\text{energy}]^2[\text{time}]^2} = \frac{[\text{mass}]}{[\text{energy}][\text{time}]^2} = \frac{[\text{mass}]}{[\text{mass}][\text{length}]^2[\text{time}]^{-22}[\text{time}]^2} = [\text{length}]^{-2}$$

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

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 $\psi_{k,\pm}(x)$:

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

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

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

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 the $-x$ 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, $\hat{\mathcal{P}}$, defined as

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

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

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

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

$$\begin{aligned} \langle \varphi | \hat{\mathcal{P}}^\dagger | \psi \rangle &= (\langle \psi | \hat{\mathcal{P}} | \varphi \rangle)^* \\ &= \left(\int_{-\infty}^{\infty} dx \langle \psi | x \rangle \langle x | \hat{\mathcal{P}} | \varphi \rangle \right)^* \\ &= \left(\int_{-\infty}^{\infty} dx \psi^*(x) \hat{\mathcal{P}}\varphi(x) \right)^* \\ &= \left(\int_{-\infty}^{\infty} dx \psi^*(x) \varphi(-x) \right)^* \end{aligned}$$

Let $u = -x$ then $dx \rightarrow -du$ and $(-\infty, \infty) \rightarrow (\infty, -\infty)$:

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

$$\begin{aligned}
&= \left(\int_{-\infty}^{\infty} du \psi^*(-u) \varphi(u) \right)^* \\
&= \int_{-\infty}^{\infty} du \varphi^*(u) \psi(-u) \\
&= \int_{-\infty}^{\infty} du \varphi^*(u) \hat{\mathcal{P}} \psi(u) \\
&= \int_{-\infty}^{\infty} du \langle \varphi | u \rangle \langle u | \hat{\mathcal{P}} | \psi \rangle \\
&= \langle \varphi | \hat{\mathcal{P}} | \psi \rangle.
\end{aligned}$$

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

8.4 Time Dependent Schrödinger Equation

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

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

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

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

Here we have used the fact that $|x\rangle$ is independent of time so commutes with d/dt . 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 exist $\partial_x \Psi(x, t)$ 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:

$$\begin{aligned}
\langle \psi | \hat{H} | \psi \rangle &= \int_{-\infty}^{\infty} dx \langle \psi | x \rangle \langle x | \hat{H} | \psi \rangle \\
&= \int_{-\infty}^{\infty} dx \psi^*(x) \hat{H} \psi(x) \\
&= \int_{-\infty}^{\infty} dx \psi^*(x) \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \psi(x) \\
&= - \int_{-\infty}^{\infty} dx \frac{\hbar^2}{2m} \psi^*(x) \frac{d^2 \psi}{dx^2} + \int_{-\infty}^{\infty} dx \psi^*(x) V(x) \psi(x)
\end{aligned}$$

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

Focusing on the first integral we can integrate by parts:

$$\begin{aligned} u &= \psi^*, & v &= \frac{d\psi}{dx}, \\ u' &= \frac{d\psi^*}{dx}, & v' &= \frac{d^2\psi}{dx^2}. \end{aligned}$$

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

$$\begin{aligned} \int dx \frac{\hbar^2}{2m} \psi^*(x) \frac{d^2\psi}{dx^2} &= \left[\psi^* \frac{d\psi}{dx} \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} dx \frac{d\psi^*}{dx} \frac{d\psi}{dx} \\ &= - \int_{-\infty}^{\infty} dx \left| \frac{d\psi}{dx} \right|^2 \end{aligned}$$

Hence the expectation value of the energy is

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

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

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

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

$$\begin{aligned} &\geq \int_{-\infty}^{\infty} dx V_{\min} |\psi(x)|^2 \\ &= V_{\min} \int_{-\infty}^{\infty} |\psi(x)|^2 \\ &= V_{\min} \end{aligned}$$

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:

$$\begin{aligned} \frac{d}{dt} \langle m\hat{X} \rangle_t &= \frac{d}{dt} \langle \Psi(t) | m\hat{X} | \Psi(t) \rangle \\ &= m \frac{d}{dt} \langle \Psi(t) | \hat{X} | \Psi(t) \rangle \\ &= \frac{mi}{\hbar} \langle \Psi(t) | [\hat{H}, \hat{X}] | \Psi(t) \rangle. \end{aligned}$$

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

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

$$\begin{aligned}
&= \frac{1}{2m} [\hat{P}^2, \hat{X}] \\
&= \frac{1}{2m} (\hat{P}^2 \hat{X} - \hat{X} \hat{P}^2) \\
&= \frac{1}{2m} (\hat{P}^2 \hat{X} - \hat{P} \hat{X} \hat{P} + \hat{P} \hat{X} \hat{P} - \hat{X} \hat{P}^2) \\
&= \frac{1}{2m} (\hat{P}(\hat{P} \hat{X} - \hat{X} \hat{P}) + (\hat{P} \hat{X} - \hat{X} \hat{P}) \hat{P}) \\
&= \frac{1}{2m} (\hat{P}[\hat{P}, \hat{X}] + [\hat{P}, \hat{X}] \hat{P}) \\
&= -\frac{1}{2m} (\hat{P}[\hat{X}, \hat{P}] + [\hat{X}, \hat{P}] \hat{P}) \\
&= -\frac{1}{2m} (\hat{P} i\hbar + i\hbar \hat{P}) \\
&= -\frac{i\hbar}{m} \hat{P}.
\end{aligned}$$

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

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

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

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

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

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

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

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

this holds for all ψ so

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

Hence

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

$$\begin{aligned}
&= -\langle \Psi(t) | \frac{d}{dx} V(\hat{X}) | \Psi(t) \rangle \\
&= -\left\langle \frac{d}{dx} V(\hat{X}) \right\rangle_t.
\end{aligned}$$

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

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

8.6 Degeneracy of the Hamiltonian

Theorem 8.1. *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 :

$$\begin{aligned}
\hat{H}|\psi_1\rangle &= E|\psi_1\rangle \\
\hat{H}|\psi_2\rangle &= E|\psi_2\rangle.
\end{aligned}$$

The Schrödinger equation gives us

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

for any wave function ψ . Rearranging we get

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

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

$$\begin{aligned}
\frac{d^2\psi_1}{dx^2} &= -\frac{2m}{\hbar^2} [E - V(x)]\psi_1(x), \\
\frac{d^2\psi_2}{dx^2} &= -\frac{2m}{\hbar^2} [E - V(x)]\psi_2(x).
\end{aligned}$$

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

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

and

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

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

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

Consider now

$$\begin{aligned}
\frac{d}{dx} \left[\frac{d\psi_1}{dx} \psi_2(x) - \psi_1(x) \frac{d\psi_2}{dx} \right] &= \frac{d^2\psi_1}{dx^2} \psi_2(x) + \frac{d\psi_1}{dx} \frac{d\psi_2}{dx} - \frac{d\psi_1}{dx} \frac{d\psi_2}{dx} - \psi_1(x) \frac{d^2\psi_2}{dx^2} \\
&= \frac{d^2\psi_1}{dx^2} \psi_2(x) - \psi_1(x) \frac{d^2\psi_2}{dx^2}.
\end{aligned}$$

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

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

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

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

Which means

$$\begin{aligned} \frac{d\psi_1}{dx} \psi_2(x) &= \psi_1(x) \frac{d\psi_2}{dx} \\ \implies \frac{\psi_1'(x)}{\psi_1(x)} &= \frac{\psi_2'(x)}{\psi_2(x)} \\ \implies \frac{d}{dx} \ln(\psi_1(x)) &= \frac{d}{dx} \ln(\psi_2(x)) \\ \implies \ln(\psi_1(x)) &= \ln(\psi_2(x)) + \tilde{\kappa} \end{aligned}$$

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. \square

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:

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

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 $E < V$.

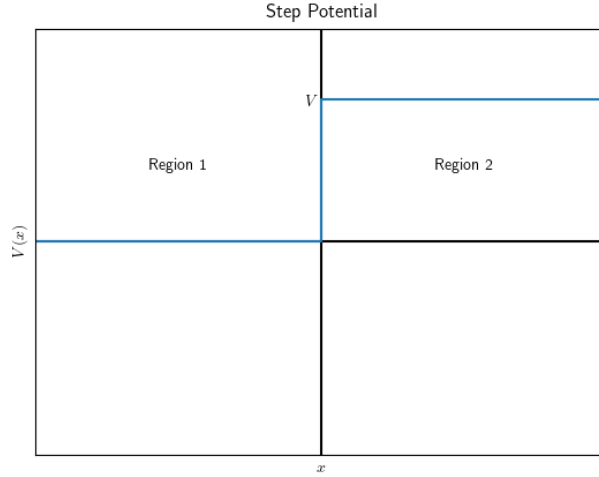


Figure 9.1: The step potential

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 sections, in region 1, $x < 0$, we have

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

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

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

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

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

and

$$\lim_{x \rightarrow 0^+} \psi'(x) = \lim_{x \rightarrow 0^-} \psi'(x).$$

The first of these gives us

$$\begin{aligned} \lim_{x \rightarrow 0^+} \psi(x) &= \lim_{x \rightarrow 0^+} [Ae^{ipx/\hbar} + Be^{-ikx/\hbar}] = A + B \\ \lim_{x \rightarrow 0^-} \psi(x) &= \lim_{x \rightarrow 0^-} Ce^{i\bar{p}x/\hbar} = C. \end{aligned}$$

Hence

$$A + B = C.$$

The second leads to

$$\begin{aligned} \lim_{x \rightarrow 0^+} \psi'(x) &= \lim_{x \rightarrow 0^+} \left[\frac{ip}{\hbar} Ae^{ipx/\hbar} - \frac{ip}{\hbar} Be^{-ikx/\hbar} \right] = \frac{ip}{\hbar} A + \frac{ip}{\hbar} B \\ \lim_{x \rightarrow 0^-} \psi'(x) &= \lim_{x \rightarrow 0^-} \frac{i\bar{p}}{\hbar} Ce^{i\bar{p}x/\hbar} = \frac{i\bar{p}}{\hbar} C. \end{aligned}$$

Hence

$$p(A + B) = \bar{p}C.$$

We can then show that

$$B = \frac{p - \bar{p}}{p + \bar{p}} A, \quad \text{and} \quad C = \frac{2p}{p + \bar{p}} A.$$

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

$$\psi(x) = Ce^{-\tilde{p}x/\hbar} + De^{\tilde{p}x/\hbar}.$$

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 $x \rightarrow \infty$. This is not allowed so we discard this term in order for ψ to be normalisable. Hence in region 2

$$\psi(x) = Ce^{-\tilde{p}x/\hbar}.$$

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 $E > V$

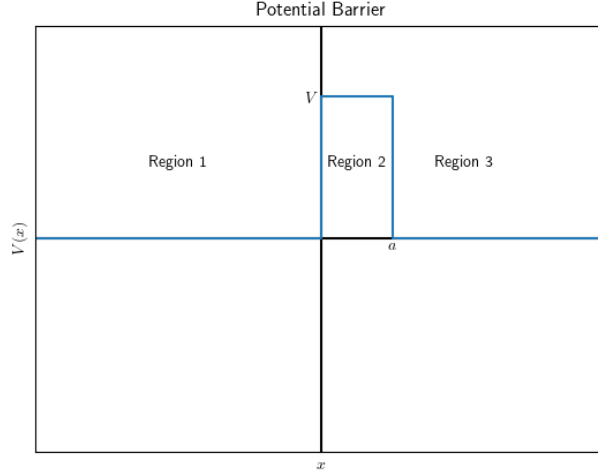


Figure 9.2: The potential barrier

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.

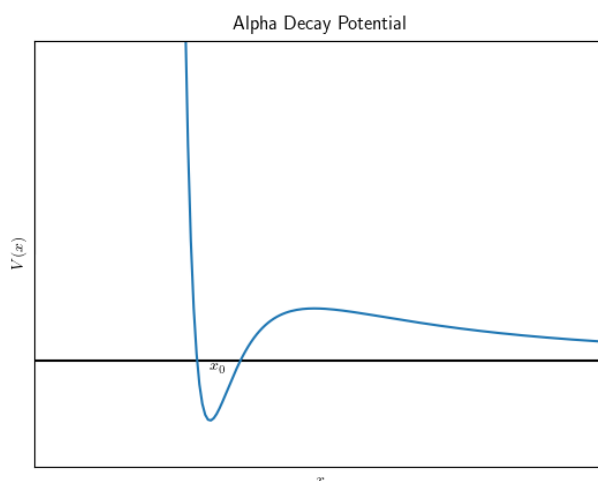


Figure 9.3: The potential modelling the alpha decay process. An alpha particle in the nucleus is in the potential well at x_0 , it can then tunnel outside of this well to some positive x value even if it has less energy than the peak on the right of the potential well.

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 $V \rightarrow \infty$. The solution in these regions is

$$\psi \sim e^{-\tilde{p}\delta/\hbar}.$$

For finite V

$$\tilde{p} = \sqrt{2m(V - E)}.$$

As $V \rightarrow \infty$ $\tilde{p} \rightarrow \infty$ and so $\psi \rightarrow 0$. 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

$$ka = n\pi \implies k = \frac{n\pi}{a}, \quad n \in \mathbb{N}.$$

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 n th 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:

$$E \sim \frac{(\Delta p)^2}{2m} \geq \frac{\hbar^2}{8ma}.$$

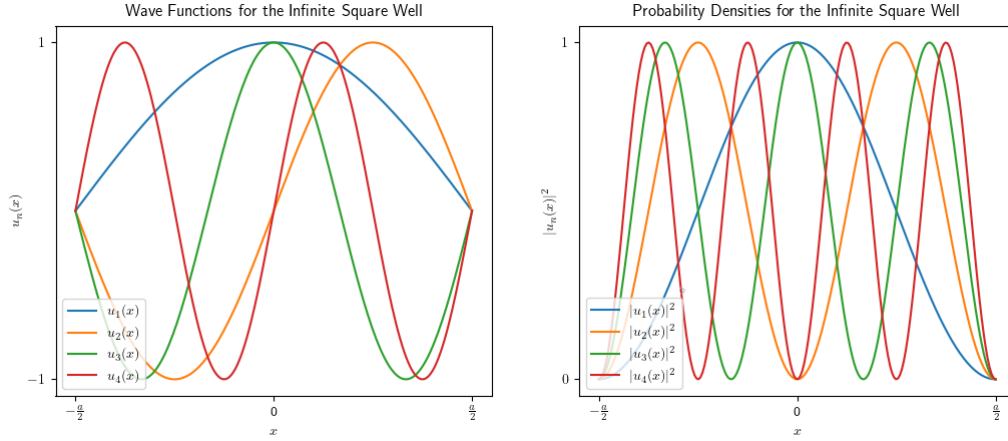


Figure 9.4: The first four wave functions, u_n , for the infinite square well and the the associated probability densities.

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 $T = 0$ 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, $\hat{\mathcal{P}}$, to have the action

$$\hat{\mathcal{P}}\psi(x) = \psi(-x),$$

then the energy eigenstates are also eigenstates of the parity operator:

$$\hat{\mathcal{P}}u_n(x) = (-1)^{n+1}u_n(x).$$

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 $\hat{\mathcal{P}}$, \hat{P} , and \hat{H} .

We can show that in general if the potential is symmetric, so $V(-x) = V(x)$, then $[\hat{H}, \hat{\mathcal{P}}] = 0$, that is \hat{H} is symmetric under $\hat{\mathcal{P}}$. This then implies the existence of a basis of simultaneous eigenstates of \hat{H} and $\hat{\mathcal{P}}$.

It is fairly simple to show the above paragraph is true. Suppose that V is a symmetric potential. Then

$$\begin{aligned} \hat{\mathcal{P}}\hat{H}\psi(x) &= \hat{\mathcal{P}}\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right]\psi(x) \\ &= \hat{\mathcal{P}}\left[-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x)\right] \\ &= -\frac{\hbar^2}{2m}\frac{d^2\psi(-x)}{dx^2} + V(-x)\psi(-x) \\ &= -\frac{\hbar^2}{2m}\frac{d^2\psi(-x)}{dx^2} + V(x)\psi(-x). \end{aligned}$$

Here we have used the fact that

$$\hat{\mathcal{P}}\frac{d}{dx} = \frac{d}{d-x} = -\frac{d}{dx}$$

and hence

$$\hat{\mathcal{P}}\frac{d^2}{dx^2} = -\frac{d}{dx}\left(-\frac{d}{dx}\right) = \frac{d^2}{dx^2}.$$

Also

$$\hat{H}\hat{\mathcal{P}}\psi(x) = \hat{H}\psi(-x)$$

$$\begin{aligned}
&= \left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] \psi(-x) \\
&= -\frac{\hbar^2}{2m} \frac{d^2 \psi(-x)}{dx^2} + V(x) \psi(-x).
\end{aligned}$$

Hence

$$[\hat{\mathcal{P}}, \hat{H}] = 0.$$

We know that this means there is a common basis of eigenstates but how do we find it? The states

$$v_n^+(x) = u_n(x) + u_n(-x), \quad \text{and} \quad v_n^-(x) = u_n(x) - u_n(-x)$$

are one such basis. These states can be compactly written as

$$v_n^\pm(x) = (1 \pm \hat{\mathcal{P}})u_n(x).$$

It is then easy to show that these are eigenstates of \hat{H} and $\hat{\mathcal{P}}$:

$$\begin{aligned}
\hat{H}v_n^\pm(x) &= \hat{H}(1 \pm \hat{\mathcal{P}})u_n(x) = \hat{H}u_n(x) \pm \hat{H}\hat{\mathcal{P}}u_n(x) = \hat{H}u_n(x) \pm \hat{\mathcal{P}}\hat{H}u_n(x) \\
&= E_n u_n(x) + E_n \hat{\mathcal{P}}u_n(x) = E_n(1 \pm \hat{\mathcal{P}})u_n = E_n v_n^\pm(x).
\end{aligned}$$

and

$$\begin{aligned}
\hat{\mathcal{P}}v_n^\pm(x) &= \hat{\mathcal{P}}(1 \pm \hat{\mathcal{P}})u_n(x) = \hat{\mathcal{P}}u_n(x) \pm \hat{\mathcal{P}}^2 u_n(x) = \hat{\mathcal{P}}u_n(x) \pm (-1)^{n+1}(-1)^{n+1}u_n(x) \\
&= \hat{\mathcal{P}}u_n(x) \pm u_n(x) = \hat{\mathcal{P}}u_n(x) \pm u_n(x) = \pm(1 \pm \hat{\mathcal{P}})u_n(x)
\end{aligned}$$

9.4 Finite Potential Well

The finite potential well has the potential

$$V(x) = \begin{cases} -V_0, & x \in (-a/2, a/2), \\ 0, & x \notin (-a/2, a/2), \end{cases}$$

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

$$\psi(x) = \begin{cases} A \cos(px/\hbar), & |x| < a/2, \\ C e^{-\bar{p}x/\hbar}, & x > a/2, \\ C e^{\bar{p}x/\hbar}, & x < -a/2. \end{cases}$$

The odd parity solutions are

$$\psi(x) = \begin{cases} A \sin(px/\hbar), & |x| < a/2, \\ C e^{-\bar{p}x/\hbar}, & x > a/2, \\ -C e^{\bar{p}x/\hbar}, & x < -a/2. \end{cases}$$

We have introduced the momenta

$$p = \sqrt{2m(E + V_0)}, \quad \text{and} \quad \bar{p} = \sqrt{-2mE}.$$

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

$$\begin{aligned} A \cos\left(\frac{pa}{2\hbar}\right) &= C e^{-\bar{p}a/2\hbar} \\ -\frac{p}{\hbar} A \sin\left(\frac{pa}{2\hbar}\right) &= -\frac{\bar{p}a}{\hbar} C e^{-\bar{p}a/2\hbar}. \end{aligned}$$

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

$$\begin{aligned} A \sin\left(\frac{pa}{2\hbar}\right) &= C e^{-\bar{p}a/2\hbar} \\ \frac{p}{\hbar} A \cos\left(\frac{pa}{2\hbar}\right) &= -\frac{\bar{p}a}{\hbar} C e^{-\bar{p}a/2\hbar}. \end{aligned}$$

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

$$E = -\frac{mV_0^2 a^2}{2\hbar^2}$$

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

$$\ddot{x} = \frac{d^2x}{dt^2} = \frac{dv}{dt} = \frac{dv}{dx} \frac{dx}{dt} = v \frac{dv}{dx}.$$

Integrating the equation of motion we then have

$$m \int v \frac{dv}{dx} dx = -k \int x dx \implies \frac{1}{2}mv^2 = -\frac{1}{2}kx^2 + E$$

where E is the total energy of the system, which appears as a constant of integration. Thus

$$\frac{1}{2}mv^2 + \frac{1}{2}kx^2 = E.$$

Identifying the first term as the kinetic energy we see that the potential energy is

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

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

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2 \right] u_n(x) = E_n u_n(x).$$

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:

$$E_n = \left(n + \frac{1}{2} \right) \hbar\omega, \quad \text{where} \quad n \in \mathbb{N}.$$

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

$$\int_{-\infty}^{\infty} ds e^{-s^2} H_m(s) H_n(s) = 2^n \sqrt{\pi} n! \delta_{mn},$$

which is only non-zero if $n = m$. The first four Hermite polynomials are

$$\begin{aligned} H_0(s) &= 1, & H_2(s) &= 4s^2 - 2, \\ H_1(s) &= 2s, & H_3(s) &= 8s^3 - 12s. \end{aligned}$$

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

$$\hat{\mathcal{P}}u_n(x) = u_n(-x) = (-1)^n u_n(x).$$

We see that the energy eigenstates are eigenstates of the parity operator, $\hat{\mathcal{P}}$, 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

$$\hat{\xi} = \sqrt{\frac{m\omega}{2\hbar}}, \quad \text{and} \quad \hat{\eta} = \frac{1}{\sqrt{2m\hbar\omega}} \hat{P}.$$

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:

$$\hat{a} = \hat{\xi} + i\hat{\eta}, \quad \text{and} \quad \hat{a}^\dagger = \hat{\xi} - i\hat{\eta}.$$

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} :

$$\begin{aligned}\hat{a} &= \sqrt{\frac{m\omega}{2\hbar}}\hat{X} + \frac{i}{\sqrt{2m\omega\hbar}}\hat{P} = \sqrt{\frac{m\omega}{2\hbar}}\left[\hat{X} + \frac{i}{m\omega}\hat{P}\right] \\ \hat{a}^\dagger &= \sqrt{\frac{m\omega}{2\hbar}}\hat{X} - \frac{i}{\sqrt{2m\omega\hbar}}\hat{P} = \sqrt{\frac{m\omega}{2\hbar}}\left[\hat{X} - \frac{i}{m\omega}\hat{P}\right]\end{aligned}$$

We can then compute

$$\begin{aligned}\hat{a}\hat{a}^\dagger &= (\hat{\xi} + i\hat{\eta})(\hat{\xi} - i\hat{\eta}) \\ &= \hat{\xi}^2 + i\hat{\eta}\hat{\xi} - i\hat{\xi}\hat{\eta} + \hat{\eta}^2 \\ &= \hat{\xi} + i[\hat{\eta}, \hat{\xi}] + \hat{\eta}^2,\end{aligned}$$

and

$$\begin{aligned}\hat{a}^\dagger\hat{a} &= (\hat{\xi} - i\hat{\eta})(\hat{\xi} + i\hat{\eta}) \\ &= \hat{\xi}^2 + i\hat{\xi}\hat{\eta} - i\hat{\eta}\hat{\xi} + \hat{\eta}^2 \\ &= \hat{\xi}^2 - i[\hat{\eta}, \hat{\xi}] + \hat{\eta}^2.\end{aligned}$$

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

$$\begin{aligned}\hat{H}(\hat{a}|n\rangle) &= \hat{a}\hat{H}|n\rangle + [\hat{H}, \hat{a}]|n\rangle \\ &= E_n\hat{a}|n\rangle - \hbar\omega\hat{a}|n\rangle \\ &= (E_n - \hbar\omega)\hat{a}|n\rangle.\end{aligned}$$

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

$$\begin{aligned}\hat{H}(\hat{a}^\dagger|n\rangle) &= \hat{a}^\dagger\hat{H}|n\rangle + [\hat{H}, \hat{a}^\dagger]|n\rangle \\ &= E_n\hat{a}^\dagger|n\rangle + \hbar\omega\hat{a}^\dagger|n\rangle \\ &= (E_n + \hbar\omega)\hat{a}^\dagger|n\rangle.\end{aligned}$$

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

$$\hat{a}|n\rangle = c_n|n-1\rangle, \quad \text{and} \quad \hat{a}^\dagger|n\rangle = d_n|n+1\rangle.$$

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:

$$\hat{H}|0\rangle = \hbar\omega \left(\hat{a}^\dagger\hat{a} + \frac{1}{2} \right) |0\rangle = \hbar\omega \underbrace{\hat{a}^\dagger\hat{a}}_{=0} |0\rangle + \frac{\hbar\omega}{2}|0\rangle = \frac{\hbar\omega}{2}|0\rangle$$

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

$$\hat{a}^{\dagger 2}|0\rangle = d_0\hat{a}^\dagger|1\rangle = d_0d_1|2\rangle.$$

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:

$$\begin{aligned}\langle n|\hat{a}^\dagger\hat{a}|n\rangle &= c_n\langle n|\hat{a}^\dagger|n-1\rangle \\ &= c_n\langle n-1|\hat{a}|n\rangle^* \\ &= c_n c_n^* \langle n-1|n-1\rangle \\ &= |c_n|^2.\end{aligned}$$

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

$$\begin{aligned}
 \langle n|\hat{a}\hat{a}^\dagger|n\rangle &= d_n\langle n|\hat{a}|n+1\rangle \\
 &= d_n\langle n+1|\hat{a}^\dagger|n\rangle^* \\
 &= d_nd_n^*\langle n+1|n+1\rangle
 \end{aligned}$$

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

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle, \quad \text{and} \quad \hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle.$$

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:

$$\begin{aligned}
 0 &= \hat{a}u_0(x) \\
 &= \sqrt{\frac{m\omega}{2\hbar}}\left[\hat{X} + \frac{i}{m\omega}\hat{P}\right]u_0(x) \\
 &= \sqrt{\frac{m\omega}{2\hbar}}\left[x - \frac{i}{m\omega}\hbar\frac{d}{dx}\right]u_0(x) \\
 \implies 0 &= xu_0(x) + \frac{\hbar}{m\omega}\frac{du_0}{dx}
 \end{aligned}$$

$$\begin{aligned}
\Rightarrow \frac{du_0}{dx} &= -\frac{m\omega}{\hbar}xu_0(x) \\
&= \alpha^2xu_0(x) \\
\Rightarrow \frac{1}{u_0}\frac{du_0}{dx} &= -\alpha x \\
\Rightarrow \int \frac{1}{u_0} du_0 &= -\alpha^2 \int x dx \\
\Rightarrow \ln(u_0) &= -\frac{1}{2}\alpha^2 x^2 + \ln(N)
\end{aligned}$$

where $\ln C$ is a constant of integration

$$\Rightarrow u_0 = Ce^{-\alpha^2 x^2/2}.$$

We can find C by requiring the state to be normalised so

$$\begin{aligned}
1 &= \int_{-\infty}^{\infty} |u_0(x)|^2 dx \\
&= C^2 \int_{-\infty}^{\infty} e^{-\alpha^2 x^2} dx \\
&= C^2 \frac{\sqrt{\pi}}{\alpha} \\
\Rightarrow C &= \frac{\sqrt{\alpha}}{\pi^{1/4}}.
\end{aligned}$$

where we have used the standard result

$$\int_{-\infty}^{\infty} e^{-p(x+q)^2} dx = \sqrt{\frac{\pi}{p}}.$$

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

$$\begin{aligned}
u_n(x) &= \frac{1}{\sqrt{n!}}\hat{a}^{\dagger n}u_0(x) \\
&= \frac{1}{\sqrt{n!}} \left[\sqrt{\frac{m\omega}{2\hbar}} \left(\hat{X} - \frac{i}{m\omega}\hat{P} \right) \right]^n u_0(x) \\
&= \frac{1}{\sqrt{n!}} \left[\sqrt{\frac{m\omega}{2\hbar}} \left(x - \frac{\hbar}{m\omega} \frac{d}{dx} \right) \right]^n u_0(x)
\end{aligned}$$

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, \mathbf{r} , so the basis is $\{|\mathbf{r}\rangle\}$. A generic state vector is then

$$|\psi\rangle = \int d^3r \langle \mathbf{r} | \psi \rangle |\mathbf{r}\rangle = \int d^3r \psi(\mathbf{r}) |\mathbf{r}\rangle$$

where

$$\psi(\mathbf{r}) = \langle \mathbf{r} | \psi \rangle.$$

The probability density is given by $|\psi(\mathbf{r})|^2$ and $|\psi(\mathbf{r})|^2 d^3r$ then gives the probability that the particle is found in the volume element d^3r at position \mathbf{r} . Introducing some time dependence a generic state becomes

$$|\Psi(t)\rangle = \int d^3r |\mathbf{r}\rangle \langle \mathbf{r} | \Psi(t) \rangle = \int d^3r \Psi(\mathbf{r}, t) |\mathbf{r}\rangle,$$

where

$$\Psi(\mathbf{r}, t) = \langle \mathbf{r} | \Psi(t) \rangle.$$

An inner product between two states, $|\Psi(t)\rangle, |\Phi(t)\rangle \in \mathcal{H}$ is

$$\langle \Phi(t) | \Psi(t) \rangle = \int d^3r \Phi^*(\mathbf{r}, t) \Psi(\mathbf{r}, t).$$

As usual we require normalised states so

$$\langle \Psi(t) | \Psi(t) \rangle = \int_{\text{all space}} d^3r |\Psi(t)|^2 = 1.$$

The only change so far from the one-dimensional case is that x becomes \mathbf{r} and dx becomes d^3r .

11.2 Observables

An observable, O , is still associated with a Hermitian operator, $\hat{O}: \mathcal{H} \rightarrow \mathcal{H}$ which we can think of as acting on the wave function with the usual identification that

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

The eigenvalue equation is still

$$\hat{O}\psi_k(\mathbf{r}) = O_k\psi_k(\mathbf{r})$$

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

$$\int d^3r \varphi^*(\mathbf{r}) \hat{O}^\dagger \psi(\mathbf{r}) = \left[\int d^3r \psi^*(\mathbf{r}) \hat{O} \varphi(\mathbf{r}) \right]^*.$$

11.2.1 Position Operator

The three-dimensional position operator is a vector of one-dimensional position operators:

$$\hat{\mathbf{X}} = (\hat{X}, \hat{Y}, \hat{Z}) = (\hat{X}_1, \hat{X}_2, \hat{X}_3).$$

Which have the expected application of multiplication by their respective Cartesian coordinate:

$$\hat{X}_k \psi(\mathbf{r}) = x_k \psi(\mathbf{r})$$

where $\mathbf{r} = (x_1, x_2, x_3)$.

11.2.2 Momentum Operator

The three-dimensional momentum operator is a vector of one-dimensional momentum operators:

$$\hat{\mathbf{P}} = (\hat{P}_x, \hat{P}_y, \hat{P}_z) = (\hat{P}_1, \hat{P}_2, \hat{P}_3).$$

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

$$\hat{P}_k \psi(\mathbf{r}) = -i\hbar \frac{\partial \psi}{\partial x_k} \Big|_{\mathbf{r}}.$$

We can concisely write the momentum operator with some vector calculus notation as

$$\hat{\mathbf{P}} = -i\hbar \nabla$$

where $\nabla = (\partial_1, \partial_2, \partial_3)$ is the gradient operator.

11.2.3 The Canonical Commutation Relation

The partial derivatives commute so

$$\hat{P}_i \hat{P}_j \psi(\mathbf{r}) = -\hbar^2 \frac{\partial^2 \psi}{\partial x_i \partial x_j} = -\hbar^2 \frac{\partial^2 \psi}{\partial x_j \partial x_i} = \hat{P}_j \hat{P}_i \psi(\mathbf{r}).$$

This means that

$$[\hat{P}_i, \hat{P}_j] = 0.$$

Clearly $x_i x_j = x_j x_i$ 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,

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

where

$$\hat{H} = \frac{\hat{\mathbf{P}} \cdot \hat{\mathbf{P}}}{2m} + V(\hat{\mathbf{X}}).$$

The action of this one a wave function is then

$$\hat{H} = \frac{-\hbar^2}{2m} \nabla^2 + V(\mathbf{r}).$$

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

$$V(\mathbf{r}) = \frac{1}{2}m\omega^2 r^2 = \frac{1}{2}m\omega^2(x^2 + y^2 + z^2).$$

We can separate the TISE into a product of three terms,

$$\psi(\mathbf{r}) = X(x)Y(y)Z(z),$$

which each depend on only one coordinate. Using this the TISE becomes

$$\begin{aligned} EX(x)Y(y)Z(z) &= \left[-\frac{\hbar^2}{2m} \nabla^2 + \frac{1}{2}m\omega^2(x^2 + y^2 + z^2) \right] X(x)Y(y)Z(z) \\ &= -\frac{\hbar^2}{2m} \left(\frac{d^2X}{dx^2} Y(y)Z(z) + X(x) \frac{d^2Y}{dy^2} Z(z) + X(x)Y(y) \frac{d^2Z}{dz^2} \right) \\ &\quad + \frac{1}{2}m\omega^2(x^2 + y^2 + z^2)X(x)Y(y)Z(z) \end{aligned}$$

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

$$E = \left[-\frac{\hbar^2}{2m} \frac{d^2X}{dx^2} \frac{1}{X} + \frac{1}{2}m\omega^2 x^2 \right] + \left[-\frac{\hbar^2}{2m} \frac{d^2Y}{dy^2} \frac{1}{Y} + \frac{1}{2}m\omega^2 y^2 \right] + \left[-\frac{\hbar^2}{2m} \frac{d^2Z}{dz^2} \frac{1}{Z} + \frac{1}{2}m\omega^2 z^2 \right].$$

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

$$E_{n_x} = \left[-\frac{\hbar^2}{2m} \frac{d^2X}{dx^2} \frac{1}{X} + \frac{1}{2}m\omega^2 x^2 \right],$$

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

$$\psi_{n_x, n_y, n_z}(\mathbf{r}) = u_{n_x}(x)u_{n_y}(y)u_{n_z}(z).$$

n	n_x	n_y	n_z	g_n
0	0	0	0	1
1	1	0	0	3
	0	1	0	
	0	0	1	
2	2	0	0	6
	0	2	0	
	0	0	2	
	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 ?? 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 n th 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 ??.

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 \mathbf{x} with momentum \mathbf{p} is

$$\mathbf{L} = \mathbf{x} \times \mathbf{p}.$$

As a vector equation this can be written as three scalar equations:

$$L_x = yp_z - zp_y, \quad L_y = zp_x - xp_z, \quad \text{and} \quad L_z = xp_y - yp_x.$$

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 \mathbf{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:

$$\hat{\mathbf{L}} = \hat{\mathbf{X}} \times \hat{\mathbf{P}}.$$

For a state $|\psi\rangle \in \mathcal{H}$ if we define $\hat{L}_k|\psi\rangle = |\varphi\rangle$ then

$$\langle \mathbf{r} | \hat{L}_k | \psi \rangle = \langle \mathbf{r} | \varphi \rangle = \varphi(\mathbf{r}) = \hat{L}_k \psi(\mathbf{r}) = -\hbar \varepsilon_{ijk} x_j \frac{\partial \psi}{\partial x_k}.$$

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 momentum 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

$$\hat{L}_y = -i\hbar \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right), \quad \text{and} \quad \hat{L}_z = -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right).$$

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:

$$\begin{aligned} \hat{L}_x &= i\hbar \left[\sin \varphi \frac{\partial}{\partial \vartheta} + \cot \vartheta \cos \varphi \frac{\partial}{\partial \varphi} \right], \\ \hat{L}_y &= i\hbar \left[-\cos \varphi \frac{\partial}{\partial \vartheta} + \cot \vartheta \sin \varphi \frac{\partial}{\partial \varphi} \right], \\ \hat{L}_z &= -i\hbar \frac{\partial}{\partial \varphi}. \end{aligned}$$

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

$$\hat{L}^2 = -\hbar^2 \left[\frac{1}{\sin \vartheta} \frac{\partial}{\partial \vartheta} \left(\sin \vartheta \frac{\partial}{\partial \vartheta} \right) + \frac{1}{\sin^2 \vartheta} \frac{\partial^2}{\partial \varphi^2} \right]. \quad (12.1)$$

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 discuss 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 $\ell = 0, 1, 2, \dots$ is the **angular momentum quantum number** and $m = -\ell, -\ell+1, \dots, \ell-1, \ell$ is the **magnetic quantum number**. P_ℓ^m are the **associated Legendre polynomials**. The first four polynomials for $\ell = 0, 1$ are

$$P_0^0(\cos \vartheta) = 1, \quad P_1^0(\cos \vartheta) = \cos \vartheta, \quad P_1^1(\cos \vartheta) = -\sin \vartheta, \quad \text{and} \quad P_1^{-1}(\cos \vartheta) = \frac{1}{2} \sin \vartheta.$$

The first four spherical harmonics are then

$$\begin{aligned} Y_0^0(\vartheta, \varphi) &= \sqrt{\frac{1}{4\pi}}, & Y_1^0(\vartheta, \varphi) &= \sqrt{\frac{3}{4\pi}} \cos \vartheta \\ Y_1^1(\vartheta, \varphi) &= -\sqrt{\frac{3}{8\pi}} \sin \vartheta e^{i\varphi} & Y_1^{-1}(\vartheta, \varphi) &= \sqrt{\frac{3}{8\pi}} \sin \vartheta e^{-i\varphi}. \end{aligned}$$

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

$$\hat{L}_z Y_\ell^m(\vartheta, \varphi) = m\hbar Y_\ell^m(\vartheta, \varphi).$$

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 $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. 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:

$$[\hat{J}_k, \hat{J}_l] = i\hbar \varepsilon_{klm} \hat{J}_m, \quad \text{and} \quad [\hat{J}^2, \hat{J}_k] = 0.$$

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

$$\begin{aligned} \hat{J}_z \hat{J}_\pm |\lambda, m\rangle &= (\hat{J}_\pm \hat{J}_z - [\hat{J}_\pm, \hat{J}_z]) |\lambda, m\rangle \\ &= \hat{J}_\pm \hat{J}_z |\lambda, m\rangle + [\hat{J}_z, \hat{J}_\pm] |\lambda, m\rangle \\ &= \hbar m \hat{J}_\pm |\lambda, m\rangle \pm \hbar \hat{J}_\pm |\lambda, m\rangle \\ &= \hbar(m \pm 1) \hat{J}_\pm |\lambda, m\rangle. \end{aligned}$$

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

$$\begin{aligned}\langle \lambda, m | (\hat{J}^2 - \hat{J}_z^2) | \lambda, m \rangle &= \langle \lambda, m | \hat{J}^2 | \lambda, m \rangle - \langle \lambda, m | \hat{J}_z^2 | \lambda, m \rangle \\ &= \hbar^2 \lambda - \hbar^2 m^2,\end{aligned}$$

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

$$\hat{J}_+ |\lambda, m_{\max}\rangle = 0, \quad \text{and} \quad \hat{J}_- |\lambda, m_{\min}\rangle = 0.$$

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:

$$\begin{aligned}\hat{J}_+ \hat{J}_- &= (\hat{J}_x + i\hat{J}_y)(\hat{J}_x - i\hat{J}_y) \\ &= \hat{J}_x^2 + \hat{J}_y^2 + i\hat{J}_y \hat{J}_x - i\hat{J}_x \hat{J}_y \\ &= \hat{J}_x^2 + \hat{J}_y^2 + i[\hat{J}_y, \hat{J}_x] \\ &= \hat{J}_x^2 + \hat{J}_y^2 - i[\hat{J}_x, \hat{J}_y] \\ &= \hat{J}_x^2 + \hat{J}_y^2 + \hbar \hat{J}_z \\ &= \hat{J}^2 - \hat{J}_z^2 + \hbar \hat{J}_z,\end{aligned}$$

and

$$\begin{aligned}\hat{J}_- \hat{J}_+ &= (\hat{J}_x - i\hat{J}_y)(\hat{J}_x + i\hat{J}_y) \\ &= \hat{J}_x^2 + \hat{J}_y^2 - i\hat{J}_y \hat{J}_x + i\hat{J}_x \hat{J}_y \\ &= \hat{J}_x^2 + \hat{J}_y^2 + i[\hat{J}_x, \hat{J}_y] \\ &= \hat{J}_x^2 + \hat{J}_y^2 - \hbar \hat{J}_z \\ &= \hat{J}^2 - \hat{J}_z^2 - \hbar \hat{J}_z,\end{aligned}$$

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

$$\begin{aligned}\hat{J}_+ \hat{J}_- |\lambda, m_{\min}\rangle &= (\hat{J}^2 - \hat{J}_z^2 + \hbar \hat{J}_z) |\lambda, m_{\min}\rangle \\ &= \hat{J}^2 |\lambda, m_{\min}\rangle - \hat{J}_z^2 |\lambda, m_{\min}\rangle + \hbar \hat{J}_z |\lambda, m_{\min}\rangle \\ &= \hbar^2 \lambda |\lambda, m_{\min}\rangle - \hbar^2 m_{\min}^2 |\lambda, m_{\min}\rangle + \hbar^2 m_{\min} |\lambda, m_{\min}\rangle.\end{aligned}$$

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

$$\begin{aligned}\hat{J}_- \hat{J}_+ |\lambda, m_{\max}\rangle &= (\hat{J}^2 - \hat{J}_z^2 - \hbar \hat{J}_z) |\lambda, m_{\max}\rangle \\ &= \hat{J}^2 |\lambda, m_{\max}\rangle - \hat{J}_z^2 |\lambda, m_{\max}\rangle - \hbar \hat{J}_z |\lambda, m_{\max}\rangle \\ &= \hbar^2 \lambda - \hbar^2 m_{\max}^2 - \hbar^2 m_{\max}\end{aligned}$$

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

$$m_{\max} - m_{\min} = k, \quad k = 0, 1, 2, \dots$$

so

$$j - (-j) = 2j = k \implies j = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$$

For a given value of j we see that m ranges over

$$m = -j, -j+1, \dots, j-1, j.$$

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

$$\begin{aligned}\langle j, m | \hat{J}_\pm^\dagger \hat{J}_\pm | j, m \rangle &= \langle j, m | \hat{J}_\mp \hat{J}_\pm | j, m \rangle \\ &= \langle j, m | (\hat{J}^2 - \hat{J}_z^2 \mp \hbar \hat{J}_z) | j, m \rangle \\ &= \hbar^2 j(j+1) - \hbar^2 m^2 \mp \hbar^2 m \\ &= \hbar^2 [j(j+1) - m(m \pm 1)].\end{aligned}$$

Now consider an alternate way of calculating this:

$$\langle j, m | \hat{J}_\pm^\dagger \hat{J}_\pm | j, m \rangle = \hbar c_\pm \langle j, m | \hat{J}_\pm^\dagger | j, m \pm 1 \rangle$$

$$\begin{aligned}
&= \hbar c_{\pm} \langle j, m \pm 1 | \hat{J}_{\pm} | j, m \rangle^* \\
&= \hbar^2 c_{\pm} c_{\pm}^* \langle j, m \pm 1 | j, m \pm 1 \rangle \\
&= \hbar^2 |c_{\pm}|.
\end{aligned}$$

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:

$$V(r) = -\frac{Ze^2}{4\pi\epsilon_0 r}.$$

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, $V: [0, \infty) \rightarrow \mathbb{R}$:

$$\left[-\frac{\hbar^2}{2\mu} \nabla^2 + V(r) \right] u(r, \vartheta, \varphi) = E u(r, \vartheta, \varphi).$$

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

$$\begin{aligned}
E u(r, \vartheta, \varphi) &= \left[-\frac{\hbar^2}{2\mu} \left(\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{1}{\hbar^2 r^2} \hat{L}^2 \right) + V(r) \right] u(r, \vartheta, \varphi) \\
&= \left[-\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{2\mu r^2} \hat{L}^2 + V(r) \right] u(r, \vartheta, \varphi).
\end{aligned}$$

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 $R: [0, \infty) \rightarrow \mathbb{C}$ 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

$$\left[-\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + V(r) + \frac{\hbar^2 \ell(\ell+1)}{2\mu r^2} \right] R(r) Y_{\ell}^m(\vartheta, \varphi) = E R(r) Y_{\ell}^m(\vartheta, \varphi).$$

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

$$\left[-\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + V(r) + \frac{\hbar^2 \ell(\ell+1)}{2\mu r^2} \right] R_\ell(r) = E R_\ell.$$

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

$$E \chi_\ell(r) = \left[-\frac{\hbar^2}{2\mu} \frac{\partial^2}{\partial r^2} + V_{\text{eff}}(r) \right] \chi_\ell(r) \quad (13.1)$$

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

¹¹strictly we just require that $\chi(r) \rightarrow 0$ faster than $1/r \rightarrow 0$ as $r \rightarrow 0$.

13.4 Behaviour Near the Origin

Suppose that the potential is such that

$$\lim_{r \rightarrow 0} (V(r)r^2) = 0,$$

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

$$\begin{aligned} Er^{s+1} &= -\frac{\hbar^2}{2\mu} \frac{\partial}{\partial r} r^{s+1} + \frac{\hbar^2}{2\mu r^2} \ell(\ell+1) r^{s+1} \\ &= -\frac{\hbar^2}{2\mu} s(s+1) r^{s-1} + \frac{\hbar^2}{2\mu} \ell(\ell+1) r^{s-1}. \end{aligned}$$

Since we are concerned with behaviour as $r \rightarrow 0$ 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 $r \rightarrow 0$ 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

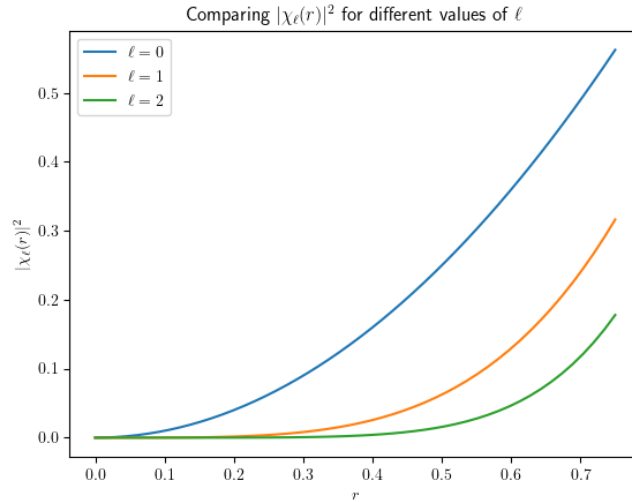


Figure 13.1: $|\chi_\ell(r)|^2$ compared for $\ell = 0, 1, 2$ and $r < 1$.

less likely the particle is to be near the origin. This can be seen 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, \mathbf{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, \mathbf{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

$$\int_a^b dr |\chi_\ell(r)|^2.$$

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

$$\hat{\mathbf{S}} = (\hat{S}_1, \hat{S}_2, \hat{S}_3),$$

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:

$$\left\{ |s, m\rangle \mid m = -s, 1-s, \dots, s-1, s \text{ and } s = \frac{n}{2} \text{ where } n \in \mathbb{N} \right\}.$$

The action of \hat{S}^2 and \hat{S}_z on these basis vectors is

$$\hat{S}^2 |s, m\rangle = s(s+1)\hbar^2 |s, m\rangle, \quad \text{and} \quad \hat{S}_z |s, m\rangle = m\hbar |s, m\rangle.$$

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

$$\left\{ \left| \frac{1}{2}, \frac{1}{2} \right\rangle, \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \right\}.$$

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 :

$$\begin{aligned} \hat{S}^2 |\uparrow\rangle &= \frac{3}{4}\hbar^2 |\uparrow\rangle, & \hat{S}^2 |\downarrow\rangle &= \frac{3}{4}\hbar^2 |\downarrow\rangle, \\ \hat{S}_z |\uparrow\rangle &= \frac{\hbar}{2}, & \text{and} & \quad \hat{S}_z |\downarrow\rangle = -\frac{\hbar}{2} |\downarrow\rangle. \end{aligned}$$

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

$$\left\{ |\mathbf{r}, m\rangle \mid \mathbf{r} \in \mathbb{R}^3 \text{ and } m = \pm \frac{1}{2} \right\}.$$

The vectors in this basis are simply the tensor products of vectors of the individual bases of $\mathcal{H}_{\text{space}}$ and $\mathcal{H}_{\text{spin}}$:

$$|\mathbf{r}, m\rangle = |\mathbf{r}\rangle |m\rangle = |\mathbf{r}\rangle \otimes |m\rangle.$$

This basis is complete, meaning

$$\int d^3 r' \sum_m |\mathbf{r}', m\rangle \langle \mathbf{r}', m| = 1.$$

A generic state, $|\psi\rangle \in \mathcal{H}$, is then given by

$$\begin{aligned} |\psi\rangle &= \int d^3 r \sum_m |\mathbf{r}', m\rangle \langle \mathbf{r}', m|\psi\rangle \\ &= \int d^3 r' \sum_m \psi_m(\mathbf{r}') |\mathbf{r}', m\rangle \\ &= \int d^3 r' [\psi_1(\mathbf{r}') |\uparrow\rangle + \psi_2(\mathbf{r}') |\downarrow\rangle] \end{aligned}$$

where $\psi_m(\mathbf{r}') = \langle \mathbf{r}', m|\psi\rangle$. Taking the inner product with a position eigenstate, $|\mathbf{r}\rangle$, we have

$$\begin{aligned} \langle \mathbf{r}|\psi\rangle &= \int d^3 r' \sum_m \langle \mathbf{r}|\mathbf{r}'\rangle |m\rangle \psi_m(\mathbf{r}') \\ &= \int d^3 r' \sum_m \delta(\mathbf{r} - \mathbf{r}') |m\rangle \psi_m(\mathbf{r}') \\ &= \sum_m \int d^3 r' \delta(\mathbf{r} - \mathbf{r}') |m\rangle \psi_m(\mathbf{r}') \\ &= \sum_m \psi_m(\mathbf{r}) \\ &= \psi_1(\mathbf{r}) |\uparrow\rangle + \psi_2(\mathbf{r}) |\downarrow\rangle. \end{aligned}$$

In the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$ for $\mathcal{H}_{\text{spin}}$ we can denote $\langle \mathbf{r}|\psi\rangle$ as a column vector:

$$\langle \mathbf{r}|\psi\rangle \rightarrow \begin{pmatrix} \psi_1(\mathbf{r}) \\ \psi_2(\mathbf{r}) \end{pmatrix}.$$

The interpretation of these values is as follows:

- $|\psi_1(\mathbf{r})|^2$ – the PDF of finding the system at \mathbf{r} with spin up.
- $|\psi_2(\mathbf{r})|^2$ – the PDF of finding the system at \mathbf{r} with spin down.
- The probability of finding the system anywhere in space with spin up is

$$P(\uparrow) = \int d^3 r |\psi_1(\mathbf{r})|^2.$$

- The probability of finding the system anywhere in space with spin down is

$$P(\downarrow) = \int d^3 r |\psi_2(\mathbf{r})|^2.$$

- The probability of finding the system at \mathbf{r} with any spin is

$$P(\mathbf{r}) = |\psi_1(\mathbf{r})|^2 + |\psi_2(\mathbf{r})|^2.$$

14.1 Matrix Elements

$\mathcal{H}_{\text{spin}}$ 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

$$\hat{S}_z|\uparrow\rangle = \frac{\hbar}{2}|\uparrow\rangle, \quad \text{and} \quad \hat{S}_z|\downarrow\rangle = -\frac{\hbar}{2}|\downarrow\rangle.$$

Hence

$$\begin{aligned} \langle\uparrow|\hat{S}_z|\uparrow\rangle &= \frac{\hbar}{2}, \\ \langle\uparrow|\hat{S}_z|\downarrow\rangle &= 0, \\ \langle\downarrow|\hat{S}_z|\uparrow\rangle &= 0, \end{aligned}$$

and

$$\langle\downarrow|\hat{S}_z|\downarrow\rangle = -\frac{\hbar}{2}.$$

This can be written compactly as

$$\langle m'|\hat{S}_z|m\rangle = \text{sign}(m)\delta_{m'm}\frac{\hbar}{2}.$$

Thus in this basis \hat{S}_z is represented by

$$\hat{S}_z \rightarrow \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

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

$$\hat{S}_+ = \hat{S}_x + i\hat{S}_y, \quad \text{and} \quad \hat{S}_- = \hat{S}_x - i\hat{S}_y.$$

These have the expected action that

$$\begin{aligned} \hat{S}_+|\uparrow\rangle &= 0, & \hat{S}_+|\downarrow\rangle &= \hbar c_+|\uparrow\rangle = \hbar\sqrt{\frac{1}{2}\left(\frac{1}{2}+1\right) + \frac{1}{2}\left(-\frac{1}{2}+1\right)}|\uparrow\rangle = \hbar|\uparrow\rangle \\ \hat{S}_-|\downarrow\rangle &= 0, & \hat{S}_-|\uparrow\rangle &= \hbar c_-|\downarrow\rangle = \hbar\sqrt{\frac{1}{2}\left(\frac{1}{2}+1\right) - \frac{1}{2}\left(\frac{1}{2}+1\right)}|\downarrow\rangle = \hbar|\downarrow\rangle. \end{aligned}$$

Hence

$$\begin{aligned} \langle\uparrow|\hat{S}_+|\uparrow\rangle &= 0, & \langle\uparrow|\hat{S}_-|\uparrow\rangle &= 0, \\ \langle\uparrow|\hat{S}_+|\downarrow\rangle &= \hbar, & \langle\uparrow|\hat{S}_-|\downarrow\rangle &= 0, \\ \langle\downarrow|\hat{S}_+|\uparrow\rangle &= 0, & \langle\downarrow|\hat{S}_-|\uparrow\rangle &= \hbar, \\ \langle\downarrow|\hat{S}_+|\downarrow\rangle &= 0, & \langle\downarrow|\hat{S}_-|\downarrow\rangle &= 0. \end{aligned}$$

Thus in this basis these operators are represented by

$$\hat{S}_+ \rightarrow \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \hat{S}_- \rightarrow \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Now we can write

$$\hat{S}_x = \frac{1}{2}(\hat{S}_+ + \hat{S}_-), \quad \text{and} \quad \hat{S}_y = \frac{1}{2i}(\hat{S}_+ - \hat{S}_-).$$

Thus the representations of \hat{S}_x and \hat{S}_y in this basis are

$$\hat{S}_x \rightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \hat{S}_y \rightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

The three matrices representing \hat{S}_k in this basis are called the **Pauli spin matrices**:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Which allows us to compactly write

$$\hat{S}_k \rightarrow \frac{\hbar}{2} \sigma_k, \quad \text{and} \quad \hat{\mathbf{S}} = \frac{\hbar}{2} \boldsymbol{\sigma}$$

where $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$. The Pauli matrices are Hermitian, as they are associated with observables, and also unitary, meaning $\sigma_k \sigma_k^\dagger = \sigma_k^2 = 1$, they also all have zero trace, $\text{tr}(\sigma_k) = 0$. 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

$$|\varphi\rangle = \hat{S}_k |\sigma\rangle = \sum_{\sigma} \varphi_{\sigma} |\sigma\rangle = \sum_{\sigma} \sum_{\sigma'} (S_k)_{\sigma\sigma'} \psi_{\sigma'} |\sigma\rangle$$

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:

$$\begin{aligned} \hat{S}^2 &= \sum_k \hat{S}_k^2 \\ &\rightarrow \sum_k \left(\frac{\hbar}{2} \sigma_k \right)^2 \\ &= \sum_k \frac{\hbar^2}{4} \sigma_k^2 \\ &= \sum_k \frac{\hbar^2}{4} 1 \\ &= \frac{3}{4} \hbar^2 1 \\ &= \frac{3}{4} \hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

14.2 Eigenvectors of \hat{S}_z and \hat{S}^2

The eigenvectors of σ_z are easy to find:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = - \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

So the eigenvalues of this representation of \hat{S}_z are given by

$$\frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{\hbar}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

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:

$$\frac{3}{4} \hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{3}{4} \hbar^2 \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad \frac{3}{4} \hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{3}{4} \hbar^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Thus we can identify

$$|s = \frac{1}{2}, m = \frac{1}{2}\rangle = |\uparrow\rangle \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \text{and} \quad |s = \frac{1}{2}, m = -\frac{1}{2}\rangle = |\downarrow\rangle \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

A generic vector, $|\psi\rangle \in \mathcal{H}_{\text{spin}}$, can then be represented by

$$|\psi\rangle = \psi_1|\uparrow\rangle + \psi_2|\downarrow\rangle \rightarrow \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \psi_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \psi_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

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

$$|\psi\rangle \rightarrow \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \implies \langle\psi| \rightarrow (\psi_1^* \quad \psi_2^*).$$

That is $\langle\psi| = (|\psi\rangle)^\dagger$. The scalar product of two states, $|\psi\rangle$ and $|\varphi\rangle$, is then

$$\langle\varphi|\psi\rangle = (\varphi_1^* \quad \varphi_2^*) \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \varphi_1^* \psi_1 + \varphi_2^* \psi_2.$$

For example if $|\psi\rangle$ is a properly normalised state then

$$\langle\varphi|\psi\rangle = (\psi_1^* \quad \psi_2^*) \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = |\psi_1|^2 + |\psi_2|^2 = 1.$$

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:

$$\begin{aligned} 0 &= \det \left(\frac{\hbar}{2} \sigma_x - \lambda 1 \right) \\ &= \begin{vmatrix} -\lambda & \hbar/2 \\ \hbar/2 & -\lambda \end{vmatrix} \\ &= \lambda^2 - \frac{\hbar^2}{4} \end{aligned}$$

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

$$\hat{S}_x |m_x = 1/2\rangle = \frac{\hbar}{2} |m_x = 1/2\rangle \rightarrow \frac{\hbar}{2} \sigma_x \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} \beta \\ \alpha \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

Hence $\alpha = \beta$ so a properly normalised state with $m_x = \hbar/2$ is

$$|m_x = 1/2\rangle = \frac{\sqrt{2}}{2} [|\uparrow\rangle + |\downarrow\rangle].$$

We can show in a similar way that if $S_x = -1/2$ then

$$|m_x = -1/2\rangle = \frac{\sqrt{2}}{2} [|\uparrow\rangle - |\downarrow\rangle].$$

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,

$$\hat{\mathbf{n}} = \sin(\vartheta)\mathbf{e}_x + \cos(\vartheta)\mathbf{e}_z,$$

then the relevant eigenstates in the representation that we are using are the eigenvalues of

$$\boldsymbol{\sigma} \cdot \hat{\mathbf{n}} = \sigma_x \sin(\vartheta) + \sigma_z \cos(\vartheta) = \begin{pmatrix} \cos \vartheta & \sin \vartheta \\ \sin \vartheta & -\cos \vartheta \end{pmatrix}.$$

15 Addition of Angular Momenta

Suppose we have two systems with independent angular momenta $\hat{\mathbf{J}}^{(1)}$ and $\hat{\mathbf{J}}^{(2)}$. 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,

$$[\hat{J}_k^{(n)}, \hat{J}_l^{(n)}] = i\hbar \varepsilon_{klm} \hat{J}_m^{(n)}, \quad \text{and} \quad [\hat{J}^{(n)2}, \hat{J}_k^{(n)}] = 0.$$

Since the two angular momenta are independent the operators associated with different angular momenta commute:

$$[\hat{J}_k^{(1)}, \hat{J}_l^{(2)}] = [\hat{J}^{(1)2}, \hat{J}^{(2)2}] = [\hat{J}^{(1)2}, \hat{J}_k^{(2)}] = 0 \quad \text{etc.}$$

We can extend this easily to any number of independent angular momenta but we won't here.

15.1 Vector Spaces

Since $\hat{\mathbf{J}}^{(1)}$ and $\hat{\mathbf{J}}^{(2)}$ 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 $\hat{J}_z^{(1)'} = \hat{J}_z^{(1)} \otimes 1$ 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

$$\begin{aligned} \hat{J}_z^{(1)'} |j_1, m_1, j_2, m_2\rangle &= (\hat{J}_z^{(1)} \otimes 1)(|j_1, m_1\rangle \otimes |j_2, m_2\rangle) \\ &= \hat{J}_z^{(1)} |j_1, m_1\rangle \otimes 1 |j_2, m_2\rangle \\ &= \hbar m_1 |j_1, m_1\rangle \otimes |j_2, m_2\rangle \\ &= \hbar m_1 |j_1, m_1, j_2, m_2\rangle \end{aligned}$$

but in practice we use the same symbol for both operators, $\hat{J}_z^{(1)}: \mathcal{H}^{(1)} \rightarrow \mathbb{C}$ and $\hat{J}_z^{(1)'}: \mathcal{H} \rightarrow \mathbb{C}$. The action of the operators on this basis is

$$\begin{aligned} \hat{J}^{(1)2} |j_1, m_1, j_2, m_2\rangle &= \hbar^2 j_1(j_1 + 1) |j_1, m_1, j_2, m_2\rangle, \\ \hat{J}_z^{(1)} |j_1, m_1, j_2, m_2\rangle &= \hbar m_1 |j_1, m_1, j_2, m_2\rangle, \\ \hat{J}^{(2)2} |j_1, m_1, j_2, m_2\rangle &= \hbar^2 j_2(j_2 + 1) |j_1, m_1, j_2, m_2\rangle, \\ \hat{J}_z^{(2)} |j_1, m_1, j_2, m_2\rangle &= \hbar m_2 |j_1, m_1, j_2, m_2\rangle. \end{aligned}$$

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:

$$\hat{\mathbf{J}} = \hat{\mathbf{J}}^{(1)} + \hat{\mathbf{J}}^{(2)}.$$

This satisfies the normal angular momentum commutation relations:

$$[\hat{J}_k, \hat{J}_l] = i\hbar \varepsilon_{klm} \hat{J}_m, \quad \text{and} \quad [\hat{J}^2, \hat{J}_k] = 0.$$

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 15.1 (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*

$$j = j_1 + j_2, j_1 + j_2 - 1, \dots, |j_1 - j_2|.$$

For each of these values we also have

$$m = -j, -j + 1, \dots, j - 1, j.$$

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

$$\begin{aligned} \dim(\mathcal{H}) &= \sum_{j=j_1-j_2}^{j_1+j_2} (2j+1) \\ &= \sum_{j=0}^{j_1+j_2} (2j+1) - \sum_{j=0}^{j_1-j_2-1} (2j+1) \\ &= 2 \sum_{j=0}^{j_1+j_2} j + (j_1+j_2) - 2 \sum_{j=0}^{j_1-j_2-1} j - (j_1-j_2-1) \\ &= (j_1+j_2)(j_1+j_2+1) + (j_1+j_2) - (j_1-j_2-1)(j_1-j_2) - (j_1-j_2-1) \\ &= j_1^2 + j_2^2 + 2j_1j_2 + j_1 + j_2 + j_1 + j_2 - j_1^2 - j_2^2 + 2j_1j_2 + j_1 - j_2 - j_1 + j_2 + 1 \\ &= 4j_1j_2 + 2j_1 + 2j_2 + 1 \\ &= (2j_1+1)(2j_2+1). \end{aligned}$$

Where we have used

$$\sum_{r=0}^n r = \frac{1}{2}n(n+1),$$

which is proven in appendix ???. 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,

$$\sum_{m_1, m_2} \langle j_1, m_1, j_2, m_2 | j_1, m_1, j_2, m_2 \rangle = 1$$

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

$$\{s_1, m_1\} = \{|\frac{1}{2}, \frac{1}{2}\rangle, |\frac{1}{2}, -\frac{1}{2}\rangle\}.$$

$\mathcal{H}^{(2)}$ is also a two dimensional vector space with basis

$$\{s_2, m_2\} = \{|\frac{1}{2}, \frac{1}{2}\rangle, |\frac{1}{2}, -\frac{1}{2}\rangle\}.$$

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

$$\{|s_1, m_1, s_2, m_2\rangle\} = \{|\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\rangle, |\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\rangle, |\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\rangle, |\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\rangle\}.$$

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

$$|1, 1\rangle = |\uparrow\uparrow\rangle, \quad \text{and} \quad |1, -1\rangle = |\downarrow\downarrow\rangle.$$

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

$$|1, 0\rangle = \frac{\sqrt{2}}{2} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle).$$

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 choose 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.

<table> <tr> <td>j</td> <td>1</td> </tr> <tr> <td>m_1, m_2</td> <td></td> </tr> <tr> <td>$1/2, 1/2$</td> <td>1</td> </tr> </table>	j	1	m_1, m_2		$1/2, 1/2$	1	<table> <tr> <td>j</td> <td>1</td> </tr> <tr> <td>m_1, m_2</td> <td></td> </tr> <tr> <td>$-1/2, -1/2$</td> <td>1</td> </tr> </table>	j	1	m_1, m_2		$-1/2, -1/2$	1	<table> <tr> <td>j</td> <td>1</td> <td>0</td> </tr> <tr> <td>m_1, m_2</td> <td></td> <td></td> </tr> <tr> <td>$1/2, -1/2$</td> <td>$\sqrt{2}/2$</td> <td>$\sqrt{2}/2$</td> </tr> <tr> <td>$-1/2, 1/2$</td> <td>$\sqrt{2}/2$</td> <td>$-\sqrt{2}/2$</td> </tr> </table>	j	1	0	m_1, m_2			$1/2, -1/2$	$\sqrt{2}/2$	$\sqrt{2}/2$	$-1/2, 1/2$	$\sqrt{2}/2$	$-\sqrt{2}/2$
j	1																									
m_1, m_2																										
$1/2, 1/2$	1																									
j	1																									
m_1, m_2																										
$-1/2, -1/2$	1																									
j	1	0																								
m_1, m_2																										
$1/2, -1/2$	$\sqrt{2}/2$	$\sqrt{2}/2$																								
$-1/2, 1/2$	$\sqrt{2}/2$	$-\sqrt{2}/2$																								
(a) $m = 1$	(b) $m = -1$	(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, $\hat{\mathcal{P}}_{12}$ which swaps particle 1 and 2. That is

$$\hat{\mathcal{P}}_{12}|\xi_1, \xi_2\rangle = |\xi_2, \xi_1\rangle.$$

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 16.1 (Spin statistics theorem). *Particles with integer spin, $s = 0, 1, 2, \dots$, 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, $s = 1/2, 3/2, 5/2, \dots$, 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

$$\hat{H}_i = \frac{\hat{\mathbf{P}} \cdot \hat{\mathbf{P}}}{2\mu} - \frac{2e^2}{r_i}$$

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:

$$\hat{H} = \hat{H}_1 + \hat{H}_2 + \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}.$$

Notice that \hat{H} is symmetric under $\hat{\mathcal{P}}_{12}$. That is $\hat{H}(1, 2) = \hat{H}(2, 1)$ where $\hat{H}(1, 2)$ is \hat{H} as defined above and $\hat{H}(2, 1) = \hat{\mathcal{P}}_{12}\hat{H}(1, 2)$. 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 $\hat{\mathcal{P}}_{12}$ on these is

$$\hat{\mathcal{P}}_{12}|\pm\rangle = \pm|\pm\rangle.$$

So $|\pm\rangle$ are eigenstates of $\hat{\mathcal{P}}_{12}$, with eigenvalues ± 1 and \hat{H} , with eigenvalues E . Hence $[\hat{H}, \hat{\mathcal{P}}_{12}] = 0$. This means that the symmetry properties of the system are preserved in time since this means that $\hat{\mathcal{P}}_{12}$ 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

$$|0, 0\rangle = \frac{\sqrt{2}}{2}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle), \quad |1, 1\rangle = |\uparrow\uparrow\rangle, \quad |1, -1\rangle = |\downarrow\downarrow\rangle, \quad \text{and} \quad |1, 0\rangle = \frac{\sqrt{2}}{2}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle).$$

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 $|\psi_{12}\rangle \in \mathcal{H}_{\text{space}}$ 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 $\hat{\mathcal{P}}_{12}$ the action of swapping the two particles is given by

$$\begin{aligned} \hat{\mathcal{P}}_{12}|\Psi\rangle &= (\hat{\mathcal{P}}_{12}|\psi_{12}\rangle) \otimes (\hat{\mathcal{P}}_{12}|s = 1, m\rangle) \\ &= (\hat{\mathcal{P}}_{12}|\psi_{12}\rangle) \otimes |s = 1, m\rangle \\ &= -|\Psi\rangle. \end{aligned}$$

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 $\hat{\mathcal{P}}_{12}$. 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 $\hat{\mathcal{P}}_{12}$ but this time $|s = 0, m = 0\rangle$ is antisymmetric under $\hat{\mathcal{P}}_{12}$ so $|\psi_{12}\rangle$ must be symmetric under $\hat{\mathcal{P}}_{12}$. These two results are tabulated in table 16.1.

s	$ s, m\rangle$	$ \psi_{12}\rangle$	$ \Psi\rangle$
0	A	S	A
1	S	A	A

Table 16.1: The symmetry of various parts of a state under $\hat{\mathcal{P}}_{12}$. S denotes symmetric and A denotes antisymmetric. The symmetry of the total system, $|\Psi\rangle$, is set by the spin statistics theorem. The symmetry of the spin part, $|s, m\rangle$, follows from the expression in the uncoupled basis. The symmetry of the spatial part, $|\psi_{12}\rangle$, 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 $|\xi_1\rangle = |\xi_2\rangle$ then by the spin statistics theorem

$$|\xi_1, \xi_2\rangle = -|\xi_2, \xi_1\rangle = -|\xi_1, \xi_2\rangle$$

where for the first equality we use the spin statistics theorem and for the second the fact that $|\xi_1\rangle = |\xi_2\rangle$. The only solution to this equation is $|\xi_1, \xi_2\rangle = 0$ but this is not a valid state, it isn't normalisable for one. Therefore we cannot have $|\xi_1\rangle = |\xi_2\rangle$ 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

$$m_p = 1.7 \times 10^{-27} \text{ kg}, \quad m_e = 9.1 \times 10^{-31} \text{ kg}, \quad q = 1.6 \times 10^{-19} \text{ C}, \quad \text{and} \quad \frac{m_p}{m_e} = 1836.15.$$

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

$$V(r) = -\frac{q^2}{4\pi\epsilon_0 r} = -\frac{e^2}{r} \quad \text{where} \quad e^2 = \frac{q^2}{4\pi\epsilon_0}.$$

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

$$\hat{H} = \frac{\hat{\mathbf{P}} \cdot \hat{\mathbf{P}}}{2\mu} + V(r).$$

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:

$$\left[-\frac{\hbar^2}{2\mu} \nabla^2 - \frac{e^2}{r} \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}).$$

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

$$\left[-\frac{\hbar^2}{2\mu} \frac{d^2}{dr^2} + \frac{\hbar^2 \ell(\ell+1)}{2\mu r^2} - \frac{e^2}{r} \right] E \chi_\ell(r) = E_\ell \chi_\ell(r).$$

Figure 17.1 shows the potential. For $r \rightarrow 0$ the potential goes as r^{-2} and for $r \rightarrow \infty$ 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.

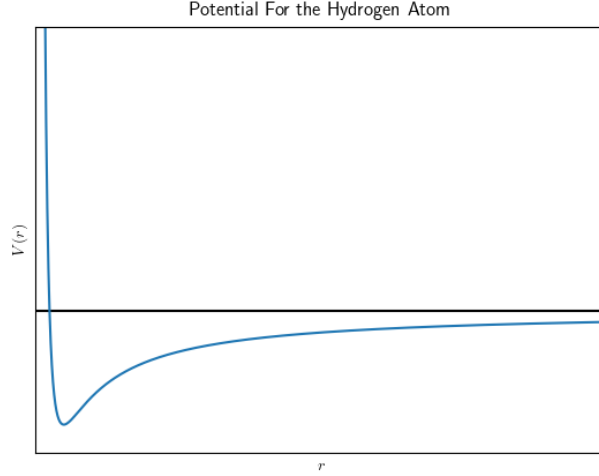


Figure 17.1: The potential for the hydrogen atom.

To tidy away even more constants we define

$$a_0 = \frac{\hbar^2}{\mu e^2} \approx 0.52 \text{ \AA}, \quad \text{and} \quad E_I = \frac{\hbar^2}{2\mu a_0^2} = \frac{\mu e^4}{2\hbar^2} \approx 13.6 \text{ eV}.$$

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:

$$\rho = \frac{r}{a_0}, \quad \text{and} \quad \lambda_\ell = \sqrt{-\frac{E_\ell}{E_I}}.$$

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

$$\begin{aligned} -\lambda_\ell^2 E_I \chi(a_0 \rho) &= -\lambda_\ell^2 E_I u_\ell(\rho) \\ &= \left[-\frac{\hbar^2}{2\mu} \frac{1}{a_0^2} \frac{d^2}{d\rho^2} + \frac{\hbar^2 \ell(\ell+1)}{2\mu a_0^2 \rho^2} - \frac{e^2}{a_0 \rho} \right] \chi_\ell(a_0 \rho) \\ &= \left[-\frac{\hbar^2}{2\mu} \frac{1}{a_0^2} \frac{d^2}{d\rho^2} + \frac{\hbar^2 \ell(\ell+1)}{2\mu a_0^2 \rho^2} - \frac{e^2}{a_0 \rho} \right] u_\ell(\rho). \end{aligned}$$

Which gives us

$$\left[-\frac{\hbar^2}{2\mu} \frac{1}{a_0^2} \frac{d^2}{d\rho^2} + \frac{\hbar^2 \ell(\ell+1)}{2\mu a_0^2 \rho^2} - \frac{e^2}{a_0 \rho} + \lambda_\ell^2 E_I \right] u_\ell(\rho) = 0.$$

Dividing through by $-\hbar^2/2\mu a_0^2$ we have

$$\left[\frac{d^2}{d\rho^2} - \frac{\ell(\ell+1)}{\rho^2} + \frac{2\mu a_0 e^2}{\hbar^2 \rho} - \lambda_\ell^2 \frac{2\mu a_0^2 E_I}{\hbar^2} \right] u_\ell(\rho) = 0.$$

Using the definitions of a_0 and E_I this becomes

$$\left[\frac{d^2}{d\rho^2} - \frac{\ell(\ell+1)}{\rho^2} + \frac{2}{\rho} - \lambda_\ell^2 \right] u_\ell(\rho) = 0.$$

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 $\rho \rightarrow \infty$, in this case the radial equation reduces to

$$\left[\frac{d^2}{d\rho^2} - \lambda^2 \right] u_\ell(\rho) = 0.$$

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

$$\begin{aligned} 0 &= \frac{d^2}{d\rho^2} [e^{-\lambda_\ell \rho} \eta_\ell(\rho)] + \left[\frac{2}{\rho} - \frac{\ell(\ell+1)}{\rho^2} - \lambda_\ell^2 \right] e^{-\lambda_\ell \rho} \eta_\ell(\rho) \\ &= \frac{d^2 \eta_\ell}{d\rho^2} e^{-\lambda_\ell \rho} + 2 \frac{d\eta_\ell}{d\rho} \frac{d}{d\rho} e^{-\lambda_\ell \rho} + \eta_\ell(\rho) \frac{d^2}{d\rho^2} e^{-\lambda_\ell \rho} + \left[\frac{2}{\rho} - \frac{\ell(\ell+1)}{\rho^2} - \lambda_\ell^2 \right] e^{-\lambda_\ell \rho} \eta_\ell(\rho) \\ &= \frac{d^2 \eta_\ell}{d\rho^2} e^{-\lambda_\ell \rho} - 2\lambda_\ell 2 \frac{d\eta_\ell}{d\rho} e^{-\lambda_\ell \rho} + \lambda_\ell^2 \eta_\ell(\rho) e^{-\lambda_\ell \rho} + \left[\frac{2}{\rho} - \frac{\ell(\ell+1)}{\rho^2} - \lambda_\ell^2 \right] e^{-\lambda_\ell \rho} \eta_\ell(\rho). \end{aligned}$$

Dividing through by $e^{-\lambda_\ell \rho}$ we have

$$\frac{d^2 \eta_\ell}{d\rho^2} - 2\lambda_\ell \frac{d\eta_\ell}{d\rho} + \left[\frac{2}{\rho} - \frac{\ell(\ell+1)}{\rho^2} \right] \eta_\ell(\rho) = 0. \quad (17.1)$$

Now as $\rho \rightarrow 0$ 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

$$\begin{aligned} 0 &= \sum_q [(q+\ell+1)(q+\ell)c_q \rho^{q+\ell-1} - 2\lambda_\ell(q+\ell+1)c_q \rho^{\ell+1} + 2c_q \rho^{q+\ell} - \ell(\ell+1)c_q \rho^{q+\ell-1}] \\ &= \sum_q q(q+2\ell+1)c_q \rho^{q+\ell-1} - \sum_q 2(\lambda_\ell[q+\ell+1]-1)c_q \rho^{q+\ell} \end{aligned}$$

relabelling, $q \rightarrow q-1$ in the second sum we have

$$= \sum_q [q(q+2\ell+1)c_q - 2[\lambda_\ell(q+\ell)-1]c_{q-1}] \rho^{q+\ell-1}.$$

Since we require this to hold for all ρ we have that for all q

$$q(q+2\ell+1)c_q - 2[\lambda_\ell(q+\ell)-1]c_{q-1} = 0. \quad (17.2)$$

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

$$\frac{c_q}{c_{q-1}} = \frac{2\lambda_\ell(q+1)-2}{q(q+2\ell+1)} \rightarrow \frac{2\lambda}{q}$$

as $q \rightarrow \infty$ 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

$$\alpha = \frac{e^2}{\hbar c} = \frac{q^2}{4\pi\epsilon_0 \hbar c} \approx \frac{1}{137}$$

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

$$E_n = -\frac{e^2}{2n^2 a_0} = -\frac{\mu e^4}{2n^2 \hbar^2} = -\frac{\mu q^4}{32\hbar^2 \pi^2 \epsilon_0^2 n^2}.$$

From $n = n_r + \ell$ since $n, n_r \in \mathbb{N}_{>0}$ we have that $\ell = 0, 1, \dots, n-1$, which corresponds to $n_r = n, n-1, \dots, 1$. 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 ???. 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

$$\begin{aligned} R_{1,0}(r) &= \frac{2}{a^{3/2}} e^{-r/a_0} \\ R_{2,0}(r) &= \frac{1}{2\sqrt{2}} \frac{1}{a^{3/2}} \left[2 - \frac{r}{a_0} \right] e^{-r/(2a_0)} \\ R_{2,1}(r) &= \frac{1}{2\sqrt{6}} \frac{1}{a^{3/2}} \frac{r}{a_0} e^{-r/(2a_0)}. \end{aligned}$$

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 $\psi = RY_\ell^m$ 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

$$\begin{aligned} \hat{H}_0|\psi_{0n}\rangle + \varepsilon(\hat{H}_0|\psi_{1n}\rangle + \hat{V}|\psi_{0n}\rangle) + \varepsilon^2(\hat{H}_0|\psi_{2n}\rangle + \hat{V}|\psi_{1n}\rangle) + \mathcal{O}(\varepsilon^3) = \\ E_{0n}|\psi_{0n}\rangle + \varepsilon(E_{0n}|\psi_{1n}\rangle + E_{1n}|\psi_{0n}\rangle) + \varepsilon^2(E_{0n}|\psi_{2n}\rangle + E_{1n}|\psi_{1n}\rangle + E_{2n}|\psi_{0n}\rangle) + \mathcal{O}(\varepsilon^3). \end{aligned}$$

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

$$(\hat{H} - E_{0n})|\psi_{1n}\rangle - (\hat{V} - E_{1n})|\psi_{0n}\rangle = 0. \quad (18.1)$$

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 n th and m th energy eigenvalues of \hat{H}_0 so are known quantities. Knowing this we can express the n th 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 n th 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

$$\Delta E_1 = \langle\varphi^{(1)}|\hat{H}'|\varphi^{(1)}\rangle$$

$$\begin{aligned}
&= \int_{-a}^a \varphi^{(1)}(x)^* V_0 \cos\left(\frac{\pi x}{2a}\right) \varphi^{(1)}(x) dx \\
&= \frac{V_0}{a} \int_{-a}^a \cos^3\left(\frac{\pi x}{2a}\right) dx \\
&= \frac{8V_0}{3\pi}.
\end{aligned}$$

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 1/2, 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

$$|s = 1, s_z = 1\rangle = |\uparrow\uparrow\rangle, \quad |s = 1, s_z = -1\rangle = |\downarrow\downarrow\rangle, \quad |s = 1, s_z = 0\rangle = \frac{\sqrt{2}}{2}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle),$$

$$\text{and} \quad |s = 0, s_z = 0\rangle = \frac{\sqrt{2}}{2}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle).$$

The spatial wave function is given by

$$\langle \mathbf{r}_1, \mathbf{r}_2 | \psi \rangle = \psi(\mathbf{r}_1, \mathbf{r}_2) = \psi^{(1)}(\mathbf{r}_1) \psi^{(2)}(\mathbf{r}_2),$$

where we assume in the last step that a separable solution exists. Each $\psi^{(k)}$ must satisfy the TISE for the relevant Hamiltonian:

$$\hat{H}_0^{(k)} \psi^{(k)}(\mathbf{r}_k) = \left[\frac{\hbar^2}{2m} \nabla_k^2 - \frac{Ze^2}{r_k} \right] \psi^{(k)}(\mathbf{r}_k) = E^{(k)} \psi^{(k)}(\mathbf{r}_k).$$

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

$$\psi^{(k)}(\mathbf{r}_k) = u_{100}(\mathbf{r}_k) = \frac{1}{\sqrt{\pi}} \left(\frac{Z}{a_0} \right)^{3/2} e^{-r/a_0}.$$

Thus

$$\langle \mathbf{r}_1, \mathbf{r}_2 | \Psi \rangle = u_{100}(\mathbf{r}_1) u_{100}(\mathbf{r}_2) \sum_{\alpha=1}^4 c_\alpha |\alpha\rangle$$

where

$$\{|\alpha\rangle\} = \{|s, s_z\rangle \mid s = 0, 1 \text{ and } s_z = -s, \dots, s\}.$$

Electrons are fermions. This means that $|\Psi\rangle$ should be antisymmetric under exchanging of 1 and 2. The radial part, $u_{100}(\mathbf{r}_1)u_{100}(\mathbf{r}_2)$, 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

$$\hat{V} = \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}.$$

The energy shift due to this perturbation is

$$\begin{aligned}\Delta E_1 &= \langle \Psi | \hat{V} | \Psi \rangle \\ &= \langle \psi; 0, 0 | \hat{V} | \psi; 0, 0 \rangle\end{aligned}$$

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, $\psi(\mathbf{r}) = u_{100}(\mathbf{r})$, or equivalently $|\psi\rangle = |100\rangle$, this becomes

$$\begin{aligned}&= \langle 100, 100 | \hat{V} | 100, 100 \rangle \\ &= \int d^3r_1 d^3r_2 u_{100}^*(\mathbf{r}_1) u_{100}^*(\mathbf{r}_2) \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} u_{100}(\mathbf{r}_1) u_{100}(\mathbf{r}_2) \\ &= \left(\frac{eZ^3}{\pi a_0^3} \right)^2 \int d^3r_1 d^3r_2 \exp\left(-\frac{2Z(r_1 + r_2)}{a_0}\right) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \\ &= \frac{5}{4} Z E_I\end{aligned}$$

where $E_I \approx 13.6 \text{ eV}$ is defined as it was in section 17.1.