Advanced Quantum Mechanics Notes

Francesco Chotuck

Abstract

This is KCL undergraduate module 6CCM436A, instructed by Neil Lambert. The formal name for this class is “Advanced Quantum Mechanics”.

Contents

1 IQT recap

Remark 1.1. In this module the notation: $(\phi, \widehat{\mathcal{O}}\Psi) = \left\langle \phi \middle| \widehat{\mathcal{O}}\Psi \right\rangle = \langle\phi| \widehat{\mathcal{O}} |\Psi\rangle$.

1.1 Hilbert spaces

Definition 1.2. A vector space is said to be **complete** if all the Cauchy sequences converge.

Definition 1.3. A vector space is said to be **separable** if for any two vectors there exists two disjoint open states such that each includes one vector but not the other.

Theorem 1.4

A vector space is separable if there exists a countable basis. In this module we assume this is ALWAYS the case.

Definition 1.5. A **Hermitian inner product** on V is a function $(\cdot,\cdot) : V \times V \to \mathbb{C}$ which satisfies the following properties

1. Positive definite:
$$(\boldsymbol{v}, \boldsymbol{v}) \geq 0 \quad \forall \boldsymbol{v} \in V$$
and $(\boldsymbol{v}, \boldsymbol{v}) \iff \boldsymbol{v} = 0$.
2. Hermiticity:
$$(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{v}, \boldsymbol{u})^* \quad \forall \boldsymbol{u}, \boldsymbol{v} \in V,$$
3. linearity in the second argument,
$$(\boldsymbol{u}, \alpha\boldsymbol{v} + \beta\boldsymbol{w}) = \alpha(\boldsymbol{u}, \boldsymbol{v}) + \beta(\boldsymbol{u}, \boldsymbol{w}) \quad \forall \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V \text{ and } \forall \alpha, \beta \in \mathbb{C},$$
4. anti-linearity in the first argument
$$(\alpha\boldsymbol{u} + \beta\boldsymbol{v}, \boldsymbol{w}) = \alpha^*(\boldsymbol{u}, \boldsymbol{w}) + \beta^*(\boldsymbol{v}, \boldsymbol{w}) \quad \forall \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V \text{ and } \forall \alpha, \beta \in \mathbb{C}.$$

Definition 1.6. A **Hilbert space** is a complete, separable complex vector space with positive definite inner product
$$\langle\psi_1|\psi_2\rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}.$$

Note 1.7. For the purpose of this module we can interpret a Hilbert space to be an infinite dimensional complex vector space which 'behaves' like $\mathbb{C}^N$ but with $N = \infty$.

Remark 1.8. If a Hilbert space is finitely dimensional then it is isomorphic to $\mathbb{C}^N$ for some N.

Example 1.9. A classic example of a Hilbert space is $L_2(\mathbb{R})$:

$$L_2(\mathbb{R}) = \left\{ \psi : \mathbb{R} \to \mathbb{C} : \int |\psi|^2 \, dx < \infty \right\}$$
$$\langle \psi_1 | \psi_2 \rangle = \int \psi_1^* \psi_2 \, dx.$$

Definition 1.10. The **state** of a system is given by an element $|\psi\rangle$ of unit norm in a Hilbert space $\mathcal{H}$.

Corollary 1.11

We will use the terms "state" and "wave function" interchangeably.

Definition 1.12. The space of linear maps from $\mathcal{H}$ to $\mathbb{C}$ is called the **dual space**, and is denoted by $\mathcal{H}^*$.

In Dirac notation elements of the dual space are denoted by "bras" $\langle \psi |$ such that

$$\langle \psi_1 | \left(|\psi_2\rangle \right) = \langle \psi_1 | \psi_2 \rangle .$$

Theorem 1.13

The dual space is a Hilbert space and is isomorphic to $\mathcal{H}$.

Note 1.14. This means that for each $|\psi\rangle$ there exists a unique $\langle \psi |$ and vice versa.

Theorem 1.15. There exists linear operators in Hilbert spaces i.e.

$$\mathcal{O} : \mathcal{H} \to \mathcal{H}$$

such that $\mathcal{O}\left(\lambda_1 |\psi_1\rangle + \lambda_2 |\psi_2\rangle\right) = \lambda_1 \mathcal{O} |\psi_1\rangle + \lambda_2 \mathcal{O} |\psi_2\rangle$ where $\lambda_1, \lambda_2 \in \mathbb{C}$.

Definition 1.16. The **adjoint** of a linear map $\mathcal{O}$ is a linear map $\mathcal{O}^\dagger : \mathcal{H} \to \mathcal{H}$ that satisfies

$$\langle \psi_1 | \mathcal{O} \psi_2 \rangle = \left\langle \mathcal{O}^\dagger \psi_1 | \psi_2 \right\rangle$$

for $|\psi_1\rangle, |\psi_2\rangle \in \mathcal{H}$.

Note 1.17. Equivalently,

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

1.2 Delta-normalisable states

Definition 1.18. The **Dirac delta function** is defined as

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

with the property that

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

> **Remark 1.19.** Although, it is called a 'function' it is not truly a function, rather a distribution.

Theorem 1.20

Properties of the δ-function:

- $\delta(x) \in \mathbb{R}$,
- $\delta(x) \geq 0$,
- $\delta(x) = \delta(-x)$ (even).
- $\int_{-\infty}^{\infty} \delta(k) f(k)\, dk = f(0)$ (sometimes we use the notation $\delta(f) = f(0)$);
- $\int_{-\infty}^{\infty} \delta(k-a) f(k)\, dk = f(a)$;
- $\int_{-\infty}^{\infty} e^{ikx}\, dx = 2\pi\delta(k)$;
- $\langle e_x | e_y \rangle = \delta(x-y)$.

Definition 1.21. We say that $|e_x\rangle$ and $|e_y\rangle$ are **delta-normalisable** if

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

Definition 1.22. We say that $|e_k\rangle$ for $k \in \mathbb{R}$ form a **delta-normalisable basis** of $\mathcal{H}$ if any $|\Psi\rangle \in \mathcal{H}$ can be represented in the form

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

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

1.3 Expectation values

Theorem 1.23

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

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

1.4 Observables

Definition 1.24. A linear operator is said to be **self-adjoint** (or **Hermitian**) if $\mathcal{O}^\dagger = \mathcal{O}$.

Definition 1.25. A linear operator which is self-adjoint is called an **observable**.

Theorem 1.26. Eigenstate of self-adjoint operators have real eigenvalues and eigenvectors with distinct eigenvalues are orthogonal.

Theorem 1.27

A self-adjoint operator admits an orthonormal basis of eigenvectors with real eigenvalues.

Definition 1.28. The linear **operator of position**, $\widehat{x}$, for a fixed time t is defined as

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

Definition 1.29. The linear **operator of momentum**, $\widehat{p}$, for a fixed time t is defined as

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

Theorem 1.30

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

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

Definition 1.31. The **Hamiltonian operator** is defined as

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

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

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

Definition 1.32. An operator, $\widehat{A}$, which satisfies the property

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

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

Definition 1.33. We call the operator $e^{-\frac{i\widehat{H}}{\hbar}t}$ the **time evolution operator** and is denoted by

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

Theorem 1.34. Properties of $\widehat{U}(t)$:

- $\widehat{U}(t_2)\widehat{U}(t_1) = \widehat{U}(t_1 + t_2)$.
- $\left(\widehat{U}(t)\right)^\dagger = \widehat{U}(-t)$.

Theorem 1.35. The time evolution operator, $\widehat{U}(t)$ is unitary.

Theorem 1.36

Unitary operators preserve the inner product i.e.

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

Theorem 1.37 (Properties of exponentials of operators)

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

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

1.4.1 Matrix elements

Proposition 1.38

The matrix elements of the operator $\widehat{O}$ in the chosen basis are given by the inner product of the basis states with the operator acting on the basis states, say $\left|\psi_m\right\rangle$:

$$O_{ij} = \left\langle\psi_i\right|\widehat{O}\left|\psi_j\right\rangle$$

where O_{ij} is the element of the matrix in the i-th row and j-th column.

Example 1.39

For a simple example, consider a 2-dimensional space with basis states $\left|\psi_1\right\rangle$ and $\left|\psi_2\right\rangle$. The matrix representation of an operator $\widehat{O}$ in this space would be a 2×2 matrix. The matrix representation of $\widehat{O}$ would be:

$$\begin{pmatrix} \left\langle\psi_1\right|\widehat{O}\left|\psi_1\right\rangle & \left\langle\psi_1\right|\widehat{O}\left|\psi_2\right\rangle \\ \left\langle\psi_2\right|\widehat{O}\left|\psi_1\right\rangle & \left\langle\psi_2\right|\widehat{O}\left|\psi_2\right\rangle \end{pmatrix}$$

Each element in this matrix represents the action of the operator $\widehat{O}$ on the basis states, and how that action projects onto the other basis states.

1.5 Heisenberg's uncertainty principle

Definition 1.40. The **uncertainty** of a normalised Hermitian operator $\widehat{A}$, denoted by $\Delta\widehat{A}$, is defined by

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

Note 1.41. Operators need to be Hermitian as these are the operators which are used to measure quantities in Quantum Mechanics.

Theorem 1.42. The uncertainty of an operator $\widehat{A}$ is real and non-negative i.e. $\Delta\widehat{A} \geq 0$.

Proof. We have

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

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

Corollary 1.43. Let $\widehat{A}$ be an operator then, $\Delta\widehat{A} = 0$ in a state $|\Psi\rangle$ if and only if $|\Psi\rangle$ is an eigenvector of $\widehat{A}$.

Proof. Recall the definition of uncertainty

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

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

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

We can write this condition as

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

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

Theorem 1.44

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

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

Example 1.45

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

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

1.6 The Schrödinger equation

Theorem 1.46

We obtain the wave function from solving the (time-dependent) **Schrödinger equation**:

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

where $U(x)$ is the potential energy function of the particle.

Theorem 1.47

The **time-independent** Schrödinger equation is given by

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

Equivalently, it can be written as

$$(\widehat{H}\psi)(x) = E\psi(x).$$

Note 1.48. Due to separation of variables the partial derivative above can be written as a total derivative and use standard techniques to solve a second order ODE.

Corollary 1.49

The general solution of the Schrödinger equation is given by

$$\Psi(x,t) = \psi(x)e^{-\frac{iE}{\hbar}t}.$$

1.7 Harmonic oscillator

Definition 1.50. A system with Hamiltonian operator

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

is called a **harmonic oscillator**.

It will be useful to consider the operators:

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

Proposition 1.51. The value of

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

Theorem 1.52

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

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

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

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

We can write

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

Proposition 1.53. Properties of $\widehat{N} = \widehat{a}^\dagger\widehat{a}$.

- $\widehat{N}$ is Hermitian i.e. $\widehat{N}^\dagger = \widehat{N}$;
- $\left|\Psi\right\rangle$ is an eigenvector of $\widehat{H}$ if and only if $\left|\Psi\right\rangle$ is an eigenvector of $\widehat{N}$.

Note 1.54. The eigenvectors are the same, but the eigenvalues are different.

Lemma 1.55. If λ is an eigenvalue of $\widehat{N}$ then $\lambda \geq 0$.

Lemma 1.56. Let $\left|\Psi\right\rangle$ be an eigenvector of $\widehat{N}$ with the eigenvalue λ. Then exactly one of the statements holds:

1. $\lambda = 0$ $\underline{\text{and}}$ $\widehat{a}\left|\Psi\right\rangle = 0$ OR,
2. $\lambda \neq 0$, $\widehat{a}\left|\Psi\right\rangle \neq 0$ and $\widehat{a}\left|\Psi\right\rangle$ is an eigenvector of $\widehat{N}$ with eigenvalue $\lambda - 1$.

Lemma 1.57. Let $\left|\Psi\right\rangle$ be an eigenvector of $\widehat{N}$ with eigenvalue λ. Then, $\lambda \in \mathbb{Z}_{\geq 0}$ and $\left|\Psi\right\rangle, \widehat{a}\left|\Psi\right\rangle,$
$\widehat{a}^2\left|\Psi\right\rangle, \ldots, \widehat{a}^\lambda\left|\Psi\right\rangle$ are eigenstates of $\widehat{N}$ with the eigenvalues $\lambda, \lambda-1, \lambda-2, \ldots, 0$ respectively. Finally, $\widehat{a}^{\lambda+1}\left|\Psi\right\rangle = 0$.

Lemma 1.58

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

Note 1.59. To remember that $\widehat{a}^\dagger$ increases the eigenvalue by 1 we can think of $\dagger$ as $+$.

Note 1.60. In conclusion, we see that we can lower the eigenvalue of $\widehat{N}$ by applying $\widehat{a}$ to the eigenstate, and similarly we can raise it by applying $\widehat{a}^\dagger$. For this reason $\widehat{a}$ and $\widehat{a}^\dagger$ are called the **lowering** and the **raising** operators.

Lemma 1.61

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

1. $\lambda \in \mathbb{Z}_{\geq 0}$,
2. there exists a state $|\Phi\rangle \equiv \widehat{a}^\lambda |\Psi\rangle \in \mathcal{H}$ which satisfies $\widehat{a}|\Phi\rangle = 0$ (in particular, it is an eigenstate of $\widehat{N}$ with eigenvalue 0),
3. we have $|\Psi\rangle = C(\widehat{a}^\dagger)^\lambda |\Phi\rangle$ for some $C \in \mathbb{C}$ where $C \neq 0$.

Note 1.62. This lemma allows us to undo the action of $\widehat{a}$ with $\widehat{a}^\dagger$ and vice versa. For example consider the state $\widehat{a}|\Psi\rangle$ with eigenvalue $\lambda - 1$, and then apply $\widehat{a}^\dagger$ to get back to the eigenvalue λ. We find the state

$$\begin{aligned}
\widehat{a}^\dagger (\widehat{a}|\Psi\rangle) &= \widehat{a}^\dagger \widehat{a}|\Psi\rangle \\
&= \widehat{N}|\Psi\rangle \\
&= \lambda|\Psi\rangle .
\end{aligned}$$

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

Example 1.63

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

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

which have eigenvalues 0 and 1 respectively. We have

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

Moreover,

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

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

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

Lemma 1.64

Let $|\Phi\rangle$ be such that

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

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

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

Note 1.65. This lemma tells us that there exists a unique normalised state $|\Phi\rangle$ which satisfies $\widehat{a} |\Phi\rangle = 0$.

Definition 1.66 (Notation)**.** We identify the state $|\Phi\rangle \equiv |0\rangle$ where $|\Phi\rangle$ is defined as above. The state $|0\rangle$ is called the **ground state** (or vacuum state).

Remark 1.67. $|0\rangle \neq 0$ and indeed $|0\rangle$ is an eigenvector.

Note 1.68. We call $|0\rangle$ the *ground state* because it has the lowest possible energy.

Theorem 1.69

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

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

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

Corollary 1.70. The state $|n\rangle$ is an eigenstate of $\widehat{N}$ with the eigenvalue n, i.e.

$$\widehat{N}|n\rangle = n|n\rangle .$$

Theorem 1.71. Let $|\Psi\rangle$ be an eigenstate of $\widehat{N}$ with the eigenvalue λ. Then, the following statements hold

1. $\lambda = n \in \mathbb{Z}_{\geq 0}$,
2. $|\Psi\rangle = C|n\rangle$ for some $C \in \mathbb{C}$ where $C \neq 0$.

Theorem 1.72

The eigenstates of $\widehat{H}$ are the states $|n\rangle$, and the corresponding eigenvalues are

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right).$$

The spectrum of $\widehat{H}$ is therefore,

$$\left\{\frac{\hbar\omega}{2}, \frac{3\hbar\omega}{2}, \frac{5\hbar\omega}{2}, \dots\right\}$$

2 Spherically symmetric potential

Note 2.1. We solve the Schrödinger equation as usual, where we do a seperation of varibales and set the two new equations equal to a constant.

In three-dimensions the Schrödinger equation takes the intimidating form

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + V(x, y, x)\Psi \tag{2.11}$$

Here we have assumed the potential is time-independent and

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \ . \tag{2.12}$$

This is in general too hard to solve but let us consider a potential with spherical symmetry $V = V(r)$ and $r = \sqrt{x^2 + y^2 + z^2}$. It is also helpful to switch to spherical coordinates:

$$\begin{aligned} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta \end{aligned} \tag{2.13}$$

In which case (you can show this by a tedious calculation just using the chain rule, see the problem sets, its easier if you know Riemannian geometry)

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \tag{2.14}$$

Furthermore we look for energy eigenstates so

$$\Psi = e^{-iEt/\hbar}\psi(r, \theta, \phi) \tag{2.15}$$

which leads us to the following equation

$$-\frac{\hbar^2}{2m}\left(\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right)\psi + V(r)\psi = E\psi \tag{2.16}$$

We can solve this equation by separation of variables. In particular we write

$$\psi(r, \theta, \phi) = u(r)Y(\theta, \phi) \tag{2.17}$$

From here we find

$$-\frac{\hbar^2}{2mr^2}\frac{d}{dr}\left(r^2\frac{du}{dr}\right)Y - \frac{\hbar^2 u}{2mr^2}\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right)Y + (V(r) - E)uY = 0 \tag{2.18}$$

By the usual logic of separation of variables (divide by uY/r^2 and see that the first and third terms only depend on r where as the second is independent of r) splits into two equations

$$\begin{aligned} -\frac{\hbar^2}{2mu}\frac{d}{dr}\left(r^2\frac{du}{dr}\right) + (V(r) - E)r^2 &= -\epsilon \\ -\frac{1}{Y}\frac{\hbar^2}{2m}\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right)Y &= \epsilon \end{aligned} \tag{2.19}$$

where ϵ is a constant. In particular the second equation knows nothing about the potential and hence nothing about the problem at hand:

$$\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}+\frac{2m\epsilon}{\hbar^2}\right)Y=0 \tag{2.20}$$

Indeed the solutions for $Y(\theta,\phi)$ can be completely solved and the solutions are related to spherical harmonics $Y_{l,m}(\theta,\phi)$, labelled by two integers l,m with $l\geq 0$ and $|m|\leq l$.

On the other hand the first equation is a single ordinary differential equation for $u(r)$ and we have some hope to solve it:

$$-\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{du}{dr}\right)+\frac{2m}{\hbar^2}\left(V(r)+\frac{\epsilon}{r^2}-E\right)u=0 \tag{2.21}$$

This is the only place which is sensitive to the particular problem at hand through the choice of potential $V(r)$.

Thus we have reduced the problem to solving two independent differential equations. And the only one specific to the particular problem is a second order ordinary linear differential equation. One can proceed with brute force (the $Y's$ can be found by separation of variables once more). You can find them on Wikipedia or Mathematica so we won't say more about them here - come on Tuesday! Although we will say that one finds

$$\epsilon=\frac{\hbar^2\mathfrak{l}(\mathfrak{l}+1)}{2m} \tag{2.22}$$

for some non-negative integer $\mathfrak{l}$. Later we will construct them in an insightful way. Indeed spherical harmonics arise from some very important physics (angular momentum) and mathematics (Lie algebras) underlying the system. So let us explore...

3 Angular momentum

Classically the angular momentum vector $\widehat{L}$ is defined as $\widehat{L}=\boldsymbol{r}\times\boldsymbol{p}$ where $\boldsymbol{r}$ and $\boldsymbol{p}$ denote the position and momentum vector respectively. In Cartesian components, this is written as

$$\begin{aligned}L_x&=yp_z-zp_y\\L_y&=zp_x-xp_z\\L_z&=xp_y-yp_x.\end{aligned}$$

3.1 Angular momentum operator

Definition 3.1. The **angular momentum** operator is the Hermitian operator defined by

$$\begin{aligned}\widehat{L}_a&=\varepsilon_{abc}\widehat{x}_b\widehat{p}_c\\&=-i\hbar\varepsilon_{abc}x_b\frac{\partial}{\partial x_c}\end{aligned}$$

Proposition 3.2
We have the following results:

$$[L_1, L_2] = i\hbar L_3, \quad [L_2, L_3] = i\hbar L_1, \quad [L_3, L_1] = i\hbar L_2.$$

Therefore, in general we have

$$[L_a, L_b] = i\hbar\varepsilon_{abc} L_c.$$

Proof. The proof relies on the following commutator relations:

- $[A, B_1B_2] = [A, B_1]B_2 + B_1[A, B_2]$;
- $[A_1A_2, B_1B_2] = A_1[A_2, B_1]B_2 + A_1B_1[A_2, B_2] + [A_1, B_1]B_2A_2 + B_1[A_1, B_2]A_2$;
- $[\widehat{p}_i, \widehat{p}_j] = 0$;
- $[\widehat{x}_i, \widehat{p}_j] = i\hbar\delta_{ij}$;
- $[\widehat{x}_i, \widehat{x}_j] = 0$.

□

From the operators $\widehat{L}_a$ we can form the operator of the square of the modulus of the angular momentum vector, which we denote by $\widehat{L}^2$.

Definition 3.3. Define
$$\widehat{L}^2 = \widehat{L}_1^2 + \widehat{L}_2^2 + \widehat{L}_3^2.$$

Theorem 3.4
The operator $\widehat{L}^2$ commutes with each of the operators $\widehat{L}_a$ i.e.

$$\left[\widehat{L}^2, \widehat{L}_a\right] = 0.$$

The relation above means that the square of the angular momentum, can have a definite value at the same time as one of its components. The most widely used basis is constructed from eigenstates of $\widehat{L}^2$ and $\widehat{L}_3$.
Instead of the operators $\widehat{L}_1$ and $\widehat{L}_2$ it is often more convenient to use the following.

Definition 3.5. Define
$$\widehat{L}_+ = \widehat{L}_1 + i\widehat{L}_2$$
$$\widehat{L}_- = \widehat{L}_1 - i\widehat{L}_2.$$

Proposition 3.6

Some properties.

- $\widehat{L}_+^\dagger = \widehat{L}_-$;
- $[\widehat{L}_+, \widehat{L}_-] = 2\hbar\widehat{L}_3$;
- $[\widehat{L}_3, \widehat{L}_+] = \hbar\widehat{L}_+$;
- $[\widehat{L}_3, \widehat{L}_-] = -\hbar\widehat{L}_-$;
- We can write
$$\begin{aligned}\widehat{L}^2 &= \widehat{L}_+\widehat{L}_- + \widehat{L}_3^2 - \hbar\widehat{L}_3\\ &= \widehat{L}_-\widehat{L}_+ + \widehat{L}_3^2 + \hbar\widehat{L}_3.\end{aligned}$$

Proof. We prove the first bullet point:

$$\widehat{L}_+^\dagger = (\widehat{L}_1 + i\widehat{L}_2)^\dagger = \widehat{L}_1^\dagger + (i\widehat{L}_2)^\dagger = \widehat{L}_1 - i\widehat{L}_2 = \widehat{L}_-.$$

□

3.2 Algebraic theory of angular momentum

3.2.1 Lie Algebra

Definition 3.7. A **Lie Algebra** is a vector space $\mathfrak{g}$ with an antisymmetric linear map $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ called the Lie Bracket. It satisfies the following properties:

- Linearity: $[ax + by, z] = a[x, z] + b[y, z]$.
- Antisymmetric: $[x, y] = -[y, x]$.
- Jacobi's identity: $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$.

Where $a, b \in \mathbb{R}$ and $x, y, z \in \mathfrak{g}$.

Definition 3.8. Let $\mathfrak{g}$ be a Lie algebra and let V be a vector space. We let $\mathrm{gl}(V)$ denote the space of endomorphisms of V, that is, the space of all linear maps of V to itself. A **representation** of $\mathfrak{g}$ on V is a Lie algebra homomorphism

$$\rho : \mathfrak{g} \to \mathrm{GL}(V).$$

Definition 3.9. Given a representation $\rho : \mathfrak{g} \to \mathrm{End}(V)$ of a Lie algebra $\mathfrak{g}$, we say that a subspace W of V is **invariant** if $\rho(X)w \in W$ for all $w \in W$ and $X \in \mathfrak{g}$.

Definition 3.10. A non-zero representation is said to be **irreducible** if the only invariant subspaces are V itself and the zero space $\{0\}$.

3.2.2 Representation of the algebra

Note 3.11. In the section we construct an irreducible unitary representation of $\mathfrak{su}(2)$. Where by 'unitary' we mean the J_i are unitary matrices.

Definition 3.12. The Hermitian operators J_x, J_y, J_z are said to satisfy the **algebra of angular momentum** if the following commutation relation holds:

$$[J_a, J_b] = i\varepsilon_{abc}J_c.$$

Note 3.13. We recover the angular momentum by setting $L_i = \hbar J_i \Rightarrow J_i = \frac{1}{\hbar}L_i$.

Definition 3.14. We define the following operators:

$$\begin{aligned} J_+ &= J_x + iJ_y \\ J_- &= J_x - iJ_y \end{aligned}$$

such that

$$(J_+)^\dagger = J_-.$$

Proposition 3.15. We have the following relation:

$$[J_z, J_\pm] = \pm J_\pm.$$

Definition 3.16. We define the operator

$$J^2 = J_x^2 + J_y^2 + J_z^2.$$

Theorem 3.17. We have that

$$[J^2, J_i] = 0.$$

Proof. It follows from the definition of the algebra and proof is similar to the one where $L_i = \hbar J_i$. ☐

Note 3.18. The operator J^2 is said to be **Casimir**, which means that it commutes with all of its generators. When the commutator of two operators is zero it means that they can be simultaneously diagonalised. That is, meaning there exists a basis where both operators are represented by diagonal matrices simultaneously.

Basis

We now try to find a basis for this representation of $\mathfrak{su}(2)$. Since $[J^2, J_z] = 0$ it means that these two operators are simultaneously diagonalised i.e.

$$J^2 \left|\lambda, \mathfrak{m}\right\rangle = \lambda \left|\lambda, \mathfrak{m}\right\rangle \quad \text{and} \quad J_z \left|\lambda, \mathfrak{m}\right\rangle = \mathfrak{m} \left|\lambda, \mathfrak{m}\right\rangle.$$

where $\lambda, \mathfrak{m} \in \mathbb{R}$. The first label of the eigenstate $|\lambda, \mathfrak{m}\rangle$ represents the eigenvalue of the J^2 operator and the second label the eigenvalue of J_z. By the construction of the ladder operator $J_\pm$ we can use it raise or lower the eigenvalues of J_z.

Proposition 3.19

We have that

$$J_+ \left|\lambda, \mathfrak{m}\right\rangle = c_{\mathfrak{m}} \left|\lambda, \mathfrak{m}+1\right\rangle \quad \text{and} \quad J_- \left|\lambda, \mathfrak{m}\right\rangle = d_{\mathfrak{m}} \left|\lambda, \mathfrak{m}-1\right\rangle$$

where $c_{\mathfrak{m}}$ and $d_{\mathfrak{m}}$ are chosen to ensure that the states are normalised. Therefore, we have

$$c_{\mathfrak{m}} = \sqrt{\lambda - \mathfrak{m}^2 - \mathfrak{m}} \quad \text{and} \quad d_{\mathfrak{m}} = \sqrt{\lambda - \mathfrak{m}^2 + \mathfrak{m}}.$$

Proof. Recall the commutator relation $[J_z, J_\pm] = J_z J_\pm - J_\pm J_z$, and notice we can write $J_z J_\pm = [J_z, J_\pm] + J_\pm J_z$. Therefore,

$$\begin{aligned} J_z \left(J_+ \left|\lambda, \mathfrak{m}\right\rangle\right) &= \left([J_z, J_\pm] + J_\pm J_z\right) \left|\lambda, \mathfrak{m}\right\rangle \\ &= \left(\pm J_\pm + \mathfrak{m} J_\pm\right) \left|\lambda, \mathfrak{m}\right\rangle \\ &= (\mathfrak{m} \pm 1) J_\pm \left|\lambda, \mathfrak{m}\right\rangle . \end{aligned}$$

The coefficients $c_{\mathfrak{m}}$ and $d_{\mathfrak{m}}$ were chosen such that the corresponding states would be normalised hence, we normalise the states. To calculate $c_{\mathfrak{m}}$ we evaluate

$$\begin{aligned} \left|c_{\mathfrak{m}}\right|^2 \left\langle \mathfrak{m}+1 \middle| \mathfrak{m}+1\right\rangle &= \left(J_+ \left|\lambda, \mathfrak{m}\right\rangle\right)^\dagger J_+ \left|\lambda, \mathfrak{m}\right\rangle \\ &= \left\langle\lambda, \mathfrak{m}\right| J_- J_+ \left|\lambda, \mathfrak{m}\right\rangle \\ &= \left\langle\lambda, \mathfrak{m}\right| J_x^2 + J_y^2 + i[J_x, J_y] \left|\lambda, \mathfrak{m}\right\rangle \\ &= \left\langle\lambda, \mathfrak{m}\right| J^2 - J_z^2 - J_z \left|\lambda, \mathfrak{m}\right\rangle \\ &= \left(\lambda - \mathfrak{m}^2 - \mathfrak{m}\right) \left\langle\lambda, \mathfrak{m} \middle| \lambda, \mathfrak{m}\right\rangle . \end{aligned}$$

Note 3.20. We do not know how J_x and J_y act on the states so, we use a relation to express them in terms of operators which we know how they act on the states, namely J^2 and J_z.

Since the states are normalised we have $\langle \mathfrak{m} | \mathfrak{m} \rangle = \langle \mathfrak{m}+1 | \mathfrak{m}+1 \rangle = 1$, and we find that

$$c_{\mathfrak{m}} = \sqrt{\lambda - \mathfrak{m}^2 - \mathfrak{m}}.$$

We do a similar process for $d_{\mathfrak{m}}$. ☐

Finite representation

If we have a finitely dimensional representation them there must exist a highest and lowest value of J_z eigenvalues, denote these by $\mathfrak{m}_h$ and $\mathfrak{m}_l$.

Theorem 3.21

The highest and lowest eigenvalues of J_z satisfy

$$J_+ \left|\lambda, \mathfrak{m}_h\right\rangle = 0 \quad \text{and} \quad J_- \left|\lambda, \mathfrak{m}_l\right\rangle = 0.$$

Note 3.22. If we repeatedly apply say, the rasing operator, we will reach a point where we exceed the total angular momentum. This cannot happen thus, there must exists a maximum state. Similar arguments apply for the lowring operator.

Corollary 3.23

The Lie Algebra $\mathfrak{su}(2)$ has a finite unitary representation of dimension $2\mathfrak{l}+1$ for $\mathfrak{l} = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$ i.e. $\mathfrak{l}$ has increments of $\frac{1}{2}$. Furthermore, the states are labelled by

$$|\mathfrak{l}, \mathfrak{m}\rangle \quad \text{where } \mathfrak{m} = -\mathfrak{l}, \mathfrak{l}+1, \dots, \mathfrak{l}-1, \mathfrak{l}.$$

Note 3.24. It is convention to set $\mathfrak{m}_h = \mathfrak{l}$ which implies that $\lambda = \mathfrak{l}(\mathfrak{l}+1)$.

Proof. The highest and lowest state require that $c_{\mathfrak{m}_h} = d_{\mathfrak{m}_l} = 0$. We have a formula for these:

$$\lambda - \mathfrak{m}_h(\mathfrak{m}_h + 1) = 0 \quad \text{and} \quad \lambda - \mathfrak{m}_l(\mathfrak{m}_l - 1) = 0.$$

As such we obtain

$$\mathfrak{m}_h(\mathfrak{m}_h + 1) = \mathfrak{m}_l(\mathfrak{m}_l - 1).$$

By inspection, we have

$$\mathfrak{m}_l = -\mathfrak{m}_h \quad \text{or} \quad \mathfrak{m}_l = \mathfrak{m}_h + 1.$$

The second solution is invalid as we require $\mathfrak{m}_l \leq \mathfrak{m}_h$ by construction. Hence, the spectrum of eigenvalues of J_z is:

$$\mathfrak{m}_h, \mathfrak{m}_h - 1, \dots, 0, \dots, \mathfrak{m}_l + 1, \mathfrak{m}_l$$

equivalently,

$$\underbrace{\mathfrak{m}_h, \mathfrak{m}_h - 1, \dots, 0, \dots, -\mathfrak{m}_h + 1, -\mathfrak{m}_h}_{2\mathfrak{m}_h+1}$$

with a <u>single state</u> assigned to each eigenvalue. Therefore, the representation has dimension $2\mathfrak{m}_h + 1 \in \mathbb{N} \setminus \{0\}$, which implies that

$$\mathfrak{m}_h = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$$

□

Example 3.25. Let us look at some examples. We recall for $\mathfrak{l} = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$ the following

$$|\mathfrak{l}, \mathfrak{m}\rangle \quad \text{where } \mathfrak{m} = -\mathfrak{l}, \mathfrak{l}+1, \dots, \mathfrak{l}-1, \mathfrak{l}.$$

- When $\mathfrak{l} = 0$ we have that $\mathfrak{m} = 0$. Therefore, we have just one state $|0, 0\rangle$ and the matrices J_i act trivially. This is the trivial representation.
- When $\mathfrak{l} = \frac{1}{2}$ we have that $\mathfrak{m} = -\frac{1}{2}, \frac{1}{2}$. Hence, we have two states. Now, any state in this basis must be in the form $|v\rangle = a\left|\frac{1}{2}, \frac{1}{2}\right\rangle + b\left|\frac{1}{2}, -\frac{1}{2}\right\rangle$ thus, we set the basis state to be as simple as possible:

$$\left|\frac{1}{2}, \frac{1}{2}\right\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

We contruct

$$J_3 = \begin{pmatrix} \left\langle \frac{1}{2}, \frac{1}{2} \right| J_3 \left| \frac{1}{2}, \frac{1}{2} \right\rangle & \left\langle \frac{1}{2}, \frac{1}{2} \right| J_3 \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \\ \left\langle \frac{1}{2}, -\frac{1}{2} \right| J_3 \left| \frac{1}{2}, \frac{1}{2} \right\rangle & \left\langle \frac{1}{2}, -\frac{1}{2} \right| J_3 \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}.$$

We can determine J_+ through

$$J_+ \left| \frac{1}{2}, \frac{1}{2} \right\rangle = 0 \quad J_+ \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = \sqrt{\frac{3}{4} - \frac{1}{4}} \left| \frac{1}{2}, \frac{1}{2} \right\rangle = \left| \frac{1}{2}, \frac{1}{2} \right\rangle.$$

so that

$$J_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

And can determine J_- through

$$J_- \left| \frac{1}{2}, \frac{1}{2} \right\rangle = \sqrt{\frac{3}{4} - \frac{1}{4}} \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = \left| \frac{1}{2}, -\frac{1}{2} \right\rangle$$

so that

$$J_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Or alternatively

$$J_1 = \frac{1}{2}(J_+ + J_-) = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$J_2 = \frac{1}{2i}(J_+ - J_-) = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

Thus we have recovered the Pauli matrices.

Note 3.26. Throughout this process we have used the results obtained from above.

3.3 The eigenvalues of the angular momentum

Note 3.27. In the previous section, we dealt with the Lie algebra of $\mathfrak{su}(2)$ and obtained an irreducible representation. To relate our previous results to the algebra of the angular momentum we only need to add a factor of $\hbar$.

Theorem 3.28

We summarise the results of the previous section in terms of the angular momentum here.

- $L^2 \left|\mathfrak{l},\mathfrak{m}\right\rangle = \hbar^2\mathfrak{l}(\mathfrak{l}+1)\left|\mathfrak{l},\mathfrak{m}\right\rangle$
- $L_3 \left|\mathfrak{l},\mathfrak{m}\right\rangle = \hbar\mathfrak{m}\left|\mathfrak{l},\mathfrak{m}\right\rangle$.
- $L_+ \left|\mathfrak{l},\mathfrak{m}\right\rangle = c_{\mathfrak{m}}\left|\mathfrak{l},\mathfrak{m}+1\right\rangle$ where

$$c_{\mathfrak{m}} = \hbar\sqrt{\mathfrak{l}(\mathfrak{l}+1)-\mathfrak{m}^2-\mathfrak{m}}.$$

- $L_+ \left|\mathfrak{l},\mathfrak{m}\right\rangle = d_{\mathfrak{m}}\left|\mathfrak{l},\mathfrak{m}-1\right\rangle$ where

$$d_{\mathfrak{m}} = \hbar\sqrt{\mathfrak{l}(\mathfrak{l}+1)-\mathfrak{m}^2+\mathfrak{m}}.$$

3.4 Spin-1/2

Proposition 3.29

We can build the representation of this by

$$\widehat{S}_i = \frac{\hbar}{2}\sigma_i$$

where σ_i are the Pauli matrices.

Note 3.30. The matrices obey the Lie algebra of the angular momentum.

3.5 The eigenvectors of the angular momentum

Proposition 3.31

We write the angular momentum operators in spherical coordinatess.

- $L_1 = i\hbar\left(\sin\phi\partial_\theta + \cot\theta\cos\phi\partial_\phi\right)$.
- $L_2 = i\hbar\left(-\cos\phi\partial_\theta + \cot\theta\sin\phi\partial_\phi\right)$.
- $L_3 = -i\hbar\partial_\phi$.
- $L_+ = \hbar e^{i\phi}\left[\partial_\theta + i\cot\theta\partial_\phi\right]$.
- $L_- = \hbar e^{-i\phi}\left[\partial_\theta + i\cot\theta\partial_\phi\right]$

Theorem 3.32

We see that $|\mathfrak{l},\mathfrak{m}\rangle = Y_{\mathfrak{l},\mathfrak{m}}$ i.e. the spherical harmonics. Hence, the set $\{Y_{\mathfrak{l},\mathfrak{m}} : \mathfrak{m} = -\mathfrak{l},\dots,\mathfrak{l}\}$ provides an irreducible $(2\mathfrak{l}+1)$-dimensional representation of $\mathfrak{su}(2)$ inside the space of all differentiable functions on S^2 (viewed as the unit sphere in $\mathbb{R}^3$).

3.6 Addition of angular momenta

In this section we consider a system of two particles. We assume that each particle carries and angular momentum. The combined system will have a single wave function $\Psi(t, \boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)})$. Thus, we define two sets of angular momentum operators $\widehat{L}^{(1)}$ and $\widehat{L}^{(2)}$ which act on $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$ separately (i.e. $L_a^{(1)} = -i\hbar\varepsilon_{abc}x_b^{(1)}\partial_c^{(1)}$).

Definition 3.33. The total angular momentum associated to the rotational symmetry of the system as a whole is

$$\widehat{L}^{(T)} = \widehat{L}^{(1)} + \widehat{L}^{(2)}.$$

Proposition 3.34. The following set of operators commutes:

$$\left(\widehat{L}^{(1)}\right)^2, \widehat{L}_z^{(1)}, \left(\widehat{L}^{(2)}\right)^2, \widehat{L}_z^{(2)}.$$

We label the eigenvectors diagonalising simultaneously these 4 operators by

$$|\mathfrak{l}_1, \mathfrak{m}_1; \mathfrak{l}_2, \mathfrak{m}_2\rangle$$

where the $\mathfrak{m}_i$ range between

$$\begin{aligned} -\mathfrak{l}_1 \leq &\mathfrak{m}_1 \leq \mathfrak{l}_1 \\ -\mathfrak{l}_2 \leq &\mathfrak{m}_2 \leq \mathfrak{l}_2. \end{aligned}$$

Therefore, for a fixed $\mathfrak{l}_1$ and $\mathfrak{l}_2$ there are a total of $(2\mathfrak{l}_1+1)(2\mathfrak{l}_2+1)$ states.

Proposition 3.35

The following set of operators commutes:

$$\left(\widehat{L}^{(T)}\right)^2, \widehat{L}_z^{(T)}, \left(\widehat{L}^{(1)}\right)^2, \left(\widehat{L}^{(2)}\right)^2.$$

We denote the eigenvectors diagonalising the operators as

$$|\mathfrak{l}, \mathfrak{m}; \mathfrak{l}_1, \mathfrak{l}_2\rangle,$$

where for a fixed $\mathfrak{l}_1$ and $\mathfrak{l}_2$ we have that

$$\mathfrak{l} = \mathfrak{l}_1 - \mathfrak{l}_2, (\mathfrak{l}_1 - \mathfrak{l}_2 + 1), (\mathfrak{l}_1 - \mathfrak{l}_2 + 2), \dots, \mathfrak{l}_1 + \mathfrak{l}_2.$$

The above representation is equivalent to a change of basis from the previous set to the following, and so we have

$$|\mathfrak{l}, \mathfrak{m}; \mathfrak{l}_1, \mathfrak{l}_2\rangle = \sum_{\mathfrak{m}_1=-\mathfrak{l}_1}^{\mathfrak{l}_1} \sum_{\mathfrak{m}_2=-\mathfrak{l}_2}^{\mathfrak{l}_2} C^{\mathfrak{l}_1,\mathfrak{l}_2,\mathfrak{l}}_{\mathfrak{m}_1,\mathfrak{m}_2,\mathfrak{m}} |\mathfrak{l}_1, \mathfrak{m}_1; \mathfrak{l}_2, \mathfrak{m}_2\rangle,$$

where the coefficients $C^{\mathfrak{l}_1,\mathfrak{l}_2,\mathfrak{l}}_{\mathfrak{m}_1,\mathfrak{m}_2,\mathfrak{m}}$ are called the **Clebsch-Gordan** coefficients.

Note 3.36. It is more natural to consider the above set of operators when dealing with addition of angular momenta.

Note 3.37. In some other texts it $|\mathfrak{l}_1 - \mathfrak{l}_2|$.

Proposition 3.38

The sum of angular momenta is an angular momentum in the tensor product:

$$\widehat{L}_a^{(T)} = \widehat{L}_a^{(1)} \otimes \mathbb{I} + \mathbb{I} \otimes \widehat{L}_a^{(2)} \text{ which satisfies } \left[\widehat{L}_a, \widehat{L}_b\right] = i\hbar\varepsilon_{abc}\widehat{L}_c \text{ acting on } \mathcal{H}_1 \otimes \mathcal{H}_2.$$

Therefore, we can write

$$|\mathfrak{l}_1, \mathfrak{m}_1; \mathfrak{l}_2, \mathfrak{m}_2\rangle = |\mathfrak{l}_1, \mathfrak{m}_1\rangle \otimes |\mathfrak{l}_2, \mathfrak{m}_2\rangle .$$

This is how the total angular momentum operator acts:

$$\begin{aligned} L_a^{(T)} |\mathfrak{l}_1, \mathfrak{m}_1; \mathfrak{l}_2, \mathfrak{m}_2\rangle &= \left(L_a^{(1)} + L_a^{(2)}\right) |\mathfrak{l}_1, \mathfrak{m}_1; \mathfrak{l}_2, \mathfrak{m}_2\rangle \\ &= L_a^{(1)} |\mathfrak{l}_1, \mathfrak{m}_1\rangle \otimes |\mathfrak{l}_2, \mathfrak{m}_2\rangle + |\mathfrak{l}_1, \mathfrak{m}_1\rangle \otimes L_a^{(2)} |\mathfrak{l}_2, \mathfrak{m}_2\rangle . \end{aligned}$$

Note 3.39. The 2 particle wavefunction lives in $L^2(\mathbb{R}^6) = L^2(\mathbb{R}^3) \otimes L^2(\mathbb{R}^3)$.

Note 3.40. All of the above is also valid for J_i, by removing the $\hbar$ factor.

Example 3.41

We have that

$$\begin{aligned} J_3^{(T)} \left(|\mathfrak{l}_1, \mathfrak{m}_1\rangle \otimes |\mathfrak{l}_2, \mathfrak{m}_2\rangle\right) &= \left(J_3^{(1)} + J_3^{(2)}\right) |\mathfrak{l}_1, \mathfrak{m}_1\rangle \otimes |\mathfrak{l}_2, \mathfrak{m}_2\rangle \\ &= J_3^{(1)} |\mathfrak{l}_1, \mathfrak{m}_1\rangle \otimes |\mathfrak{l}_2, \mathfrak{m}_2\rangle + |\mathfrak{l}_1, \mathfrak{m}_1\rangle \otimes J_3^{(2)} |\mathfrak{l}_2, \mathfrak{m}_2\rangle \\ &= \mathfrak{m}_1 |\mathfrak{l}_1, \mathfrak{m}_1\rangle \otimes |\mathfrak{l}_2, \mathfrak{m}_2\rangle + |\mathfrak{l}_1, \mathfrak{m}_1\rangle \otimes \mathfrak{m}_2 |\mathfrak{l}_2, \mathfrak{m}_2\rangle \\ &= (\mathfrak{m}_1 + \mathfrak{m}_2) |\mathfrak{l}_1, \mathfrak{m}_1\rangle \otimes |\mathfrak{l}_2, \mathfrak{m}_2\rangle . \end{aligned}$$

Example 3.42. We consider the case when $\mathfrak{l}_1 = \mathfrak{l}_2 = \frac{1}{2}$. There are $(2\left(\frac{1}{2}\right) + 1)^2 = 4$ states. Since we are in this case we suppress some notation since the values are the same and instead write:

$$|\mathfrak{l}_1, \mathfrak{m}_1; \mathfrak{l}_2, \mathfrak{m}_2\rangle = |\mathfrak{m}_1, \mathfrak{m}_2\rangle .$$

So,

$$\begin{aligned}
\left|\frac{1}{2},\frac{1}{2};\frac{1}{2},\frac{1}{2}\right\rangle &= \left|\frac{1}{2},\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2},\frac{1}{2}\right\rangle = \left|\frac{1}{2},\frac{1}{2}\right\rangle \\
\left|\frac{1}{2},\frac{1}{2};\frac{1}{2},-\frac{1}{2}\right\rangle &= \left|\frac{1}{2},\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2},-\frac{1}{2}\right\rangle = \left|\frac{1}{2},-\frac{1}{2}\right\rangle \\
\left|\frac{1}{2},-\frac{1}{2};\frac{1}{2},\frac{1}{2}\right\rangle &= \left|\frac{1}{2},-\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2},\frac{1}{2}\right\rangle = \left|-\frac{1}{2},\frac{1}{2}\right\rangle \\
\left|\frac{1}{2},-\frac{1}{2};\frac{1}{2},-\frac{1}{2}\right\rangle &= \left|\frac{1}{2},-\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2},-\frac{1}{2}\right\rangle = \left|-\frac{1}{2},-\frac{1}{2}\right\rangle
\end{aligned}$$

These are eigenstates of $J_3^{(1)} + J_3^{(2)}$

$$\begin{aligned}
(J_3^{(1)} + J_3^{(2)})\left|\frac{1}{2},\frac{1}{2}\right\rangle &= J_3^{(1)}\left|\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}\right\rangle + \left|\frac{1}{2}\right\rangle \otimes J_3^{(2)}\left|\frac{1}{2}\right\rangle \\
&= \frac{1}{2}\left|\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}\right\rangle + \frac{1}{2}\left|\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}\right\rangle \\
&= \left|\frac{1}{2},\frac{1}{2}\right\rangle \\
(J_3^{(1)} + J_3^{(2)})\left|\frac{1}{2},-\frac{1}{2}\right\rangle &= J_3^{(1)}\left|\frac{1}{2}\right\rangle \otimes \left|-\frac{1}{2}\right\rangle + \left|\frac{1}{2}\right\rangle \otimes J_3^{(2)}\left|-\frac{1}{2}\right\rangle \\
&= \frac{1}{2}\left|\frac{1}{2}\right\rangle \otimes \left|-\frac{1}{2}\right\rangle - \frac{1}{2}\left|\frac{1}{2}\right\rangle \otimes \left|-\frac{1}{2}\right\rangle \\
&= 0 \\
(J_3^{(1)} + J_3^{(2)})\left|-\frac{1}{2},\frac{1}{2}\right\rangle &= J_3^{(1)}\left|-\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}\right\rangle + \left|-\frac{1}{2}\right\rangle \otimes J_3^{(2)}\left|\frac{1}{2}\right\rangle \\
&= -\frac{1}{2}\left|-\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}\right\rangle + \frac{1}{2}\left|-\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}\right\rangle \\
&= 0 \\
(J_3^{(1)} + J_3^{(2)})\left|-\frac{1}{2},-\frac{1}{2}\right\rangle &= J_3^{(1)}\left|-\frac{1}{2}\right\rangle \otimes \left|-\frac{1}{2}\right\rangle + \left|-\frac{1}{2}\right\rangle \otimes J_3^{(2)}\left|-\frac{1}{2}\right\rangle \\
&= -\frac{1}{2}\left|-\frac{1}{2}\right\rangle \otimes \left|-\frac{1}{2}\right\rangle - \frac{1}{2}\left|-\frac{1}{2}\right\rangle \otimes \left|-\frac{1}{2}\right\rangle \\
&= -\left|-\frac{1}{2},-\frac{1}{2}\right\rangle
\end{aligned}$$

Thus we find the $J_3^{(T)}$ eigenvalues $\mathfrak{m} = -1, 0, 0, +1$. This doesn't correspond to any irreducible representation of $\mathfrak{su}(2)$ but we can split it as 0 and $(-1, 0, 1)$ which are the eigenvalues of the $\mathfrak{l} = 0$ and $\mathfrak{l} = 1$ representations. Thus we expect to find

$$2 \otimes 2 = 1 \oplus 3$$

Indeed the correct basis is

$$\begin{aligned}
|0,0\rangle &= \frac{1}{\sqrt{2}}\left|\frac{1}{2},-\frac{1}{2}\right\rangle - \frac{1}{\sqrt{2}}\left|-\frac{1}{2},\frac{1}{2}\right\rangle \\
|1,-1\rangle &= \left|-\frac{1}{2},-\frac{1}{2}\right\rangle \\
|1,0\rangle &= \frac{1}{\sqrt{2}}\left|\frac{1}{2},-\frac{1}{2}\right\rangle + \frac{1}{\sqrt{2}}\left|-\frac{1}{2},\frac{1}{2}\right\rangle \\
|1,1\rangle &= \left|\frac{1}{2},\frac{1}{2}\right\rangle .
\end{aligned}$$

Note 3.43. Here we wrote $|\mathfrak{l},\mathfrak{m}\rangle$ instead of $\left|\mathfrak{l},\mathfrak{m},\frac{1}{2},\frac{1}{2}\right\rangle$.

3.6.1 Example

1 Week Five: Addition of Angular Momentum

Consider the tensor product of two $\mathfrak{l} = 1/2$ representations of $su(2)$. We want to show that the states

$$\begin{aligned}
|0,0\rangle &= |\tfrac{1}{2},-\tfrac{1}{2}\rangle - \frac{1}{\sqrt{2}}|-\tfrac{1}{2},\tfrac{1}{2}\rangle \\
|1,-1\rangle &= |-\tfrac{1}{2},-\tfrac{1}{2}\rangle \\
|1,0\rangle &= \frac{1}{\sqrt{2}}|\tfrac{1}{2},-\tfrac{1}{2}\rangle + \frac{1}{\sqrt{2}}|-\tfrac{1}{2},\tfrac{1}{2}\rangle \\
|1,1\rangle &= \frac{1}{\sqrt{2}}|\tfrac{1}{2},\tfrac{1}{2}\rangle
\end{aligned}$$

form a basis for the $\mathfrak{l} = 0$ and $\mathfrak{l} = 1$ representations of $su(2)$. Note that here we have we have suppressed the $\mathfrak{l} = 1/2$ eigenvalue labels as they are common to all states *i.e.*

$$|\tfrac{1}{2},-\tfrac{1}{2}\rangle = |\tfrac{1}{2},\tfrac{1}{2};\tfrac{1}{2},-\tfrac{1}{2}\rangle = |\tfrac{1}{2},\tfrac{1}{2}\rangle \otimes |\tfrac{1}{2},-\tfrac{1}{2}\rangle \tag{1}$$

We can read off the Clebsch-Gordan coefficients:

$$|\mathfrak{l},\mathfrak{m},\mathfrak{l}_1,\mathfrak{l}_2\rangle = \sum_{\mathfrak{m}_1=-\mathfrak{l}_1}^{\mathfrak{l}_1}\sum_{\mathfrak{m}_2=-\mathfrak{l}_2}^{\mathfrak{l}_2} C^{\mathfrak{l}_1,\mathfrak{l}_2,\mathfrak{l}}_{\mathfrak{m}_1,\mathfrak{m}_2,\mathfrak{m}}|\mathfrak{l}_1,\mathfrak{m}_1;\mathfrak{l}_2,\mathfrak{m}_2\rangle \tag{2}$$

So here we find:

$$\begin{aligned}
C^{\frac{1}{2}\frac{1}{2}0}_{\frac{1}{2}-\frac{1}{2}0} &= -C^{\frac{1}{2}\frac{1}{2}0}_{-\frac{1}{2}\frac{1}{2}0} = \frac{1}{\sqrt{2}} \\
C^{\frac{1}{2}\frac{1}{2}1}_{\frac{1}{2}\frac{1}{2}1} &= 1 \\
C^{\frac{1}{2}\frac{1}{2}1}_{\frac{1}{2}-\frac{1}{2}0} &= C^{\frac{1}{2}\frac{1}{2}0}_{-\frac{1}{2}\frac{1}{2}0} = \frac{1}{\sqrt{2}} \\
C^{\frac{1}{2}\frac{1}{2}1}_{-\frac{1}{2}-\frac{1}{2}-1} &= 1
\end{aligned}$$

It is generically true that $C^{\mathfrak{l}_1,\mathfrak{l}_2,\mathfrak{l}}_{\mathfrak{l}_1,\mathfrak{l}_2,\mathfrak{m}} = \delta_{\mathfrak{m},\mathfrak{l}_1+\mathfrak{l}_2}$ and $C^{\mathfrak{l}_1,\mathfrak{l}_2,\mathfrak{l}}_{-\mathfrak{l}_1,-\mathfrak{l}_2,\mathfrak{m}} = \delta_{\mathfrak{m},-\mathfrak{l}_1-\mathfrak{l}_2}$ as there is a unique way to construct the highest and lowest weight states

We saw in the lectures that

$$J_3|0,0\rangle = 0 \qquad J_3|1,\mathfrak{m}\rangle = \mathfrak{m}|1,\mathfrak{m}\rangle$$

Let's look at

$$\begin{aligned}
J_+^{(T)}|0,0\rangle =& \frac{1}{\sqrt{2}}J_+^{(1)}|\tfrac{1}{2}\rangle \otimes |-\tfrac{1}{2}\rangle + \frac{1}{\sqrt{2}}|\tfrac{1}{2}\rangle \otimes J_+^{(2)}|-\tfrac{1}{2}\rangle \\
&- \frac{1}{\sqrt{2}}J_+^{(1)}|-\tfrac{1}{2}\rangle \otimes |\tfrac{1}{2}\rangle - \frac{1}{\sqrt{2}}|-\tfrac{1}{2}\rangle \otimes J_+^{(2)}|\tfrac{1}{2}\rangle \\
&= 0 + \frac{1}{\sqrt{2}}|\tfrac{1}{2}\rangle \otimes |\tfrac{1}{2}\rangle - \frac{1}{\sqrt{2}}|\tfrac{1}{2}\rangle \otimes |\tfrac{1}{2}\rangle + 0 \\
&= 0 \\
J_-^{(T)}|0,0\rangle =& \frac{1}{\sqrt{2}}J_-^{(1)}|\tfrac{1}{2}\rangle \otimes |-\tfrac{1}{2}\rangle + \frac{1}{\sqrt{2}}|\tfrac{1}{2}\rangle \otimes J_-^{(2)}|-\tfrac{1}{2}\rangle \\
&- \frac{1}{\sqrt{2}}J_-^{(1)}|-\tfrac{1}{2}\rangle \otimes |\tfrac{1}{2}\rangle - \frac{1}{\sqrt{2}}|-\tfrac{1}{2}\rangle \otimes J_-^{(2)}|\tfrac{1}{2}\rangle \\
&= \frac{1}{\sqrt{2}}|-\tfrac{1}{2}\rangle \otimes |-\tfrac{1}{2}\rangle + 0 - 0 - \frac{1}{\sqrt{2}}|-\tfrac{1}{2}\rangle \otimes |-\tfrac{1}{2}\rangle \\
&= 0
\end{aligned}$$

So $|0,0\rangle$ is indeed invariant under all the J's.

Next consider the rest $|1,\mathfrak{m}\rangle$:

$$\begin{aligned}
J_+^{(T)}|1,-1\rangle &= J_+^{(1)}|-\tfrac{1}{2}\rangle \otimes |-\tfrac{1}{2}\rangle + |-\tfrac{1}{2}\rangle \otimes J_+^{(2)}|-\tfrac{1}{2}\rangle \\
&= |\tfrac{1}{2}\rangle \otimes |-\tfrac{1}{2}\rangle + |-\tfrac{1}{2}\rangle \otimes |\tfrac{1}{2}\rangle \\
&= \sqrt{2}|1,0\rangle \\
J_-^{(T)}|1,-1\rangle &= J_-^{(1)}|-\tfrac{1}{2}\rangle \otimes |-\tfrac{1}{2}\rangle + |-\tfrac{1}{2}\rangle \otimes J_-^{(2)}|-\tfrac{1}{2}\rangle \\
&= 0
\end{aligned}$$

$$\begin{aligned}
J_+^{(T)}|1,0\rangle =& \frac{1}{\sqrt{2}}J_+^{(1)}|\tfrac{1}{2}\rangle \otimes |-\tfrac{1}{2}\rangle + \frac{1}{\sqrt{2}}|\tfrac{1}{2}\rangle \otimes J_+^{(2)}|-\tfrac{1}{2}\rangle \\
&+ \frac{1}{\sqrt{2}}J_+^{(1)}|-\tfrac{1}{2}\rangle \otimes |\tfrac{1}{2}\rangle + \frac{1}{\sqrt{2}}|-\tfrac{1}{2}\rangle \otimes J_+^{(2)}|\tfrac{1}{2}\rangle \\
&= 0 + \frac{1}{\sqrt{2}}|\tfrac{1}{2}\rangle \otimes |\tfrac{1}{2}\rangle + \frac{1}{\sqrt{2}}|\tfrac{1}{2}\rangle \otimes |\tfrac{1}{2}\rangle + 0 \\
&= \sqrt{2}|1,1\rangle \\
J_-^{(T)}|1,0\rangle =& \frac{1}{\sqrt{2}}J_-^{(1)}|\tfrac{1}{2}\rangle \otimes |-\tfrac{1}{2}\rangle + \frac{1}{\sqrt{2}}|\tfrac{1}{2}\rangle \otimes J_-^{(2)}|-\tfrac{1}{2}\rangle \\
&+ \frac{1}{\sqrt{2}}J_-^{(1)}|-\tfrac{1}{2}\rangle \otimes |\tfrac{1}{2}\rangle + \frac{1}{\sqrt{2}}|-\tfrac{1}{2}\rangle \otimes J_-^{(2)}|\tfrac{1}{2}\rangle \\
&= \frac{1}{\sqrt{2}}|-\tfrac{1}{2}\rangle \otimes |-\tfrac{1}{2}\rangle + 0 + 0 + \frac{1}{\sqrt{2}}|-\tfrac{1}{2}\rangle \otimes |-\tfrac{1}{2}\rangle \\
&= \sqrt{2}|1,-1\rangle
\end{aligned}$$

$$\begin{aligned}
J_+^{(T)}|1,1\rangle &= J_+^{(1)}|\tfrac{1}{2}\rangle \otimes |\tfrac{1}{2}\rangle + |\tfrac{1}{2}\rangle \otimes J_+^{(2)}|\tfrac{1}{2}\rangle \\
&= 0 \\
J_-^{(T)}|1,1\rangle &= J_-^{(1)}|\tfrac{1}{2}\rangle \otimes |\tfrac{1}{2}\rangle + |\tfrac{1}{2}\rangle \otimes J_-^{(2)}|\tfrac{1}{2}\rangle \otimes |\tfrac{1}{2}\rangle + |\tfrac{1}{2}\rangle \otimes |-\tfrac{1}{2}\rangle \\
&= \sqrt{2}|1,0\rangle
\end{aligned}$$

4 The Hydrogen atom

Chapter 3

The Hydrogen atom

3.1 Motion in a spherically symmetric field

The problem of the motion of two interacting particles (like an electron and proton) can be reduced in quantum mechanics, as is true also in classical mechanics, to that of one particle in a potential. The Hamiltonian of the two particles (of masses m_1, m_2) is of the form

$$\hat{H} = -\frac{\hbar^2}{2m_1}\Delta_1 - \frac{\hbar^2}{2m_2}\Delta_2 + U(r) \tag{3.1}$$

where $r = |\vec{r}_1 - \vec{r}_2|$ is the distance between the particles and $U(r)$ is the interaction potential. Δ_1 and Δ_2 are the Laplacians with respect to the coordinates $\vec{r}_1$ and $\vec{r}_2$. Exactly like in classical mechanics, one can introduce the center of mass $\vec{R}$ and relative position $\vec{r}$ by

$$\vec{r} = \vec{r}_2 - \vec{r}_1\,, \qquad \vec{R} = \frac{m_1\vec{r}_1 + m_2\vec{r}_2}{m_1 + m_2}\,. \tag{3.2}$$

In these new coordinates the Hamiltonian is

$$\hat{H} = \hat{H}_R + \hat{H}_r\,, \quad \text{where} \quad \hat{H}_R = -\frac{\hbar^2}{2(m_1+m_2)}\Delta_R\,, \qquad \hat{H}_r = -\frac{\hbar^2}{2\mu}\Delta_r + U(r) \tag{3.3}$$

where $\mu = m_1 m_2/(m_1 + m_2)$ is the reduced mass. Thus the Hamiltonian decouples into two noninteracting parts and we can consider them separately. The first one is just the free motion of the centre of mass and the second is effectively the particle in a central potential, as presented before, see (2.2).

Having studied angular momentum, we notice that we can express all the angular derivatives in terms of the square of the angular momentum $\hat{L}^2$, so we find the eigenvalue problem

$$\frac{1}{r^2}\partial_r(r^2\partial_r\psi) - \frac{\hat{L}^2}{\hbar^2 r^2}\psi + \frac{2\mu}{\hbar^2}(E - U(r))\psi = 0\,. \tag{3.4}$$

Classically the angular momentum is conserved during motion in a centrally symmetric field. The analogous statement in quantum mechanics is that the angular momentum observables compute with the Hamiltonian. The reason is that the time dependence of a state is governed by the Hamiltonian so then the expectation value of an operator $\hat{A}$ that commutes with $\hat{H}$, *i.e.*, $\langle\psi(t)|A^n|\psi(t)\rangle$ does not depend on time).[1]

As we showed, $\hat{L}_i$ in spherical coordinates depend only on the angular directions ϕ and θ and commute with $\hat{r}$ and p_r (as well as $\hat{L}^2$). Therefore

$$[\hat{L}^2, \hat{H}] = 0\,, \qquad [\hat{L}_i, \hat{H}] = 0\,. \tag{3.5}$$

This means that we can chose $\hat{H}, \hat{L}^2, \hat{L}_3$ to be simultaneously diagonal.

From the previous lecture we know that the eigenfunctions of $\hat{L}^2$ and $\hat{L}_3$ with the eigenvalues $l(l+1)$ and m correspondingly are

$$Y_{lm}(\theta, \phi) \tag{3.6}$$

thus we seek solutions of (3.4) in the form

$$\psi = R(r)Y_{lm}(\theta, \phi) \tag{3.7}$$

it is convenient to redefine $R(r) = \chi(r)/r$, then the equation for χ is

$$\chi''(r) + \left[\frac{2\mu}{\hbar^2}(E - U(r)) - \frac{l(l+1)}{r^2}\right]\chi(r) = 0 \tag{3.8}$$

Since $R(r)$ should be regular at the origin, the function $\chi(r)$ should vanish for $r = 0$. In all other aspects, the problem is now reduced to a one dimensional Schödinger equation for the motion in a potential

$$U_l(r) = U(r) + \frac{\hbar^2}{2\mu}\frac{l(l+1)}{r^2}\,. \tag{3.9}$$

notice that the potential does not depend on m, the eigenvalue of $\hat{L}_3$. So for any value of l we will have $l+1$ states with $m = -l, -l+1, \cdots, l$ with the same l and the same energy. States with different l will generically have different energies (since the equation is different). However, in the case of the Hydrogen atom studied below, there is additional degeneracy (due to some additional symmetries) as we will see below, and there will be states with different l and the same energy..

The spectrum (eigenstates of $\hat{H}$) contains in general a discrete part of negative energies, which are bound states of the two particles and a continuum of solutions with positive energy, corresponding to the particles being unbound and propagating to infinity. We shall study only the bound states and arrange the eigenvalues of $\hat{H}$ in order of increasing magnitude, and assign them number n (starting form 0),

$$\psi_{n,l,m} = \frac{\chi_{n.l}(r)}{r}Y_{lm}(\theta, \phi)\,. \tag{3.10}$$

Then n determines the number of nodes (zeros) of the radial part of the wave function (with $r > 0$). The number n is called the radial quantum number. The number l is called azimuthal quantum number, and m is called the magnetic quantum number.[2]

---

[1] Prove this using that $\partial_t|\psi(t)\rangle = +\frac{1}{i\hbar}\hat{H}|\psi(t)\rangle$ and $\partial_t\langle\psi(t)| = -\frac{1}{i\hbar}\langle\psi(t)|\hat{H}$.

[2] The states with $l = 0, 1, 2, 3, 4, \ldots$ are usually denoted by Latin letters: $s, p, d, f, g, \ldots$

3.2 Hydrogen atom in coordinate basis

In this section we specialize to the Coulomb potential $U(r) = -\frac{e^2}{r}$ and solve for $\chi_{n,l}(r)$.

Small r limit. At small r we can neglect the potential as the $\frac{l(l+1)}{r^2}$ term dominates, and solve

$$\chi''(r) - \frac{l(l+1)}{r^2}\chi(r) \simeq 0 \tag{3.11}$$

this equation can be explicitly solved by a power $\chi(r) \simeq r^\alpha$ with

$$\alpha(\alpha-1) = l(l+1)\,. \tag{3.12}$$

This has the solutions

$$\chi(r) \simeq C_1 r^{l+1} + C_2 r^{-l}\,. \tag{3.13}$$

Note, however, that since $\chi(0) = 0$ we can only allow ourselves the first solution *i.e.*, $C_2 = 0$.

Large r limit. In this limit the dominant term is the energy and we can ignore the potential and angular momentum. Denoting $\epsilon = -\frac{2\mu E}{\hbar^2}$, the equation is

$$\chi''(r) - \epsilon\chi(r) \simeq 0\,. \tag{3.14}$$

The general solution to this equation is

$$\chi(r) = c_1 e^{-r\sqrt{\epsilon}} + c_2 e^{+r\sqrt{\epsilon}}\,. \tag{3.15}$$

Again we need to set $c_2 = 0$, or else the wave function will be non-normalizable (*i.e.*, not in the Hilbert space). The conclusion of this is that for small r we have $\chi \sim r^{l+1}$ and for large r we must have $\chi \sim e^{-r\sqrt{\epsilon}}$.

Ansatz for the exact solution. Based on these facts we take the ansatz

$$\chi(r) = e^{-\sqrt{\epsilon}r} r^{l+1} v(r) \tag{3.16}$$

where $v(r)$ should go to a constant for small r and should not grow too fast at infinity. Using that

$$\begin{aligned} \chi'(r) &= e^{-\sqrt{\epsilon}r} r^l \left((-r\sqrt{\epsilon} + l + 1)v(r) + rv'(r)\right)\,, \\ \chi''(r) &= e^{-\sqrt{\epsilon}r} r^{l-1} \left((r^2\epsilon - 2(l+1)r\sqrt{\epsilon} + l(l+1))v(r) + 2r(-r\sqrt{\epsilon} + l + 1)v'(r) + r^2 v''(r)\right)\,, \end{aligned} \tag{3.17}$$

we get the equation

$$v''(r) + 2\left(\frac{l+1}{r} - \sqrt{\epsilon}\right)v'(r) + \left(\frac{e^2\mu}{\hbar^2} - (l+1)\sqrt{\epsilon}\right)\frac{2v(r)}{r} = 0\,. \tag{3.18}$$

Representing $v(r)$ by its Taylor series

$$v(r) = \sum_{k=0}^{\infty} C_k r^k \tag{3.19}$$

we get a simple recursion relation on C_k

$$C_{k+1} = 2C_k \frac{\sqrt{\epsilon}(k+l+1) - \frac{e^2\mu}{\hbar^2}}{(k+1)(k+2l+2)} \,. \tag{3.20}$$

There are two cases to consider: If $v(r)$ is a polynomial of degree k_0, then $C_k = 0$ for $k > k_0$. Otherwise all $C_k \neq 0$ and we can look at the large k asymptotics, where we can approximate the recursion as

$$\frac{C_{k+1}}{C_k} \sim \frac{2\sqrt{\epsilon}}{k} \,. \tag{3.21}$$

This can be solved explicitly

$$C_k \sim \frac{2^k \epsilon^{k/2}}{(k-1)!} \tag{3.22}$$

which obviously gives $v(r) \sim r e^{2\sqrt{\epsilon} r}$. This grows exponentially and leads to a non-normalizable solution. Thus we must assume $C_{k_0+1} = 0$

$$0 = C_{k_0+1} = 2C_{k_0} \frac{\sqrt{\epsilon}(k_0+1+l) - \frac{e^2\mu}{\hbar^2}}{(k_0+1)(k_0+2l+2)} \,. \tag{3.23}$$

which gives

$$\sqrt{\epsilon} = \frac{e^2\mu}{\hbar^2(k_0+l+1)} \tag{3.24}$$

or

$$E = -\frac{e^4\mu}{2\hbar^2(k_0+l+1)^2} \,, \qquad k_0 = 0,1,\dots\,, \quad l = 0,1,\dots \tag{3.25}$$

We call $n = k_0 + l + 1$ and the allowed values of E are therefore

$$E_n = -\frac{e^4\mu}{2\hbar^2 n^2} \,, \qquad n = 1,\dots, \quad l = 0,\dots n-1, \quad m = -l,\dots l \,. \tag{3.26}$$

The number of states of energy E_n is

$$\sum_{l=0}^{n-1} \sum_{m=-l}^{l} 1 = \sum_{l=0}^{n-1} (2l-1) = n^2 \,. \tag{3.27}$$

This is the famous spectrum of the Hydrogen atom found by Bohr. Plugging in the values of $\hbar$, and the masses and charges, one finds

$$E_n \approx -\frac{1}{n^2} 13.605 Ev \approx -\frac{1}{n^2} 2.1798 \cdot 10^{-18} J \,. \tag{3.28}$$

In fact equation (3.18) is closely related to the Laguerre differential equation[3] and its polynomial solutions are known as Laguerre polynomials $L^{2l+1}_{n-l-1}(\rho)$, so

$$\psi_{n.l,m} = \exp\left(-\frac{e^2\mu r}{\hbar^2 n}\right) r^l \, L^{2l+1}_{n-l-1}\left(\frac{2e^2\mu}{\hbar^2 n} r\right) \, Y_{l,m}(\theta,\phi)\,, \tag{3.29}$$

(the e in this equation is the charge of the electron, not $\exp(1)$). As discussed in the previous section, the spherical harmonics $Y_{l,m}$ are solutions of the Legendre differential equation times $\exp(im\phi)$. So we have a fully explicit solution to the problem of the Hydrogen atom.

3.3 Hydrogen atom from algebra

Conservation of angular momentum, means that the classical motion lies in a plane, whose normal is defined by $\vec{L}$. Classically a bound orbit is in general an ellipse, which has another fixed direction, namely the major axis. So one might suspect that there was another conserved vector which points in that direction. There is indeed such a vector, the Runge-Lenz vector.

We will see that an analog of this invariant can be also defined in quantum mechanics. We will show that very exceptionally for the case of the Coulomb potential $U(r) = -e^2/r$ there is an additional commuting observable, *i.e.*, an operator which computes with $\hat{H}$ which is functionally independent from $\hat{L}^2$ and $\hat{L}_i$. This is so-called Runge-Lenz vector.

Definition RUNGE-LENZ VECTOR

$$\vec{K} = \frac{\vec{p}\times\vec{L} - \vec{L}\times\vec{p}}{2\mu} - e^2\frac{\vec{r}}{r} \tag{3.30}$$

we should see that it commutes with the Hamiltonian (for simplicity we will assume $\hbar = 1$, also we omit hats over the operators)

$$H = +\frac{p^2}{2\mu} - \frac{e^2}{r} \tag{3.31}$$

Theorem 3.3.1 *The commutator of L_i with any vector operator A_j is always the same*

$$[L_i, A_j] = i\epsilon_{ijk}A_k \tag{3.32}$$

where A_j could be x_j, p_j, L_j or K_j.

For each of these cases this is easy to check by explicit calculation, however, this is simply a consequence of the transformation of any vector under rotation. See Problem Sheet 1 for more details.

---

[3] To get it in the usual form, one should replace r with $\rho = \frac{2e^2\mu}{\hbar^2 n} r$.

Theorem 3.3.2

$$[\vec{K}, H] = 0 \tag{3.33}$$

Proof The first term in $\vec{K}$ commutes with the first term in H, since p^2 is a scalar (*i.e.*, compute with the generator of rotations $\vec{L}$).[4] The main strategy in computing the commutators is to move all p_i to the right. One of the two nontrivial commutators is

$$\begin{aligned}\frac{1}{2\mu}\left[\vec{p}\times\vec{L}-\vec{L}\times\vec{p}, -\frac{e^2}{r}\right] &= -\frac{1}{2\mu}\left[\vec{p}, \frac{e^2}{r}\right]\times\vec{L}+\frac{1}{2\mu}\vec{L}\times\left[\vec{p}, \frac{e^2}{r}\right] = \frac{ie^2}{2\mu r^3}\left(\vec{L}\times\vec{x}-\vec{x}\times\vec{L}\right)\\ &= \frac{ie^2}{\mu r^3}\left(i\vec{x}+r^2\vec{p}-\vec{x}(\vec{x}\cdot\vec{p})\right)\end{aligned} \tag{3.34}$$

where we use that[5] $[\vec{L}, r^n] = 0$ (again r is invariant under rotations). Another one is

$$\left[-e^2\frac{\vec{r}}{r}, \frac{p^2}{2\mu}\right] = -\frac{e^2}{2\mu}\left[\frac{\vec{r}}{r}, p^2\right] = -\frac{ie^2}{\mu r^3}\left(i\vec{x}+r^2\vec{p}-\vec{x}(\vec{x}\cdot\vec{p})\right) \tag{3.35}$$

Combining them together we get zero. In general these calculations are quite straightforward and boring. For this kind of calculations God gave us computers. We discuss a short *Mathematica* program which helps a lot below. ∎

Two other important relations are

$$[K_i, K_j] = -\frac{2H}{\mu}i\epsilon_{ijk}L_k \tag{3.36}$$

$$\vec{K}\cdot\vec{K}-\frac{2H}{\mu}(\vec{L}\cdot\vec{L}+1) = e^4 \tag{3.37}$$

we can easily check this with *Mathematica.* Finally one can show that the Runge-Lenz vector is "orthogonal" to the angular momentum as an operator:

$$\vec{K}\cdot\vec{L} = \vec{L}\cdot\vec{K} = 0\,. \tag{3.38}$$

Thus the six operators $\vec{L}$ and $\vec{K}$ close under commutation, forming a larger algebra than $SU(2)$. The structure of this algebra can be exposed after a couple of redefinitions. First let us define

$$\vec{K'} \equiv \vec{K}\sqrt{-\frac{2H}{\mu}} \tag{3.39}$$

(since the discreet spectrum is negative the square root is well defined). Equations (3.36) and (3.37) look nicer with this redefinition

$$\begin{aligned}[K'_i, K'_j] &= i\epsilon_{ijk}L_k\\ \vec{K'}\cdot\vec{K'}+\vec{L}\cdot\vec{L}+1 &= -\frac{\mu e^4}{2H}\end{aligned} \tag{3.40}$$

---

[4] ✓Check that $[\vec{L}, p^2] = 0$.

[5] ✓Check that $[\vec{L}, r^2] = 0$ and then conclude about $[\vec{L}, r^n]$.

A further redefinition,

$$\vec{J}^{(1)} \equiv \frac{1}{2}(\vec{L} + \vec{K}')\,, \quad \vec{J}^{(2)} \equiv \frac{1}{2}(\vec{L} - \vec{K}') \tag{3.41}$$

leads to the decoupled commutation relations

$$\begin{aligned} [J_i^{(1)}, J_j^{(1)}] &= i\epsilon_{ijk} J^{(1)} \\ [J_i^{(2)}, J_j^{(2)}] &= i\epsilon_{ijk} J^{(2)} \\ [J_i^{(2)}, J_j^{(1)}] &= 0 \end{aligned} \tag{3.42}$$

The operators $\vec{J}^{(1)}$ and $\vec{J}^{(2)}$ each separately satisfy an $SU(2)$ algebra. The algebra can be characterized as $SU(2) \times SU(2)$ or $SO(4)$ – the group of rotations in 4D.

(3.38) in the new notations gives

$$(\vec{J}^{(1)})^2 - (\vec{J}^{(2)})^2 = 0\,. \tag{3.43}$$

So if the eigenvalue of $(\vec{J}^{(1)})^2$ is denoted $j_1(j_1+1)$ and likewise j_2, then this equation implies $j_1 = j_2 \equiv j$. Then from (3.40)

$$-\frac{\mu e^4}{2H} = \vec{K}'\cdot\vec{K}' + \vec{L}\cdot\vec{L} + 1 = 2\vec{J}^{(1)}\cdot\vec{J}^{(1)} + 2\vec{J}^{(2)}\cdot\vec{J}^{(2)} + 1 = 1 + 4j(j+1) = (2j+1)^2 \equiv n^2 \tag{3.44}$$

where the integer $2j+1$ we denoted by n.

It is also easy to see that the degeneracy is n^2.

Chapter 5

The Hydrogen Atom

Let us put what we learnt above to use in solving for an electron moving around a positively charged nucleus. Here we can be a bit general and consider it as a two-body problem where there is the position of the electron $\underline{r}_e$ and nucleus $\underline{r}_n$. Thus our wavefunction is

$$\Psi(t, \underline{r}_e, \underline{r}_n) = e^{-iEt/\hbar}\psi(\underline{r}_e, \underline{r}_n) \tag{5.1}$$

Furthermore the potential arising from the electrostatic force between them is

$$V = -\frac{Ze^2}{|\underline{r}_n - \underline{r}_e|} \tag{5.2}$$

where Z is the nucleon number and e the charge of an electron in appropriate units. Thus the time dependent Schrödinger equation is

$$E\psi = -\frac{\hbar^2}{2m_e}\nabla^2_{\underline{r}_e}\psi - \frac{\hbar^2}{2m_n}\nabla^2_{\underline{r}_n}\psi - \frac{Ze^2}{|\underline{r}_n - \underline{r}_e|}\psi \tag{5.3}$$

where m_e and m_n are the masses of the electron and nucleon respectively.

5.1 The Quantum 2-Body Problem

Thus we need to solve for a wavefunction of six coordinates. As in classical mechanics we can exploit the fact that the potential only depends on $\underline{r}_n - \underline{r}_e$ to reduce the problem to two one-body problems:

$$\psi(\underline{r}_e, \underline{r}_n) = \psi_{CoM}(\underline{R})\psi_{rel}(\underline{r}_{12}) \tag{5.4}$$

for suitable choices of $\underline{R}$ and $\underline{r}_{12}$ namely:

$$\underline{r}_{12} = \underline{r}_e - \underline{r}_n \qquad \underline{R} = \frac{m_e\underline{r}_e + m_n\underline{r}_n}{m_e + m_n} \tag{5.5}$$

In particular it is a small exercise (see the problem set) to show that

$$-\frac{\hbar^2}{2m_e}\nabla^2_{\underline{r}_e} - \frac{\hbar^2}{2m_n}\nabla^2_{\underline{r}_n} = -\frac{\hbar^2}{2M}\nabla^2_{\underline{R}} - \frac{\hbar^2}{2\mu}\nabla^2_{\underline{r}_{12}} \tag{5.6}$$

where $M = m_e + m_n$ and $\mu = m_e m_n / M$.

Thus our time-independent Schrödinger equation is

$$E\psi = -\frac{\hbar^2}{2M}\nabla^2_{\underline{R}}\psi - \frac{\hbar^2}{2\mu}\nabla^2_{\underline{r}_{12}}\psi - \frac{Ze^2}{|\underline{r}_{12}|}\psi \tag{5.7}$$

Now we can use seperation of variables and write

$$\psi(\underline{r}_e, \underline{r}_n) = \psi_{CoM}(\underline{R})\psi_{rel}(\underline{r}_{12}) \tag{5.8}$$

and hence find two eigenvalue equations

$$\begin{aligned} \epsilon_{CoM}\psi_{CoM} &= -\frac{\hbar^2}{2M}\nabla^2_{\underline{R}}\psi_{CoM} \\ \epsilon_{rel}\psi_{rel} &= -\frac{\hbar^2}{2\mu}\nabla^2_{\underline{r}_{12}}\psi_{rel} - \frac{Ze^2}{|\underline{r}_{12}|}\psi_{rel} \end{aligned} \tag{5.9}$$

with $E = \epsilon_{CoM} + \epsilon_{rel}$.

5.2 The Hydrogen Atom

The centre of mass part is trivial to solve:

$$\psi_{CoM} = e^{i\underline{k}\cdot\underline{R}}\chi(\underline{k}) \tag{5.10}$$

with $\epsilon_{CoM} = \hbar^2|\underline{k}|^2/2M$. These just describe a free particle, the atom, in a basis where the linear momentum is fixed. We can then construct wave packets through

$$\psi_{CoM}(\underline{R}) = \int \frac{d^3k}{(2\pi)^3} e^{i\underline{k}\cdot\underline{R}}\chi(\underline{k}) \tag{5.11}$$

which will not be energy eigenstates but can be localized in position to an arbitrary accuracy.

The more interesting part is to solve for $\psi_{rel}(\underline{r}_{12})$. Here we can switch from $\underline{r}_{12}$ to spherical coordinates and use separation of variables yet again:

$$\psi_{rel} = u_{\mathfrak{l}}(r) Y_{\mathfrak{l},\mathfrak{m}}(\theta, \phi) \tag{5.12}$$

where $r = |\underline{r}_{12}|$ *etc.*. We already know what the $Y_{\mathfrak{l},\mathfrak{m}}$'s are and we know that $u_{\mathfrak{l}}$ satisfies:

$$-\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{du_{\mathfrak{l}}}{dr}\right) + \frac{2\mu}{\hbar^2}\left(-\frac{Ze^2}{r} + \frac{\hbar^2\mathfrak{l}(\mathfrak{l}+1)}{2\mu r^2} - \epsilon_{rel}\right)u_{\mathfrak{l}} = 0 \tag{5.13}$$

So solving the Hydrogen atom (which corresponds to $Z = 1$ but we can be more general without causing any more pain) comes down to solving this equation and finding $u_{\mathfrak{l}}$, E and ψ_{rel} and $\psi(\underline{r}_e, \underline{r}_n)$.

To continue we write

$$u_{\mathfrak{l}}(r) = r^{\mathfrak{l}} f_{\mathfrak{l}}(r) \tag{5.14}$$

to find

$$-\frac{d^2 f_\mathfrak{l}}{dr^2} - \frac{2(l+1)}{r}\frac{df_\mathfrak{l}}{dr} - \frac{2\mu Z e^2}{\hbar^2 r} f_\mathfrak{l} - \frac{2\mu\epsilon_{rel}}{\hbar^2} f_\mathfrak{l} = 0 \tag{5.15}$$

This substitution is relatively standard in spherically symmetric examples as it removes the $\mathfrak{l}(\mathfrak{l}+1)$ term from the equation.

To continue we look at the large r limit where only the first and last terms are important. If $\epsilon_{rel} > 0$ then we find oscillating solutions:

$$f \sim C_1 \cos(-\sqrt{2\mu\epsilon_{rel}}\, r/\hbar) + C_2 \sin(\sqrt{2\mu\epsilon_{rel}}\, r/\hbar) \tag{5.16}$$

which will not be normalisable. Solutions with $\epsilon_{rel} = 0$ will also not be normaliable. Thus we conclude that $\epsilon_{rel} < 0$. In this case we expect solutions, at large r, to be of the form

$$f \sim C_1 e^{-\sqrt{-2\mu\epsilon_{rel}}\, r/\hbar} + C_2 e^{\sqrt{-2\mu\epsilon_{rel}}\, r/\hbar} \tag{5.17}$$

where now we must set $C_2 = 0$ to find a normalizable solution.

Thus our last step is to write $f_\mathfrak{l} = P_\mathfrak{l}(r) e^{-\sqrt{-2\mu\epsilon_{rel}}\, r/\hbar}$ which leads to

$$-\frac{d^2 P_\mathfrak{l}}{dr^2} - \frac{2(l+1)}{r}\frac{dP_\mathfrak{l}}{dr} - \frac{2\mu Z e^2}{\hbar^2 r} P_\mathfrak{l} + \frac{2}{\hbar}\sqrt{-2\mu\epsilon_{rel}}\frac{dP_\mathfrak{l}}{dr} + \frac{2(\mathfrak{l}+1)\sqrt{-2\mu\epsilon_{rel}}}{\hbar r} P_\mathfrak{l} = 0 \tag{5.18}$$

This probably doesn't look much better but it is! Indeed it has polynomial solutions. To see this let us assume we have a polynomial solution and look at the highest order term $P_\mathfrak{l} \sim r^\mathfrak{n} + \ldots$. The differential equation has a leading order $r^{\mathfrak{n}-1}$ term with coefficient

$$-\frac{2\mu Z e^2}{\hbar} + 2\sqrt{-2\mu\epsilon_{rel}}\,\mathfrak{n} + 2(\mathfrak{l}+1)\sqrt{-2\mu\epsilon_{rel}} = 0 \tag{5.19}$$

which must vanish. We can rearrange this to give

$$\epsilon_{rel} = -\frac{\mu Z^2 e^4}{2\hbar^2}\left(\frac{1}{\mathfrak{n}+\mathfrak{l}+1}\right)^2 \tag{5.20}$$

Let us return to the differential equation which we now write as

$$\frac{d^2 P_\mathfrak{l}}{dr^2} + \frac{2(l+1)}{r}\frac{dP_\mathfrak{l}}{dr} + \frac{2\mu Z e^2/\hbar^2}{r} P_\mathfrak{l} - 2\frac{\mu Z e^2/\hbar^2}{\mathfrak{n}+\mathfrak{l}+1}\frac{dP_\mathfrak{l}}{dr} - 2(\mathfrak{l}+1)\frac{\mu Z e^2/\hbar^2}{\mathfrak{n}+\mathfrak{l}+1}\frac{1}{r} P_\mathfrak{l} = 0 \tag{5.21}$$

next introduce $\rho = 2\mu Z e^2 r/\hbar^2$ to find

$$(\mathfrak{n}+\mathfrak{l}+1)\left(\frac{d^2 P_\mathfrak{l}}{d\rho^2} + \frac{2(l+1)}{\rho}\frac{dP_\mathfrak{l}}{d\rho} + \frac{1}{\rho}P_\mathfrak{l}\right) - \frac{dP_\mathfrak{l}}{d\rho} - \frac{(\mathfrak{l}+1)}{\rho}P_\mathfrak{l} = 0 \tag{5.22}$$

This is more amenable to solving through an expansion

$$P_{\mathfrak{n},\mathfrak{l}} = \sum_{k=0}^{\mathfrak{n}} C_k \rho^k \tag{5.23}$$

Substituting back in simply leads to a recursion relation obtained from setting the coefficient of ρ^{k-2} to zero:

$$\begin{aligned}(\mathfrak{n}+\mathfrak{l}+1)(k(k-1)+2k(\mathfrak{l}+1))C_k + ((\mathfrak{n}+\mathfrak{l}+1)-(k-1)-(\mathfrak{l}+1))\,C_{k-1} = 0\\ \Longrightarrow C_{k-1} = -\frac{(\mathfrak{n}+\mathfrak{l}+1)(k(k-1)+2k(\mathfrak{l}+1))}{\mathfrak{n}-k+1}C_k\end{aligned} \tag{5.24}$$

We see that taking $k=0$ leads to $C_{-1}=0$ and hence the series indeed terminates giving a polynomial. These polynomials are known as Laguerre Polynomials. We haven't shown here that these are the only normalizable solutions but this does turn out to be the case.

In summary our solutions are of the form

$$\begin{aligned}\psi(\underline{r}_e,\underline{r}_n) &= \psi_{CoM}(\underline{R})\psi_{rel}(\underline{r}_e-\underline{r}_n)\\ \psi_{rel}(\underline{r}_e-\underline{r}_n) &= r^l e^{-\frac{\mu Z e^2 r}{(\mathfrak{n}+\mathfrak{l}+1)\hbar^2}} P_{\mathfrak{n},\mathfrak{l}}(r)\Theta_{\mathfrak{l}}(\cos\theta)e^{i\mathfrak{m}\phi}\end{aligned} \tag{5.25}$$

and ψ_{CoM} is a generic free wave packet. It's fun to plot $|\psi_{rel}|^2$ for various choices of $\mathfrak{n},\mathfrak{l}$ and $\mathfrak{m}$. For example look here `https://en.wikipedia.org/wiki/Atomic_orbital#/media/File:Hydrogen_Density_Plots.png` On the other hand there are actual pictures of atoms such as here: `https://www.nature.com/articles/498009d`

We can also reproduce the famous formula postulated by Bohr in the earliest days of Quantum Mechanics (before all this formalism that we have discussed was formulated):

$$E_{Bohr} = -\frac{\mu Z^2 e^4}{2\hbar^2}\frac{1}{N^2} \tag{5.26}$$

for some integer $N=\mathfrak{n}+\mathfrak{l}+1=1,2,3,...,$. In particular since a proton is 2000 times more massive than an electron we have $m_n >> m_e$ and so $\mu = m_n$ to very high accuracy. Thus one finds (for Hydrogen where $Z=1$)

$$E_{Bohr} \sim -\frac{13.6eV}{N^2} \sim -\frac{2.18\times 10^{-18}J}{N^2} \tag{5.27}$$

This in turn also leads to the Rydberg formula for the energy spectrum of light coming from excited atoms - known as spectral lines. In particular the energy of a photon (a single particle of light) corresponds to the energy lost when an electron drops from a higher orbital to a lower one ($N_1 > N_2$):

$$\begin{aligned}E_{photon} &= -\frac{m_p Z^2 e^4}{2\hbar^2}\frac{1}{N_1^2} + \frac{m_p Z^2 e^4}{2\hbar^2}\frac{1}{N_2^2}\\ &= R\left(\frac{1}{N_2^2}-\frac{1}{N_1^2}\right)\end{aligned} \tag{5.28}$$

where we have computed the Rydberg constant $R = m_n Z^2 e^2/2\hbar^4$

We can also ask how many states there are at each energy level:

- $N=1$: $\mathfrak{n}=\mathfrak{l}=0$ 1 state

- $N = 2$: $\mathfrak{n} = 1, \mathfrak{l} = 0$ or $\mathfrak{n} = 0, \mathfrak{l} = 1$ 1+3=4 states
- $N = 3$: $\mathfrak{n} = 2, \mathfrak{l} = 0$ or $\mathfrak{n} = 1, \mathfrak{l} = 1$ or $\mathfrak{n} = 0, \mathfrak{l} = 2$ 1+3+5 = 9 states

In fact we find N^2 states for each energy level (the sum of the first N odd numbers is N^2).

However what are the lowest energy states with multiple electrons. In fact there are twice as many as these as each electron can be spin up or spin down (more on this later). This is just an additional quantum number (which only takes two values: up or down) and corresponds to the fact that the correct rotation group of Nature isn't $SO(3)$ but $SU(2) = Spin(3)$. Note also that the terminology of spin up and spin down has no physical interpretation in terms of up and down: if you take an electron of spin up to Australia it does not become spin down. It has more to do with the fact that we write vectors as column matrices and the up state are on top and then down state on the bottom. We must know another fact: no two electrons can be in the same state (the multi-electron wavefunction must be anti-symmetric - odd - under interchange of any two electrons). Thus the degeneracies of low energy multi-electron states are (this ignores inter electron interactions which become increasingly important)

- 2: $\mathfrak{n} = l = 0$
- 8: 2 states with $\mathfrak{n} = 1, \mathfrak{l} = 0$ and 6 from $\mathfrak{n} = 0, \mathfrak{l} = 1$
- 18: 2 from $\mathfrak{n} = 2$ 6 from $\mathfrak{n} = 1, \mathfrak{l} = 1$ and 10 from $\mathfrak{n} = 0, \mathfrak{l} = 2$,

This pattern is evident in the periodic table whose rows have $2, 8$ and 18 elements. Which we have now predicted based on some crazy idea that states are vectors in a Hilbert space and observables are self-adjoint linear maps.

5 Perturbation theory

Few problems in quantum mechanics — with either time-independent or time-dependent Hamiltonians — can be solved exactly. Inevitably we are forced to resort to some form of approximation method.
We are going to consider systems that have a Hamiltonian:

$$H = H_0 + V$$

where H_0 is the Hamiltonian of the unperturbed and $V = f(\varepsilon)$ is the potential describing the perturbation. We will assume V is independent of time. Furthermore, let us assume that we can solve the time-independent Schrödinger equation for H i.e. that we know its eigenvalues and eigenfunctions:

$$H_0 \left|E_n^{(0)}\right\rangle = E_n^{(0)} \left|E_n^{(0)}\right\rangle$$

Since ε is a small parameter we expain in terms of ε:

$$\begin{aligned} H &= H^{(0)} + \varepsilon H^{(1)} + \varepsilon^2 H^{(2)} + \cdots \\ E_n &= E_n^{(0)} + \varepsilon E_n^{(1)} + \varepsilon^2 E_n^{(2)} + \cdots \\ |E_n\rangle &= \left|E_n^{(0)}\right\rangle + \varepsilon \left|E_n^{(1)}\right\rangle + \varepsilon^2 \left|E_n^{(2)}\right\rangle + \cdots \end{aligned}$$

We want to solve the equation:

$$H |E_n\rangle = E_n |E_n\rangle$$

by substituting the expansions we have from above and obtain

$$\left(H^{(0)} + \varepsilon H^{(1)} + \varepsilon^2 H^{(2)} + \cdots\right) = \left(E_n^{(0)} + \varepsilon E_n^{(1)} + \varepsilon^2 E_n^{(2)} + \cdots\right) \left(\left|E_n^{(0)}\right\rangle + \varepsilon \left|E_n^{(1)}\right\rangle + \varepsilon^2 \left|E_n^{(2)}\right\rangle + \cdots\right).$$

There are many cases to analyse, we begin with non-degenerate cases (i.e. all the eigenvalues are distinct).

5.1 Non-degenerate: first order

We consider the expansion:

$$\left(H^{(0)} + \varepsilon H^{(1)} + \cdots\right) = \left(E_n^{(0)} + \varepsilon E_n^{(1)} + \cdots\right) \left(\left|E_n^{(0)}\right\rangle + \varepsilon \left|E_n^{(1)}\right\rangle + \cdots\right).$$

The zeroth order ε terms cancel as we have already solved the unperturbed system. Thus, ignoring all ε terms higher than the linear one we have

$$H^{(1)} \left|E_n^{(0)}\right\rangle + H^{(0)} \left|E_n^{(1)}\right\rangle = E_n^{(1)} \left|E_n^{(0)}\right\rangle + E_n^{(0)} \left|E_n^{(1)}\right\rangle.$$

Now we know $H^{(0)}, E_n^{(0)}$ and $E_n^{(1)}$ so we need to solve for $E_n^{(1)}$.

Proposition 5.1

We have that the first order energy correction is given by

$$E_n^{(1)} = \left\langle E_n^{(0)}\right| H^{(1)} \left|E_n^{(0)}\right\rangle$$

and that

$$\begin{aligned}\left|E_n^{(1)}\right\rangle &= \sum_m \left\langle E_m^{(0)}\middle|E_n^{(1)}\right\rangle \left|E_m^{(0)}\right\rangle \\ &= \sum_{m\neq n} \frac{\left\langle E_m^{(0)}\right| H^{(1)} \left|E_n^{(0)}\right\rangle}{E_n^{(0)} - E_m^{(0)}} \left|E_m^{(0)}\right\rangle .\end{aligned}$$

Remark 5.2. We can rewrite this in a more concise manner by setting $H^{(1)} = V$ and use the notation

$$V_{mn} = \left\langle E_m^{(0)}\right| V \left|E_n^{(0)}\right\rangle .$$

With this, we have that

$$E_n^{(1)} = V_{nn}$$

and

$$\left|E_n^{(1)}\right\rangle = \sum_{m\neq n} \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} \left|E_m^{(0)}\right\rangle .$$

Proof. We begin from

$$H^{(1)} \left|E_n^{(0)}\right\rangle + H^{(0)} \left|E_n^{(1)}\right\rangle = E_n^{(1)} \left|E_n^{(0)}\right\rangle + E_n^{(0)} \left|E_n^{(1)}\right\rangle .$$

We use the fact that $\left\{\left|E_n^{(0)}\right\rangle\right\}$ are an orthonormal basis. So

$$\left\langle E_m^{(0)}\right| H^{(1)} \left|E_n^{(1)}\right\rangle + \left\langle E_m^{(0)}\right| H^{(0)} \left|E_n^{(1)}\right\rangle = E_n^{(1)} \left\langle E_m^{(0)}\middle|E_n^{(0)}\right\rangle + E_n^{(0)} \left\langle E_m^{(0)}\middle|E_n^{(1)}\right\rangle$$

However we also know that $\left\langle E_m^{(0)}\right| H^{(0)} = \left(H^{(0)} \left|E_m^{(0)}\right\rangle\right)^\dagger = \left(E_m^{(0)} \left|E_m^{(0)}\right\rangle\right)^\dagger = \left\langle E_m^{(0)}\right| E_m^{(0)}$ and hence

$$\left\langle E_m^{(0)}\right| H^{(1)} \left|E_n^{(1)}\right\rangle = E_n^{(1)}\delta_{mn} + (E_n^{(0)} - E_m^{(0)}) \left\langle E_m^{(0)}\middle|E_n^{(1)}\right\rangle .$$

We can now take $n = m$ to find

$$E_n^{(1)} = \left\langle E_n^{(0)}\right| H^{(1)} \left|E_n^{(0)}.\right\rangle$$

So we know the first order correction to the energies.

We can now write, for $m \neq n$,

$$\left\langle E_m^{(0)}\middle|E_n^{(1)}\right\rangle = \frac{1}{E_n^{(0)} - E_m^{(0)}} \left\langle E_m^{(0)}\right| H \left|E_n^{(0)}\right\rangle$$

Here is where the non-degeneracy is important. On the other hand we want orthonormal states $|E_n\rangle$:

$$\begin{aligned}\delta_{mn} &= \langle E_n|E_m\rangle \\ &= \left(\left\langle E_n^{(0)}\right| + \varepsilon \left\langle E_n^{(1)}\right| + \dots\right)\left(\left|E_m^{(0)}\right\rangle + \varepsilon \left|E_m^{(1)}\right\rangle + \dots\right) \\ 0 &= \varepsilon\left(\left\langle E_n^{(1)}\middle|E_m^{(0)}\right\rangle + \left\langle E_n^{(0)}\middle|E_m^{(1)}\right\rangle\right)\end{aligned}$$

Thus we learn that

$$\langle E_n^{(1)}|E_m^{(0)}\rangle = -\left(\langle E_m^{(0)}|E_n^{(1)}\rangle\right)^* = -\langle E_m^{(0)}|E_n^{(1)}\rangle$$

In particular we can take

$$\boxed{\left\langle E_n^{(1)} \middle| E_n^{(0)} \right\rangle = \left\langle E_n^{(0)} \middle| E_n^{(1)} \right\rangle = 0}$$

This tells us that the first order correction is orthogonal to the unperturbed eignevalue. In particular we can write $\left|E_n^{(1)}\right\rangle$ in a basis of $\left|E_m^{(0)}\right\rangle$ i.e.

$$|E_n^{(1)}\rangle = \sum_m c_{mn} |E_m^{(0)}\rangle .$$

We recall that we can write

$$\mathbb{I} = \sum_m |E_m^{(0)}\rangle \langle E_m^{(0)}|$$

so multiplying both sides by $\left|E_n^{(1)}\right\rangle$ we have

$$\begin{aligned} |E_n^{(1)}\rangle &= \sum_m |E_m^{(0)}\rangle \underbrace{\langle E_m^{(0)}|E_n^{(1)}\rangle}_{c_{mn}} \\ &= \sum_m \langle E_m^{(0)}|E_n^{(1)}\rangle |E_m^{(0)}\rangle \\ &= \sum_{m\neq n} \frac{\left\langle E_m^{(0)}\middle| H^{(1)} \middle| E_n^{(0)}\right\rangle}{E_n^{(0)} - E_m^{(0)}} |E_m^{(0)}\rangle \end{aligned}$$

□

5.2 Non-degenerate: second order

We can assume that the zeroth and first order equations have been solved so we find, at second order,

$$H^{(2)}|E_n^{(0)}\rangle = H^{(1)}|E_n^{(1)}\rangle + H^{(0)}|E_n^2\rangle = E_n^{(0)}|E_n^{(2)}\rangle + E_n^{(1)}|E_n^{(1)}\rangle + E_n^{(2)}|E_n^{(0)}\rangle .$$

Proposition 5.3

The second order energy correction to E_n is given by

$$E_n^{(2)} = \langle E_n^{(0)}|H^{(2)}|E_n^{(0)}\rangle + \sum_{m\neq n} \frac{\left\langle E_m^{(0)}\middle| H^{(1)} \middle| E_n^{(0)}\right\rangle \left\langle E_n^{(0)}\middle| H^{(1)} \middle| E_m^{(0)}\right\rangle}{E_n^{(0)} - E_m^{(0)}}$$

Remark 5.4. In the nice notation we can write if we set $H^{(2)} = 0$ then:

$$E_n^{(2)} = \sum_{m \neq n} \frac{V_{mn} V_{nm}}{E_n^{(0)} - E_m^{(0)}}.$$

Proof. Remember that the unknowns are $E_n^{(2)}$ and $\left\langle E_n^{(0)} \middle| E_n^{(2)} \right\rangle$, everything else is known. So we take matrix elements again:

$$\begin{aligned} &\left\langle E_m^{(0)} \right| H^{(2)} \left| E_n^{(0)} \right\rangle + \left\langle E_m^{(0)} \right| H^{(1)} \left| E_n^{(1)} \right\rangle + \left\langle E_m^{(0)} \right| H^{(0)} \left| E_n^{(2)} \right\rangle \\ &= E_n^{(2)} \left\langle E_m^{(0)} \middle| E_n^{(0)} \right\rangle + E_n^{(1)} \left\langle E_m^{(0)} \middle| E_n^{(1)} \right\rangle + E_n^{(2)} \underbrace{\left\langle E_m^{(0)} \middle| E_n^{(0)} \right\rangle}_{\delta_{mn}}. \end{aligned}$$

Again we first look at $m = n$ and find

$$\left\langle E_n^{(0)} \right| H^{(2)} \left| E_n^{(0)} \right\rangle + \left\langle E_n^{(0)} \right| H^{(1)} \left| E_n^{(1)} \right\rangle + E_n^{(0)} \left\langle E_n^{(0)} \middle| E_n^{(2)} \right\rangle = E_n^{(0)} \left\langle E_n^{(0)} \middle| E_n^{(2)} \right\rangle + E_n^{(2)}$$

where we have used $\left\langle E_n^{(0)} \middle| E_n^{(1)} \right\rangle = 0$. The $\left\langle E_n^{(0)} \middle| E_n^{(2)} \right\rangle$ terms cancel so this tells us the second order correction to E_n:

$$\begin{aligned} E_n^{(2)} &= \left\langle E_n^{(0)} \right| H^{(2)} \left| E_n^{(0)} \right\rangle + \left\langle E_n^{(0)} \right| H^{(1)} \left| E_n^{(1)} \right\rangle \\ &= \left\langle E_n^{(0)} \right| H^{(2)} \left| E_n^{(0)} \right\rangle + \left\langle E_n^{(0)} \right| H^{(1)} \left(\sum_{m \neq n} \frac{\left\langle E_m^{(0)} \middle| H^{(1)} \middle| E_n^{(0)} \right\rangle}{E_n^{(0)} - E_m^{(0)}} \left| E_m^{(0)} \right\rangle \right) \\ &= \left\langle E_n^{(0)} \right| H^{(2)} \left| E_n^{(0)} \right\rangle + \sum_{m \neq n} \frac{\left\langle E_m^{(0)} \middle| H^{(1)} \middle| E_n^{(0)} \right\rangle \left\langle E_n^{(0)} \middle| H^{(1)} \middle| E_m^{(0)} \right\rangle}{E_n^{(0)} - E_m^{(0)}} \end{aligned}$$

Note 5.5. We multplied the $\left\langle E_n^{(0)} \right| H^{(1)}$ with $\left| E_m^{(0)} \right\rangle$ inside the bracket.

□

Proposition 5.6

We have that for $H^{(2)} = 0$:

$$\left| E_n^{(2)} \right\rangle = \sum_m c_{nm} \left| E_m^{(0)} \right\rangle$$

where for $m \neq n$

$$c_{nm} = \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} + \sum_{p \neq n} \left(\frac{V_{pm} V_{mp}}{\left(E_n^{(0)} - E_m^{(0)}\right)\left(E_n^{(0)} - E_p^{(0)}\right)} \right) - \frac{V_{nn} V_{mn}}{\left(E_n^{(0)} - E_m^{(0)}\right)\left(E_n^{(0)} - E_m^{(0)}\right)}$$

and

$$c_{nn} = -\frac{1}{2} \sum_{m \neq n} \frac{V_{mn} V_{mn}}{\left(E_n^{(0)} - E_m^{(0)}\right)^2}$$

Proof. We recall

$$\begin{aligned}&\left\langle E_m^{(0)}\right| H^{(2)} \left|E_n^{(0)}\right\rangle + \left\langle E_m^{(0)}\right| H^{(1)} \left|E_n^{(1)}\right\rangle + \left\langle E_m^{(1)}\right| H^{(1)} \left|E_n^{(0)}\right\rangle \\ &= \left\langle E_n^{(0)}\right| H^{(0)} \left|E_n^{(2)}\right\rangle + \left\langle E_n^{(1)}\right| H^{(0)} \left|E_n^{(1)}\right\rangle + \left\langle E_n^{(2)}\right| H^{(0)} \left|E_n^{(0)}\right\rangle \delta_{nm}.\end{aligned}$$

To determine $\left|E_n^{(2)}\right\rangle$ we need to know the coefficients c_{nm} where

$$\left|E_n^{(2)}\right\rangle = \sum_m c_{nm} \left|E_m^{(0)}\right\rangle$$

and we know that $c_{nm} = \left\langle E_m^{(0)}\middle|E_n^{(2)}\right\rangle$. Thus since we have a non-degenerate spectrum we so, by rearranging the first equaion, we have that for $n \neq m$,

$$\begin{aligned}c_{nm} &= \left\langle E_m^{(0)}\middle|E_n^{(2)}\right\rangle \\ &= \frac{1}{E_n^{(0)} - E_m^{(0)}} \left(\left\langle E_m^{(0)}\right| H^{(2)} \left|E_n^{(0)}\right\rangle + \left\langle E_m^{(1)}\right| H^{(1)} \left|E_n^{(1)}\right\rangle - E_n^{(1)} \left\langle E_m^{(1)}\middle|E_n^{(0)}\right\rangle\right).\end{aligned}$$

However we know that

$$\left|E_n^{(1)}\right\rangle = \sum_{p\neq n} \frac{1}{E_n^{(0)} - E_p^{(0)}} \left\langle E_p^{(0)}\right| H^{(1)} \left|E_n^{(0)}\right\rangle \left|E_p^{(0)}\right\rangle$$

so

$$\left\langle E_m^{(0)}\right| H^{(1)} \left|E_n^{(1)}\right\rangle = \sum_{p\neq n} \frac{1}{E_n^{(0)} - E_p^{(0)}} \left\langle E_m^{(0)}\right| H^{(1)} \left|E_p^{(0)}\right\rangle \left\langle E_p^{(0)}\right| H^{(1)} \left|E_n^{(0)}\right\rangle$$

We also know

$$E_n^{(1)} = \left\langle E_n^{(0)}\right| H^{(1)} \left|E_n^{(0)}\right\rangle$$

so

$$\begin{aligned}E_n^{(1)} \left\langle E_m^{(1)}\middle|E_n^{(0)}\right\rangle &= \left\langle E_n^{(0)}\right| H^{(1)} \left|E_n^{(0)}\right\rangle \sum_{p\neq n} \frac{1}{E_n^{(0)} - E_p^{(0)}} \left\langle E_p^{(0)}\right| H^{(1)} \left|E_n^{(0)}\right\rangle \delta_{mp} \\ &= \frac{1}{E_n^{(0)} - E_m^{(0)}} \left\langle E_n^{(0)}\right| H^{(1)} \left|E_n^{(0)}\right\rangle \left\langle E_m^{(0)}\right| H^{(1)} \left|E_n^{(0)}\right\rangle\end{aligned}$$

Thus substituting each term we have worked out, (for $m \neq n$)

$$\begin{aligned}c_{nm} = \frac{\left\langle E_m^{(0)}\middle| H^2 \middle|E_n^{(0)}\right\rangle}{E_n^{(0)} - E_m^{(0)}} + \sum_{p\neq n} &\left(\frac{\left\langle E_p^{(0)}\middle| H^{(1)} \middle|E_m^{(0)}\right\rangle \left\langle E_m^{(0)}\middle| H^{(1)} \middle|E_p^{(0)}\right\rangle}{(E_n^{(0)} - E_p^{(0)})(E_n^{(0)} - E_m^{(0)})}\right) \\ &- \frac{\left\langle E_n^{(0)}\middle| H^{(1)} \middle|E_n^{(0)}\right\rangle \left\langle E_m^{(0)}\middle| H^{(1)} \middle|E_n^{(0)}\right\rangle}{(E_n^{(0)} - E_m^{(}0))(E_n^{(0)} - E_m^{(0)})}\end{aligned}$$

However we still need to know c_{nn}. We fix this by asking that

$$\begin{aligned}1 &= \langle E_n|E_n\rangle \\ &= (\left\langle E_n^{(0)}\right| + \epsilon \left\langle E_n^{(1)}\right| + \epsilon^2 \left\langle E_n^{(2)}\right|)(\left|E_n^{(0)}\right\rangle + \epsilon \left|E_n^{(1)}\right\rangle + \epsilon^2 \left|E_n^{(2)}\right\rangle)\end{aligned}$$

which at order ϵ^2 gives

$$\begin{aligned}0 &= \left\langle E_n^{(0)}\middle|E_n^{(2)}\right\rangle + \left\langle E_n^{(2)}\middle|E_n^{(0)}\right\rangle + \left\langle E_n^{(1)}\middle|E_n^{(1)}\right\rangle \\ &= c_{nn} + c_{nn}^* + \left\langle E_n^{(1)}\middle|E_n^{(1)}\right\rangle\end{aligned}$$

Thus we take

$$c_{nn} = -\frac{1}{2}\left\langle E_n^{(1)} \middle| E_n^{(1)} \right\rangle = -\frac{1}{2}\sum_{p\neq n}\frac{\left|\left\langle E_p^{(0)}\right| H^{(1)} \left|E_n^{(0)}\right\rangle\right|^2}{(E_n^{(0)} - E_p^{(0)})^2}.$$

□

5.3 Degenerate perturbation theory

Proposition 5.7

In degenerate perturbation theory we have that for $m = n$

$$E_n^{(1)} = \left\langle E_n^{(0)}\right| H^{(1)} \left|E_n^{(0)}\right\rangle.$$

Suppose we have some subset $\{n', m'\}$ of degenerate energy states $E_{n'}^{(0)} = E_{m'}^{(0)}$ we have that

$$E_{n'}^{(1)}\delta_{n'm'} = \left\langle E_{m'}^{(0)}\right| H^{(1)} \left|E_{n'}^{(0)}\right\rangle.$$

Furthermore,

$$\begin{aligned} E_{n'} &= E_{n'}^{(0)} + \varepsilon\left\langle E_{n'}^{(0)}\right| H^{(1)} \left|E_{n'}^{(0)}\right\rangle + \cdots \\ &= E_{n'}^{(0)} + \varepsilon E_{n'}^{(1)} + \cdots \end{aligned}$$

where $H^{(1)}\left|E_{n'}^{(0)}\right\rangle = E_{n'}^{(1)}\left|E_{n'}^{(0)}\right\rangle$. Then the eigenvectors are

$$\left|E_{n'}^{(1)}\right\rangle = \sum_{m''}\frac{\left\langle E_{m''}^{(0)}\right| H^{(1)} \left|E_{n'}^{(0)}\right\rangle}{E_{n'}^{(0)} - E_{m''}^{(0)}}\left|E_{m''}^{(0)}\right\rangle.$$

where n'' label non-degenerate eigenstates.

5.4 Time dependent perturbation theory

In this section we assume that the Hamiltonian is mostly time-independent adn the time dependence enters through a small perturbation:

$$H = H_0 + \varepsilon H_1(t) + \cdots$$

We recall that the time-independent Schrödinger equation is given by

$$i\hbar\frac{\partial}{\partial t}\left|\Psi\right\rangle = H\left|\Psi\right\rangle$$

and is solved by

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

To solve the peturbed Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\left|\Psi\right\rangle = H^{(0)}\left|\Psi\right\rangle + \varepsilon H^{(1)}(t)\left|\Psi\right\rangle$$

we allow c_n to be a function of time.

Proposition 5.8

We can expand $c_m(t)$ as

$$c_m(t) = c_m(0) + \varepsilon c_m^{(1)}(t) + \varepsilon^2 c_m^{(2)}(t)_+ \cdots$$

with the boundary condition that $c_m^{(k)}(0) = 0$ for all $k \in \mathbb{N}$ and $c_m = \delta_{mk}$. We have that

$$c_m^{(1)} = -\frac{i}{\hbar}\int_0^t e^{i(E_m - E_k)\tau/\hbar} \left\langle \psi_m^{(0)} \right| H^{(1)}(t) \left| \psi_k^{(0)} \right\rangle \, d\tau.$$

Note 5.9. We can set the lower bound to be anything other than 0 like $-\infty$ and have $c_m^{(k)}(-\infty) = 0$ for all k.

Proof. We substitute the solution of the Schrödinger equation to

$$i\hbar\frac{\partial}{\partial t}\left|\Psi\right\rangle = H^{(0)}\left|\Psi\right\rangle + \varepsilon H^{(1)}(t)\left|\Psi\right\rangle$$

and obtain

$$\begin{aligned}\sum_n \left(i\hbar\frac{\mathrm{d}}{\mathrm{d}t}c_n(t) + E_n c_n(t)\right) e^{-iE_n t/\hbar}\left|\psi_n^{(0)}\right\rangle &= \sum_n c_n(t) e^{-iE_n t/\hbar} H^{(0)}\left|\psi_n^{(0)}\right\rangle \\ &+ \varepsilon \sum_n c_n(t) e^{-iE_n t/\hbar} H^{(1)}\left|\psi_n^{(0)}\right\rangle.\end{aligned}$$

We can cancel out the second term on the LHS and the first term on the RHS since we know that $H^{(0)}\left|\psi_n^{(0)}\right\rangle = E_n\left|\psi_n^{(0)}\right\rangle$. Therefore, we have

$$i\hbar\sum_n \frac{\mathrm{d}}{\mathrm{d}t}c_n(t) e^{-iE_n t/\hbar}\left|\psi_n^{(0)}\right\rangle = \varepsilon\sum_n c_n(t) e^{-iE_n t/\hbar} H^{(1)}\left|\psi_n^{(0)}\right\rangle.$$

Since $\left|\psi_n^{(0)}\right\rangle$ are an orthonormal basis of the Hilber space we can take teh inner product with $\left\langle\psi_m^{(0)}\right|$ and obtain

$$\begin{aligned} i\hbar\frac{\mathrm{d}}{\mathrm{d}t}c_m(t)e^{-iE_m t/\hbar} &= \varepsilon\sum_n c_n(t)e^{-iE_m t/\hbar}\left\langle\psi_m^{(0)}\right| H^{(1)}(t)\left|\psi_n^{(0)}\right\rangle \\ \frac{\mathrm{d}}{\mathrm{d}t}c_m(t) &= -\frac{i\varepsilon}{\hbar}\sum_n c_n(t)e^{i(E_m-E_n)t/\hbar}\left\langle\psi_m^{(0)}\right| H^{(1)}(t)\left|\psi_n^{(0)}\right\rangle.\end{aligned}$$

Thus we obtain a first order differential equation for each $c_n(t)$. □

Proposition 5.10

We have that $\left|c_m(t)\right|^2$ gives the probability that after the perturbation the system will lie in the m-th energy state ($m \neq k$) at tiem t. We call $c_m^{(1)}(t)$ the **first order transition amplitude**.

> **Remark 5.11.** We do not expect $\sum_n |c_n(t)|^2 = 1$ since the perturbation has introduced (or taken away) energy from the system which then gets redistributed.

5.5 Degenerate perturbation theory

In the previous sections we discussed the approach to perturbation theory for distinct eigenvalues. In this section, we discuss when some (or all) are not distinct. We only consider in first-order:

$$\widehat{H}_0\psi_n^{(1)}(x) + \widehat{V}\psi_n^{(0)}(x) = E_n^{(0)}\psi_n^{(1)}(x) + E_n^{(1)}\psi_n^{(0)}(x)$$

and, we take the inner product with $\langle \psi_m^0|$:

$$\left\langle \psi_m^{(0)}\right| \widehat{H}_0 \left|\psi_n^{(1)}\right\rangle + \left\langle \psi_m^{(0)}\right| \widehat{V} \left|\psi_n^{(0)}\right\rangle = E_n^{(0)} \left\langle \psi_m^{(0)}\middle|\psi_n^{(1)}\right\rangle + E_n^{(1)} \left\langle \psi_m^{(0)}\middle|\psi_n^{(0)}\right\rangle .$$

Proposition 5.12. If $m = n$ we have (the same as before):

$$E_n^{(1)} = \left\langle E_n^{(0)}\right| \widehat{V} \left|E_n^{(0)}\right\rangle .$$

Let us assume that we have that we 2 degenerates states labelled by a and b i.e. $E_a^{(0)} = E_b^{(0)}$. We need

$$\left\langle \psi_b^{(0)}\right| \widehat{V} \left|\psi_a^{(0)}\right\rangle = E_a^{(1)}\delta_{ab}$$

where we need to choose $\left|\psi_a^{(0)}\right\rangle$ to be an eigenstate of $\widehat{V}$ with eigenvalue $E_a^{(1)}$. We then find that the degeneracy is lifted by the eigenvalues of $\hat{V}$ in the degenerate subspace

$$E_{n'} = E_{n'}^{(0)} + \varepsilon \left\langle E_{n'}^{(0)}\right| \hat{V} \left|E_{n'}^{(0)}\right\rangle + \dots$$

$$= E_{n'}^{(0)} + \varepsilon E_{n'}^{(1)} + \dots$$

where $\hat{V}\left|E_{n'}^{(0)}\right\rangle = E_{n'}^{(1)}\left|E_{n'}^{(0)}\right\rangle$. Whereas for non-degenerate eigenvalues we find

$$E_{n'} = E_{n'}^{(0)} + \varepsilon \left\langle E_{n'}^{(0)}\right| \hat{V} \left|E_{n'}^{(0)}\right\rangle + \dots$$

where n' labels the non-degenerate eigenstates. We then solve for the eigenvectors by

Proposition 5.13

$$\left|E_{n'}^{(1)}\right\rangle = \sum_{m'\neq n'} \frac{\left\langle E_{m'}^{(0)}\right| \hat{V} \left|E_{n'}^{(0)}\right\rangle}{E_{n'}^{(0)} - E_{m'}^{(0)}} \left|E_{m'}^{(0)}\right\rangle$$

$$\left|E_{n'}^{(1)}\right\rangle = \sum_{m'} \frac{\left\langle E_{m'}^{(0)}\right| \hat{V} \left|E_{n'}^{(0)}\right\rangle}{E_{n'}^{(0)} - E_{m'}^{(0)}} \left|E_{m'}^{(0)}\right\rangle$$

6 Semi-classical quantisation

Definition 6.1. The **WKB** (Wentzel, Kramers, Brillouin) method is a technique for obtaining approximate solutions to the time-independent Schrödinger equation in one dimension.

We will consider wave-functions of the form

$$\psi(x) = e^{\frac{i}{\hbar}\sigma(x)},$$

where $\sigma(x)$ has units in $\hbar$. Using

$$\begin{aligned}\psi'(x) &= \frac{i}{\hbar}\sigma'(x)e^{\frac{i}{\hbar}\sigma(x)}\\ \psi''(x) &= \frac{i}{\hbar}\sigma''(x)e^{\frac{i}{\hbar}\sigma(x)} - \frac{(\sigma'(x))^2}{\hbar^2}e^{\frac{i}{\hbar}\sigma(x)}\end{aligned}$$

We can rewrite the Schrödinger equation: $-\frac{\hbar^2}{2m}\psi'' + (U(x) - E)\psi = 0$ as

$$\frac{1}{2m}\left((\sigma')^2 - i\hbar\sigma''\right) = E - U(x).$$

The equation above can be solved by a formal power series in $\hbar$:

$$\sigma = \sigma_0 + \hbar\sigma_1 + \hbar^2\sigma_2 + \cdots$$

Plugging the expansion into the Schrödinger equations, and we have

$$(\sigma_0' + \hbar\sigma_1' + \hbar^2\sigma_2' + \cdots)^2 - i\hbar\left(\sigma_0' + \hbar\sigma_1' + \hbar^2\sigma_2' + \cdots\right) = 2m(E - U(x)).$$

6.1 Zeroth order

Omitting all terms containing $\hbar$ gives

$$(\sigma_0')^2 = 2m(E - U(x)).$$

Hence, we find

$$\sigma_0 = \pm\int_{x_0}^{x} p(y)\,dy, \quad \text{where} \quad p(y) = \sqrt{2m(E - U(y))}.$$

Note that $p(x)$ is nothing but the momentum of the classical particle of energy E moving in the external potential $U(x)$:

$$E = \frac{p^2_{\text{classical}}(x)}{2m} + U(x).$$

Thus, we found the solution

$$\begin{aligned}\psi(x)_{\text{semiclassical}} &= e^{\frac{i}{\hbar}\sigma(x)}\\ &= e^{\pm\frac{i}{\hbar}\int p(y)dy}\end{aligned}$$

The fact that there are two solutions is expected as the starting point is the second order differential equation. The general solution is a linear combination of these two

$$\psi(x) = Ae^{+\frac{i}{\hbar}\int p(y)dy} + Be^{-\frac{i}{\hbar}\int p(y)dy}.$$

The coefficients A and B should be selected to satisfy the boundary conditions. We will discuss some particular examples below.

6.2 Validity of the approximation

We should have that

$$\hbar\sigma'' \ll (\sigma')^2.$$

We see that

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\sigma_0 = \pm\frac{1}{2}\sqrt{\frac{2m}{E-V}}\frac{\mathrm{d}V}{\mathrm{d}x} \quad \text{and} \quad \left(\frac{\mathrm{d}}{\mathrm{d}x}\sigma\right)^2 = 2m(E-V).$$

Assuming V iw well behaved i.e. $\frac{\mathrm{d}V}{\mathrm{d}x}$ is bounded, then we find the semi-classical approximation break down at the **'turning points'** where $V = E$. Here the momentum p vanishes so nothing can be smaller than it.

6.3 First order

Recalling the Schrödinger equation:

$$\begin{aligned}(\sigma_0' + \hbar\sigma_1' + \hbar^2\sigma_2' + \cdots)^2 - i\hbar\left(\sigma_0' + \hbar\sigma_1' + \hbar^2\sigma_2' + \cdots\right) - 2m(E-U(x)) = 0\\ (\sigma_0')^2 - 2m(E-U(x)) + \hbar(2\sigma_0'\sigma_1' - i\sigma_0'') + O(h^2) = 0\end{aligned}$$

Now we keep all the terms up to $\hbar$:

$$2\sigma_0'\sigma_1' - i\sigma_0'' = 0$$

equivalently

$$\begin{aligned}\sigma_1' &= \frac{i}{2}\frac{\sigma_0''}{\sigma_0'}\\ &= \frac{i}{2}\frac{p'}{p}\\ &= \frac{i}{2}\frac{\mathrm{d}}{\mathrm{d}x}\ln(p)\end{aligned}$$

So that

$$\text{constant} + \sigma_1 = \frac{i}{2}\ln p(x)$$

We can absorb the constant in the normalisation coefficients A, B and hence, to first order

$$\boxed{\psi_{\text{WKB}} = \frac{A}{\sqrt{p(x)}}e^{+\frac{i}{\hbar}\int p(y)\,dy} + \frac{B}{\sqrt{p(x)}}e^{-\frac{i}{\hbar}\int p(y)\,dy}.}$$

6.4 Quantisation I: Particle in a box

We investigate the WKB approximation for a particle in a box, i.e.

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

Note 6.2. The wave function vanishes for $x \leq 0$ and $x \geq L$.

From the previous section we know the form of the wave function:

$$\begin{aligned}\psi(x) &= \frac{A}{\sqrt{p}} e^{+\frac{i}{\hbar}\int p\, dy} + \frac{B}{\sqrt{p}} e^{-\frac{i}{\hbar}\int p\, dy} \\ &= A' e^{\frac{i}{\hbar}px} + B' e^{-\frac{i}{\hbar}px},\end{aligned}$$

where $p = \sqrt{2mE}$. To find the coefficients we need to impose the boundary condition $\psi(0) = \psi(L) = 0$ which leads to

$$\begin{aligned}A' + B' &= 0 \\ A' e^{\frac{ipL}{\hbar}} + B' e^{-\frac{ipL}{\hbar}} &= 2A' \sin\left(\frac{pL}{\hbar}\right) = 0.\end{aligned}$$

We find that $p = \frac{n\pi\hbar}{L}$ for $n \in \mathbb{N} \setminus \{0\}$. Therefore,

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

which agrees with the exact answer.

Remark 6.3. This agrees with teh answer because $\dfrac{\mathrm{d}^2}{\mathrm{d}x^2}\sigma_0 = \dfrac{\mathrm{d}p}{\mathrm{d}x} = 0$.

6.5 Quantisation II: Smooth potentials

In this situation we have a continuous potential where we could encounter $p(x) = 0$. Near a turning point we can assume that the potential is roughly linear. Before, starting we must discuss the following.

6.5.1 Airy's function

Consider a particle experiencing a constant force F then, the potential is

$$U(x) = -Fx.$$

The corresponding Schrödinger equation is

$$\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\Psi}{\mathrm{d}x^2} + (E + Fx)\Psi = 0.$$

Making the change of variables

$$x = y + \frac{E}{F} \Rightarrow -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\Psi}{\mathrm{d}x^2} + yF\Psi = 0$$

and next we rescale with $y = z\left(\frac{\hbar^2}{2mF}\right)^{1/3}$ to obtain

$$\boxed{\frac{\mathrm{d}^2\Psi}{\mathrm{d}z^2} - z\Psi = 0.}$$

The solution to his equation which decays at large z is

$$\Psi = c\text{Ai}(z)$$

for some constant c.

Definition 6.4. **Airy's function** is defined as

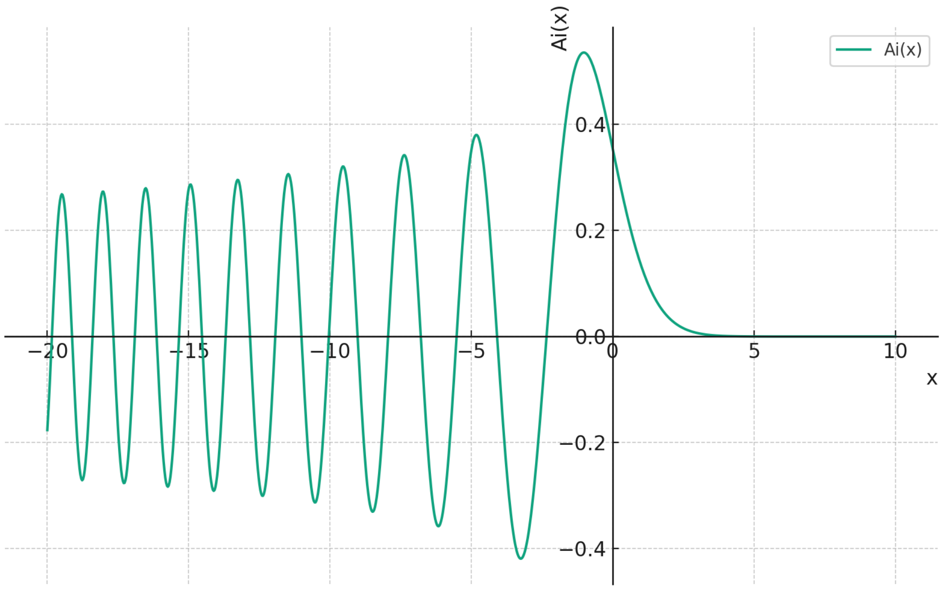
$$\text{Ai}(z) = \frac{1}{\sqrt{\pi}} \int_0^\infty \cos\left(\frac{1}{3}u^3 + zu\right) \, du$$

Proposition 6.5

Some results regarding Airy's function.

- $\text{Ai}'(z) = -\frac{1}{\sqrt{\pi}} \int_0^\infty u \sin\left(\frac{1}{3}u^3 + zu\right) \, du;$
- $\text{Ai}''(z) = -\frac{1}{\sqrt{\pi}} \int_0^\infty u^2 \cos\left(\frac{1}{3}u^3 + zu\right) du;$
- $\text{Ai}''(z) - z\text{Ai}(z) = 0.$

Below is a plot of Airy's function.



It oscillates to the left of the turning point (where $E > Fx$) and then decays exponentially to the right (where $E < Fx$). This is what we expect near a turning point.

> **Remark 6.6.** To the right of the turning point (i.e. when $U = Fx$ and $U > E_{\text{total}}$) Airy's function and in turn the wave function are not zero. Thus, there is a non-zero probability

> to find the particle in a region which is strictly forbidden in the classical world.

6.5.2 Bohr-Sommerfeld quantisation

Let us do the WKB procedure for a potential of the form

$$U = E + F(x-b).$$

This is a potential that arises to the right ($F > 0$) with the turning point at $x = b$. We have that

$$p(x) = \sqrt{2m(Fb - Fx)}$$

and hence (we pick the lower bound of integration constant to make things simple as the effect is just a constant that can be absorbed elsewhere)

$$\begin{aligned}\int_b^x p(y)\,dy = \sigma^{(0)} &= \int_b^x \sqrt{2m(Fb-Fy)}\,dy \\ &= -\frac{2}{3}\sqrt{2mF}(b-x)^{3/2}\end{aligned}$$

and we can have that the WKB wave function is

$$\psi_{\text{WKB}} = \frac{A}{\sqrt{p(x)}}e^{\frac{i}{\hbar}\frac{2}{3}\sqrt{2mF}(b-x)^{3/2}} + \frac{B}{\sqrt{p(x)}}e^{-\frac{i}{\hbar}\frac{2}{3}\sqrt{2mF}(b-x)^{3/2}}.$$

However, we do not trust this approximation near $x = b$ and in particular when $p = 0$. Rather we choose the following:

$$\psi = \begin{cases} \phi_{\text{WKB}} & \text{if } x > b \\ c\text{Ai}\left(\frac{2mF}{\hbar}\right)^{1/3}(x-b) & \text{if } x \sim b \\ \chi_{\text{WKB}} & \text{if } x < b \end{cases}$$

Where ϕ_{WKB} and χ_{WKB} are the same as ψ_{WKB} but with different normalisation constants.

Let us consider a more general situation with a potential with turning points on the left and right. In the middle region $a < x < b$ we need to match the two solution we find (assuming a general potential):

$$\begin{aligned}\psi_b &= \frac{c}{\sqrt{p(x)}}\sin\left(-\frac{1}{\hbar}\int_b^x p(y)\,dy + \frac{\pi}{4}\right) \quad \text{if } x < b \\ \psi_a &= \frac{c}{\sqrt{p(x)}}\sin\left(-\frac{1}{\hbar}\int_a^x p(y)\,dy - \frac{\pi}{4}\right) \quad \text{if } x > a\end{aligned}$$

which are equal when

$$-\frac{1}{\hbar}\int_b^x p(y)\,dy + \frac{\pi}{4} = -\frac{1}{\hbar}\int_a^x p(y)\,dy - \frac{\pi}{4} + n\pi$$

for some $n \in \mathbb{Z}$. Rearranging this gives the

Proposition 6.7

Bohr-Sommerfeld quantisation rule:

$$\int_a^b p(y)\,dy = \pi\hbar\left(n - \frac{1}{2}\right) \quad \text{for } n \in \mathbb{Z}$$

where a and b are two turning points.

6.6 Quantisation III: Harmonic oscillator

In this example the potential is $U(x) = \frac{k}{2}x^2$ so that

$$p(x) = \sqrt{2m\left(E - \frac{k}{2}x^2\right)}$$

we have to find the turning points a and b thus, we need to find when $E = V$. We have

$$E = \frac{k}{2}x^2 \iff a, b = \pm\sqrt{\frac{2E}{k}}$$

Using the Bohr-Sommerfeld quantisation rule we have

$$\int_{-\sqrt{\frac{2E}{k}}}^{\sqrt{\frac{2E}{k}}} \sqrt{2m\left(E - \frac{k}{2}x^2\right)}\, dx = \pi\hbar\left(n - \frac{1}{2}\right)$$

which gives

$$\pi\sqrt{\frac{E}{\frac{m}{k}}} = \pi\hbar\left(n - \frac{1}{2}\right)$$

equivalently

$$E = \hbar\sqrt{\frac{k}{m}}\left(n - \frac{1}{2}\right) = \hbar\omega\left(n - \frac{1}{2}\right)$$

which in fact gives the exact spectrum in this case! (In general it will be only valid for large n's).

7 Density matrices, thermal states and entanglement

Proposition 7.1

The trace of a matrix A in bra-ket notation is expressed as:

$$\text{Tr}(A) = \sum_{i=1}^{n} \langle i|A|i\rangle$$

Here, $|i\rangle$ represents the basis vectors, and $\langle i|A|i\rangle$ denotes the diagonal elements of the matrix.

Proposition 7.2

$\text{Tr}(|\psi\rangle\langle\psi|) = \dim(\mathcal{H}) \times \langle\psi|\psi\rangle$.

7.1 Density matrices

Definition 7.3. Define the operator $\rho = |\psi\rangle\langle\psi|$. This operator is a matrix which we call the **density matrix**.

Remark 7.4. This is indeed a matrix (considering a 3-dimensional Hilbert space) we have:

$$\begin{aligned} |\Phi\rangle\langle\Psi| &= \begin{pmatrix}\Phi_1\\ \Phi_2\\ \Phi_3\end{pmatrix}\begin{pmatrix}\Psi_1^* & \Psi_2^* & \Psi_3^*\end{pmatrix} \\ &= \begin{pmatrix}\Phi_1\Psi_1^* & \Phi_1\Psi_2^* & \Phi_1\Psi_3^*\\ \Phi_2\Psi_1^* & \Phi_2\Psi_2^* & \Phi_2\Psi_3^*\\ \Phi_3\Psi_1^* & \Phi_3\Psi_2^* & \Phi_3\Psi_3^*\end{pmatrix}\end{aligned}$$

Proposition 7.5. We have that $\mathrm{tr}(\rho) = 1$.

Proof. The simplest way to see that is to assume that $|\psi\rangle$ is an element of some basis $\{|\psi_i\rangle\}$, and calculate the trace as

$$\mathrm{tr}(\rho) = \sum_i \langle\psi_i|\psi\rangle\langle\psi|\psi_i\rangle$$

This sum gets a contribution only from the i for which $|\psi\rangle = |\psi_i\rangle$, and gives indeed 1. □

Proposition 7.6

Given any observable $\mathcal{O}$ we have that the expectation value of $\mathcal{O}$ is

$$\langle\mathcal{O}\rangle = \mathrm{tr}(\mathcal{O}\rho).$$

Proof. Assume that we have an orthonormal basis which is given by $\{|e_i\rangle\}_i$, the trace is given by

$$\begin{aligned}\mathrm{tr}(\mathcal{O}\rho) &= \sum_n \langle e_n|\mathcal{O}\rho|e_n\rangle \\ &= \sum_n \langle e_n|\mathcal{O}|\psi\rangle\langle\psi|e_n\rangle \\ &= \sum_n \langle\psi|e_n\rangle\langle e_n|\mathcal{O}|\psi\rangle \\ &= \langle\psi|\mathcal{O}|\psi\rangle .\end{aligned}$$

This is because the $\sum_n |e_n\rangle\langle e_n| = 1$ as discussed above. □

7.1.1 Pure and mixed states

Definition 7.7. A **pure state** of a quantum system is denoted by a vector (ket) with unit length, so the density matrix of a pure state $|\psi\rangle$ is given by $\rho = |\psi\rangle\langle\psi|$.

Definition 7.8. A state where the density matrix is given by $\rho = \sum_n p_n |\psi_n\rangle\langle\psi_n|$ is called a **mixed state** where $0 \le p_n \le 1$.

Note 7.9. A few things.

- We need $\sum p_i = 1$.
- This is a sum of matrices.

- If given a basis of eigenvectors then p_i are the eigenvalues of each state.

Theorem 7.10. A pure state, $|\psi\rangle$, has density matrix $|\psi\rangle\langle\psi|$ hence,

$$\operatorname{tr}(\rho^2)=\operatorname{tr}(\rho)=1.$$

Proof. For a pure state $|\psi\rangle$ we have that $\rho=|\psi\rangle\langle\psi|$ thus,

$$\begin{aligned}\rho^2&=(|\psi\rangle\langle\psi|)(|\psi\rangle\langle\psi|)\\&=|\psi\rangle\langle\psi|\psi\rangle\langle\psi|\\&=|\psi\rangle\langle\psi|=\rho.\end{aligned}$$

☐

Proposition 7.11
For a mixed state $|\phi\rangle$, with orthonormal basis $\{|\phi_i\rangle\}_i$ the density matrix is given by

$$\rho=\sum_n p_n|\phi_n\rangle\langle\phi_n|\,.$$

Theorem 7.12. For a mixed stated we have that $\operatorname{tr}(\rho^2)\leq 1$.

Note 7.13. The above is a methodology to distinguish between pure and mixed states.

Proof. The density matrix of a mixed state is give by $\rho=\sum_n p_n|\phi_n\rangle\langle\phi_n|$. Therefore,

$$\begin{aligned}\rho^2&=\sum_n\sum_m(p_n|\phi_n\rangle\langle\phi_n|)(p_m|\phi_m\rangle\langle\phi_m|)\\&=\sum_n p_n^2|\phi_n\rangle\langle\phi_n|\,.\end{aligned}$$

We have that $0\leq p_n\leq 1$ and so, $p_n^2\leq p_n$ for each p_n. Hence,

$$\operatorname{tr}(\rho^2)=\sum_n\rho_n^2\leq\sum_n p_n=1.$$

☐

7.2 Thermal states

The following is the most classic example of a mixed state.

Definition 7.14. The **Boltzmann distribution** is a mixed state described by density matrix:

$$\rho_{\text{thermal}}=\frac{1}{Z}\sum_n e^{-\frac{E_n}{k_BT}}|E_n\rangle\langle E_n|\quad\text{and}\quad Z=\sum_n e^{-\frac{E_n}{k_BT}}$$

where k_B is the Boltzmann constant (converts temperature into energy), T is temperature and $|E_n\rangle$ are the energy eigenstates.

Note 7.15. To find the normalisation Z we need to impose that $\mathrm{tr}(\rho) = 1$.

Proof. We use the basis for the density matrix given by the eigenstates $|E_m\rangle$. We compute the trace:

$$\begin{aligned}1 = \mathrm{tr}(\rho) &= \frac{1}{Z}\sum_{m,n} e^{-\frac{E_n}{k_B T}} \langle E_m|\, \rho \,|E_m\rangle \\ &= \frac{1}{Z}\sum_{m,n} e^{-\frac{E_n}{k_B T}} \langle E_m|E_n\rangle \langle E_n|E_m\rangle \\ &= \frac{1}{Z}\sum_{m,n} e^{-\frac{E_n}{k_B T}} \delta_{mn}\delta_{nm} \\ &= \frac{1}{Z}\sum_{n} e^{-\frac{E_n}{k_B T}}.\end{aligned}$$

Rearranging, gives us the required Z. □

Definition 7.16. The **partition function** counts the number of states available at each energy E_n:

$$Z = \sum_{E_n} d(n) e^{-\frac{E_n}{k_B T}},$$

where $d(n)$ counts the degeneracy of states at energy level E_n.

Note 7.17. By degeneracy we mean the mutliplicity of the eigenvalue E_n.

Proposition 7.18. At low temperatures, $T \to 0$, the density matrix will peak around the lowest energy state:

$$\lim_{T\to 0} \rho_{\text{thermal}} \,|E_0\rangle \langle E_0|$$

hence, becomes a pure state.

Proposition 7.19. At high temperatures, $T \to \infty$, all the energy states contribute (more or less) equally and so

$$\lim_{T\to\infty} \rho_{\text{thermal}} = \frac{1}{\dim(\mathcal{H})} \sum_n |E_n\rangle \langle E_n|\,.$$

Theorem 7.20

The expected energy of a thermal state is

$$\langle H\rangle = \mathrm{tr}(H\rho) = -\frac{\partial}{\partial\beta}\ln(Z),$$

where $\rho = \rho_{\text{thermal}}$ and $\beta = \frac{1}{k_B T}$.

Proof. We compute the trace

$$\begin{aligned}
\mathrm{tr}(H\rho) &= \sum_m \langle E_m| H\rho |E_m\rangle \\
&= \frac{1}{Z}\sum_m\sum_n e^{-\beta E_n} \langle E_m| H |E_n\rangle \langle E_n|E_m\rangle \\
&= \frac{1}{Z}\sum_m\sum_n e^{-\beta E_n} E_n \langle E_m|E_n\rangle \langle E_n|E_m\rangle \\
&= \frac{1}{Z}\sum_m\sum_n e^{-\beta E_n} E_n \delta_{mn}\delta_{mn} \\
&= \frac{1}{Z}\sum_n E_n e^{-\beta E_n} \\
&= -\frac{1}{Z}\frac{\partial}{\partial\beta}\sum_n e^{-\beta E_n} \\
&= -\frac{\partial}{\partial\beta}\ln(Z).
\end{aligned}$$

Where we have applied the fact that $H|E_n\rangle = E_n|E_n\rangle$. □

7.3 (Von Neumann) Entropy

We found a way to characterize the non-purity of a state in terms of the trace of the square of the matrix. In this section we define a new way to do so.

Definition 7.21. **Entropy** is a measure of disorder/chaos.

Definition 7.22. The most basic definition of entropy is

$$S(E_n) = \ln(d(n))$$

i.e. the log of the number of states with the same energy.

Definition 7.23. The **Von Neumann entropy** is defined as

$$S = -\mathrm{tr}(\rho\ln\rho) = -\sum_n p_n \ln p_n,$$

where p_n are the eigenvalues of the the density matrix ρ.

> **Remark 7.24.** We also call it S_{vN} for Von Neumann entropy.

> **Note 7.25.** To compute the log of a matrix we diagonalise it then, take the log of each diagonal entry.

Proposition 7.26

We have that

- for a pure state $S = 0$;
- for mixed states $S > 0$.

Proof. For a pure state $S = 0$, since either $p_i = 1$, with vanishing ln or $p_n = 0$ (notice that we use $\lim_{x\to 0} x \ln x = 0$). Since $0 \leq p_n \leq 1$, the entropy is always positive. ☐

Theorem 7.27. We have that $S \leq \ln(\dim(\mathcal{H}))$.

Definition 7.28. A mixed state with $S = \ln(\dim(\mathcal{H}))$ is said to be **maximally mixed**.

Example 7.29. This happens when $p_n = \frac{1}{\dim(\mathcal{H})}$ for all n.

7.4 Entanglement

Consider a system comprised of two non-interacting particles, or more generally a tensor product Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, as discussed in the context of products of $\mathfrak{su}(2)$ representations.
The most general state in our tensor Hilbert space can be written as (in the standard basis)

$$|\Psi\rangle = \sum_{n_1,n_2} c_{n_1 n_2} |e_{n_1}\rangle \otimes |e_{n_2}\rangle$$

with $|e_{n_1}\rangle$ and $|e_{n_2}\rangle$ are the standard basis vectors of $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. Note that the pure state over $\mathcal{H}$ takes the form

$$\begin{aligned}\rho &= |\Psi\rangle\langle\Psi| \\ &= \sum_{n_1,n_2,m_1,m_2} c_{n_1 n_2} c^*_{m_1 m_2} |e_{n_1}\rangle \otimes |e_{n_2}\rangle \langle e_{m_1}| \otimes \langle e_{m_2}| \\ &= \sum_{n_1,n_2,m_1,m_2} c_{n_1 n_2} c^*_{m_1 m_2} |e_{n_1}\rangle\langle e_{m_1}| \otimes |e_{n_2}\rangle\langle e_{m_2}| .\end{aligned}$$

Let us assume that we are asking questions only about the first particle and have no knowledge of the second particle. We should then average over $\mathcal{H}_2$. Doing so amounts to taking the trace

$$\begin{aligned}\rho_{\text{reduced}} &= \mathrm{tr}_{\mathcal{H}_2}(\rho) \\ &= \sum_p \langle e_p| \left(\sum_{n_1,n_2,m_1,m_2} c_{n_1 n_2} c^*_{m_1 m_2} |e_{n_1}\rangle\langle e_{m_1}| \otimes |e_{n_2}\rangle\langle e_{m_2}| \right) |e_p\rangle \\ &= \sum_{n_1,m_1} \left(\sum_{n_2,m_2} c_{n_1 n_2} c^*_{m_1 m_2} \right) |e_{n_1}\rangle\langle e_{m_1}| \\ &= \sum_{n_1,m_1} a_{n_1 m_1} |e_{n_1}\rangle\langle e_{m_1}| .\end{aligned}$$

Note 7.30. Averaging over $\mathcal{H}_2$ sets the trace of $|e_{n_2}\rangle\langle e_{m_2}|$ to 1.

This is a mixed state for the first Hilbert space known as **reduced density matrix**. Note that $a_{n_1 m_1}$ defines a self-adjoint operator on $\mathcal{H}_1$. Thus, we can find a new basis of $\mathcal{H}_1$ given by $|e'_{n_1}\rangle$ such that

$$\rho_{\text{reduced}} = \sum_{n_1} p_{n_1} |e'_{n_1}\rangle\langle e'_{n_1}| .$$

Furthermore, the p_{n_1} which are the eigenvalue of $a_{n_1 m_1}$ are real and in fact positive and less than one.

Theorem 7.31
The **entanglement entropy** given by

$$S = -\mathrm{tr}_{\mathcal{H}_1}\left(\rho_{\text{reduced}}\ln(\rho_{\text{reduced}})\right)$$

is the same in both $\mathcal{H}_1$ and $\mathcal{H}_2$.

8 Relativistic QM

Definition 8.1. The following transformation is called a **Galilean boost**:

$$\boldsymbol{x}_i \to \boldsymbol{x}_i + \boldsymbol{v}t$$

Note 8.2. In Galilean relativity the speed of light is not constant under a boost since velocities are shifted i.e. $\dot{\boldsymbol{x}} = \dot{\boldsymbol{x}} + \boldsymbol{v}$.

Definition 8.3. A **Lorentz boost** are transformation of the form

$$\begin{aligned} ct' &= \gamma ct + \gamma\boldsymbol{\beta}\cdot\boldsymbol{x} \\ \boldsymbol{x}' &= \gamma\boldsymbol{x} + \gamma\boldsymbol{\beta}ct \\ c\cdot \mathrm{d}t' &= \gamma \mathrm{d}ct + \gamma\boldsymbol{\beta}\cdot \mathrm{d}\boldsymbol{x} \\ \mathrm{d}\boldsymbol{x}' &= \gamma \mathrm{d}\boldsymbol{x} + \gamma\boldsymbol{\beta}c\cdot \mathrm{d}t \end{aligned}$$

where c is the speed of light.

Note 8.4. In special relativity we have that the invariant notion of length is, for infinitesimal displacement

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - \mathrm{d}x^2 = \mathrm{d}y^2 - \mathrm{d}z^2.$$

8.1 Relativistic wave equation

In special relativity time and space are interchangeable. The first attempt to find a relativistic wave equations is to simply make time second order and consider

$$\frac{\hbar^2}{c^2}\frac{\partial^2\Psi}{\partial t^2} - \hbar^2\nabla^2\Psi + m^2c^2\Psi = 0$$

This is invariant under Special relativity if $\Psi(x) \to \Psi'(x) = \Psi(x')$. It is known as the **Klein-Gordon equation**. But there are big problems:

1. If we look at energy and momentum eigenstates $\Psi = e^{-iEt/\hbar + i\boldsymbol{p}\cdot\boldsymbol{x}/\hbar}$ we find

   $$E^2 = |\boldsymbol{p}|^2c^2 + m^2c^4$$

   This allows for both positive and negative energy.

2. We would like to interpret $\Psi^*\Psi$ as a probability density but it is not conserved:
$$\frac{d}{dt}\int \Psi^*\Psi d^3x \neq 0$$
One can find a conserved density: $\rho = i\hbar(\Psi^*\frac{\partial\Psi}{\partial t} - \Psi\frac{\partial\Psi^*}{\partial t})$ but this is not positive definite.

3. Knowing the state at time $t = 0$ is not enough information to determine it at later times, one also needs $\frac{\partial\Psi}{\partial t}$ at $t = 0$.

An alternative approach is by Dirac. He proposed an equation that was first order in both time and space. First we define

Definition 8.5. The **anti-commutator** $\{\gamma^i, \gamma^j\} = \gamma^i\gamma^j + \gamma^j\gamma^i$.

Definition 8.6. The **Dirac equation** is given by:
$$\left(i\hbar\gamma^0\frac{\partial}{c\partial t} - i\hbar\gamma\cdot\nabla + mc\right)\Psi = 0.$$
Where
$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}$$
where $\mathbb{I}$ is the 2×2 unit matrix and σ^i, the Pauli matrices.

Proposition 8.7

We recover the Klein-Gordon equation if we square the Dirac equation and set
$$(\gamma^0)^2 = 1, \quad \{\gamma^0, \gamma^i\} = 0, \quad \{\gamma^i, \gamma^j\} = -2\delta^{ij}$$

Corollary 8.8. The wave function Ψ is a complex vector of the form
$$\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}.$$

8.2 Relativistic notation

Definition 8.9. We will use the following notation, x^μ with $\mu \in \{0, 1, 2, 3\}$ where
$$x^0 = ct, \quad x^1 = x, \quad x^2 = y, \quad x^3 = z$$
and c is the speed of light.

Theorem 8.10. With the notation above the invariant length is written as
$$ds^2 = \eta_{\mu\nu}dx^\mu dx^\nu$$
where
$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$
and repeated indices are summed over.

Corollary 8.11. The inverse to $\eta_{\mu\nu}$ is given by

$$\eta^{\mu\nu} = (\eta_{\mu\nu})^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Note 8.12. Numerically, they are the same but this is a coincidence.

Corollary 8.13

We can use η to raise and lower indices i.e.

$$\eta_{\mu\nu} s^{\mu} = s_{\nu} \quad \text{and} \quad \eta^{\mu\nu} s_{\nu} = s^{\mu}.$$

Proposition 8.14

Using this notation the Dirac equation is written as

$$-i\hbar\gamma^{\mu}\frac{\partial}{\partial x^{\mu}}\Psi + mc\Psi = 0.$$

Note 8.15. Later on we define the conjugate, we the dirac equation is the is

$$i\hbar\frac{\partial}{\partial x^{\mu}}\overline{\Psi}\gamma^{\mu} + mc\overline{\Psi} = 0.$$

8.3 Spinors

In this section we will talk about the notation we will use for Ψ.

Definition 8.16. We call Ψ a **spinor**, and we denote its entries by Ψ_{α} for $\alpha \in \{1, 2, 3, 4\}$.

Definition 8.17. The γ-matrices are now denoted by

$$\gamma^{\mu} = (\gamma^{\mu})_{\alpha}{}^{\beta}$$

where α and β are called the spinor indices.

Theorem 8.18. The condition on the anti-commutator of the γ matrices is now written as

$$\{\gamma^{\mu}, \gamma^{\nu}\} = 2\eta^{\mu\nu}\mathbb{I} \iff \{\gamma^{\mu}, \gamma^{\nu}\}_{\alpha}{}^{\beta} = 2\eta^{\mu\nu}\delta_{\alpha}{}^{\beta}$$

Theorem 8.19. We have that $(\gamma^{\mu})^{\dagger} = \gamma^{0}\gamma^{\mu}\gamma^{0}$.

Theorem 8.20
We can write a Lorentz transformation as

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu}$$

which implies that $ds'^2 = ds^2$ can be written as

$$\Lambda^{\mu}{}_{\rho} \Lambda^{\nu}{}_{\sigma} \eta_{\mu\nu} = \eta_{\rho\sigma}$$

i.e. a Lorentz transformation is invariant if it satisfies the above (Einstein summation is used).

Proof. We want a transformation that can keep the inner product invariant. Then if $X' = \Lambda X$ and $Y' = \Lambda Y$ then

$$X' \cdot Y' = \eta_{\mu\nu} X'^{\mu} Y'^{\nu}$$

$$= \eta_{\mu\nu} (\Lambda^{\mu}{}_{\rho} X^{\rho})(\Lambda^{\nu}{}_{\sigma} Y^{\sigma})$$

$$= \eta_{\lambda\xi} (\Lambda^{\lambda}{}_{\mu} X^{\mu})(\Lambda^{\xi}{}_{\nu} Y^{\nu})$$

where we have relabelled the dummy indices in the last line in order to make a direct comparison with

$$X \cdot Y = \eta_{\mu\nu} X^{\mu} Y^{\nu}.$$

If Λ leaves the Minkowski inner product invariant then it must satisfy

$$\eta_{\lambda\xi} \Lambda^{\lambda}{}_{\mu} \Lambda^{\xi}{}_{\nu} = \eta_{\mu\nu}$$

or in matrix notation,

$$\Lambda^{T} \eta \Lambda = \eta$$

□

8.4 Back to QM

From all the above we now have a wave equation that is relativistic and first order in time.

Proposition 8.21. We can write the Dirac equation as follows:

$$i\hbar \frac{\partial}{\partial t} \Psi = \widehat{H} \Psi \quad \text{where} \quad \widehat{H} = -i\hbar c \gamma^0 \boldsymbol{\gamma} \cdot \nabla + mc^2 \gamma^0.$$

Therefore, $\widehat{H}$ is self-adjoint but not positive definite and γ^0 has eigenvalues ± 1.

Definition 8.22. The **Dirac conjugate** $\overline{\Psi} = \Psi^{\dagger} \gamma^0$.

Proposition 8.23

An infinitesimal Lorentz transformation of Ψ is given by

$$\delta\Psi = \frac{1}{4}\omega_{\mu\nu}\gamma^{\mu\nu}\Psi,$$

where $w_{\mu\nu} = -w_{\nu\mu}$.

Note 8.24. To find δ of something, take its derivative i.e. for J^{μ}.

Note 8.25. The ω is obtained by consider $\Lambda^{\mu}_{\ \nu} = \delta^{\mu}_{\nu} + \omega^{\mu}_{\ \nu}$ (i.e. expanding inoto the infinitesimal form). Plug this in to the Lorentz transformation and after a contraction of indices we have the result.

Theorem 8.26

We have that

$$\partial_{\mu}\Psi = \frac{\partial x'^{\nu}}{\partial x^{\mu}}\frac{\partial\Psi}{\partial x'^{\nu}} = \Lambda^{\nu}_{\ \mu}\partial'_{\nu}\Psi.$$

9 The Feynman path integral

This section presents another way of formulating quantum mechanics which is known as the **path integral formulation**.

Definition 9.1. We define the **propagator**

$$\begin{aligned} K(a_2, t_2; a_1, t_1) &= \langle a_2, t_2 | a_1, t_1 \rangle \\ &= \langle a_2 | \, e^{i\widehat{H}(t_2 - t_1)/\hbar} \, | a_1 \rangle \, . \end{aligned}$$

It takes a particle at position a_1 at time t_1 and computes the amplitude that it is found at a_2 at time t_2.

Theorem 9.2

For Hamiltonians of the form $\widehat{H} = \frac{1}{2m}\widehat{p}^2 + V(\widehat{q})$ the propagator takes the form of a path-integral

$$K(a_2, t_2; a_1, t_1) = \int \exp\left(\frac{i}{\hbar}\int \left(k(t)\dot{a}(t) + H(k(t), a(t))\right)\, dt\right) \left[\frac{dk}{2\pi\hbar}\right][da],$$

where

- $k(t)$ and $a(t)$ are arbitrary paths such that $a(t_1) = a_1$ and $a(t_2) = a_2$;
- $\dot{a}(t)$ is the derivative with respect to time of $a(t)$;
- $H(k(t), a(t)) = \frac{1}{2m}k(t)^2 + V(a(t))$.

The square brackets indicate that this is an infinite dimensional integral where we are integrating over all paths $k(t)$ and $a(t)$.

9.1 Gaussian path integral

Definition 9.3. Gaussian integrals are of the form

$$\int e^{-Ak^2+Bk}\, dk = \sqrt{\frac{\pi}{A}} e^{\frac{B^2}{4A}}.$$

Theorem 9.4. The integral of a Gaussian gives another Gaussian.

Theorem 9.5. From the propagator the integral over $k(t)$ is a Gaussian integral.

Proposition 9.6. We can write the propagator as

$$K(a_2, t_2; a_1, t_1) = \int e^{-\frac{i}{\hbar}S[a(t)]}\, [da],$$

where

$$S[a] = \int_{t_1}^{t_2} \frac{1}{2}m\dot{a}^2 - V(a)\, dt.$$

9.2 Computing determinants

Let us look at a simple example $H = \frac{1}{2m}\widehat{p}^2 + \frac{1}{2}k\widehat{q}^2$. The corresponding Lagrangian is of course just

$$L = \frac{1}{2}m\dot{q}^2 - \frac{1}{2}kq^2$$

It is helpful to rewrite the action as

$$S = -\frac{1}{2}\int q\mathcal{E}q\, dt \quad \mathcal{E} = m\frac{\mathrm{d}^2}{\mathrm{d}t^2} + k$$

where we view $\mathcal{E}$ as a differential operator. Let us expand $q(t)$ in a real orthonormal basis of functions

$$q = \sum_n q_n e_n(t)$$

with

$$\int e_n(t)e_m(t)\,dt = \delta_{nm}$$

Thus

$$S = -\sum_n q_n q_m \int \frac{1}{2} e_n(t)\mathcal{E}e_m(t)\,dt$$

Let us choose $e_n(t)$ to be eigenstates of $\mathcal{E}$ with eigenvalues λ_n:

$$\begin{aligned}
S &= -\sum_n q_n q_m \int \frac{1}{2} e_n(t)\lambda_n e_m(t)\,dt \\
&= -\frac{1}{2}\sum_n \lambda_n q_n^2
\end{aligned}$$

Thus

$$\begin{aligned}
\int e^{-\frac{i}{\hbar}S[q]}\,[dq] &= \int \prod_n e^{\frac{i}{2\hbar}\sum_n \lambda_n q_n^2}\,dq_n \\
&= \prod_n \int e^{\frac{i}{2\hbar}\lambda_n q_n^2}\,dq_n \\
&= \prod_n \sqrt{\frac{\pi\hbar}{\lambda_n}} \\
&= \frac{N}{\sqrt{\det \mathcal{E}}} \quad \text{where} \quad \det \mathcal{E} = \prod_n \lambda_n
\end{aligned}$$

Here $\mathcal{N}$ is another normalization constant that we are not interested in. We can generalise this: for a quadratic action with many coordinates q_a for $a = 1, \ldots, n$ of the form

$$S = -\frac{1}{2}\int q_a \mathcal{E}_{ab} q_b\,dt$$

we find

$$\int \prod_a e^{\frac{i}{\hbar}S[q_1,\ldots,q_n]}\,[dq_a] = \frac{\mathcal{N}}{\sqrt{\det \mathcal{E}_{ab}}}$$

where the determinant is the infinite product of eigenvalues of the operator $\mathcal{E}_{ab}$. More generally still if one has

$$S = -\frac{1}{2}\int q_a \mathcal{E}_{ab} q_b + J_a q_a\,dt$$

we find

$$\int \prod_a e^{\frac{i}{\hbar}S[q_1,\ldots,q_n]}\,[dq_a] = \frac{\mathcal{N}}{\sqrt{\det \mathcal{E}_{ab}}} e^{-\frac{1}{2}J_a \mathcal{E}_{ab}^{-1} J_b}.$$

Appendix

A Tensors

A.1 Kronecker δ function

Definition A.1. The Kronecker delta function is defined as

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

Note A.2. The function, δ_{ij}, can also be thought as the identity tensor.

A.2 ε-tensor

Definition A.3. In three dimensions, the ε-tensor (or the Levi-Civita symbol) is defined by

$$\varepsilon_{ijk} = \begin{cases} 1 & \text{if } (i,j,k) = (1,2,3),(2,3,1),(3,1,2) \\ -1 & \text{if } (i,j,k) = (3,2,1),(1,3,2),(2,1,3) \\ 0 & \text{otherwise.} \end{cases}$$

Note A.4. That is $\varepsilon_{ijk} = 1$ if (i,j,k) are a clockwise (or even) permutation of $(1,2,3)$ whereas, $\varepsilon_{ijk} = -1$ if (i,j,k) are an anti-clockwise (or odd) permutation of $(1,2,3)$ and 0 if any of the indices are the same.

Theorem A.5

The ε-tensor obeys the following relation:

$$\sum_{i=1}^{3} \varepsilon_{ijk}\varepsilon_{inm} = \delta_{jn}\delta_{km} - \delta_{jm}\delta_{kn},$$

where δ_{ab} is the Kronecker-δ function.

B Notation

Definition B.1. Let $\boldsymbol{v}$ have components v^i, as $\boldsymbol{v}$ is a column vector then the index i labels the rows of the single column in the vector.

Definition B.2. We have $\boldsymbol{v}^\top$ is a row vector, the components are labelled by v_i where i labels the column of the single row vector.

Note B.3. In conclusion, upper indices representes row and lower indices columns.

Corollary B.4. With the notation above we may write $\boldsymbol{v} \cdot \boldsymbol{v} = \boldsymbol{v}^\top \boldsymbol{v} = \sum_{i=1}^{3} v_i v^i$.

Definition B.5. The notation ${A^i}_j$ indicates the component of the matrix A in the i-th row and j-th column.

Corollary B.6

With this notation, we can write $A\boldsymbol{v} = \boldsymbol{u}$ as

$$(A\boldsymbol{v})^i = \sum_{j=1}^{3} {A^i}_j v^j = u^i.$$

Definition B.7. Taking the transpose of a matrix switches the indices i.e.

$$({A^\top)^i}_j = {A_j}^i.$$

Definition B.8. The **Euclidean metric tensor**, δ_{ij}, is used to lower indices. That is,

$$\sum_j \delta_{ij} v^j = v_i.$$

Its inverse is given by $\delta^{-1} = \delta^{ij}$ which is used to raise the indices. That is,

$$\delta^{ij} v_j = v^i.$$

Note B.9. This is the Kronecker-δ function.

C Vector identities

Proposition C.1

Dot product.

1. **Commutative Property:** $\boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{b} \cdot \boldsymbol{a}$
2. **Distributive Property over Addition:** $\boldsymbol{a} \cdot (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{a} \cdot \boldsymbol{c}$
3. **Scalar Multiplication:** $(k\boldsymbol{a}) \cdot \boldsymbol{b} = k(\boldsymbol{a} \cdot \boldsymbol{b}) = \boldsymbol{a} \cdot (k\boldsymbol{b})$
4. **Dot Product with Zero Vector:** $\boldsymbol{a} \cdot \boldsymbol{0} = 0$
5. **Dot Product with Itself:** $\boldsymbol{a} \cdot \boldsymbol{a} = \|\boldsymbol{a}\|^2$
6. **Orthogonality:** If $\boldsymbol{a} \perp \boldsymbol{b}$ (i.e., $\boldsymbol{a}$ and $\boldsymbol{b}$ are orthogonal), then $\boldsymbol{a} \cdot \boldsymbol{b} = 0$.

Proposition C.2

Cross product.

1. **Anticommutativity:** $\boldsymbol{a} \times \boldsymbol{b} = -(\boldsymbol{b} \times \boldsymbol{a})$
2. **Distributive Property over Addition:** $\boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{a} \times \boldsymbol{c}$
3. **Scalar Triple Product:** $\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b} \cdot (\boldsymbol{c} \times \boldsymbol{a}) = \boldsymbol{c} \cdot (\boldsymbol{a} \times \boldsymbol{b})$
4. **Cross Product with Zero Vector:** $\boldsymbol{a} \times \mathbf{0} = \mathbf{0} \times \boldsymbol{a} = \mathbf{0}$
5. **Cross Product with Itself:** $\boldsymbol{a} \times \boldsymbol{a} = \mathbf{0}$
6. **Cross Product with Parallel Vectors:** If $\boldsymbol{a}$ and $\boldsymbol{b}$ are parallel, then $\boldsymbol{a} \times \boldsymbol{b} = \mathbf{0}$.
7. **BAC-CAB Identity (Triple Cross Product Identity):**
$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c}$$
8. **Scalar Quadruple Product:**
$$\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) \times \boldsymbol{d} = (\boldsymbol{a} \cdot \boldsymbol{d})(\boldsymbol{b} \cdot \boldsymbol{c}) - (\boldsymbol{a} \cdot \boldsymbol{c})(\boldsymbol{b} \cdot \boldsymbol{d})$$

Theorem C.3

The i^{th} component of the resultant vector of the cross product is given by

$$(\boldsymbol{a} \times \boldsymbol{b})_i = \sum_{j=1}^{3} \sum_{k=1}^{3} \varepsilon_{ijk} a_j b_k.$$

Note C.4. By the i^{th} component we mean

$$(\boldsymbol{a} \times \boldsymbol{b}) = \begin{pmatrix} (\boldsymbol{a} \times \boldsymbol{b})_1 \\ (\boldsymbol{a} \times \boldsymbol{b})_2 \\ (\boldsymbol{a} \times \boldsymbol{b})_3 \end{pmatrix}.$$

D The Pauli Matrices

The Pauli matrices are denoted by σ_1, σ_2, and σ_3, and they form a basis for the vector space of 2×2 Hermitian matrices. Each matrix is traceless and has a determinant of -1. They can be represented explicitly as:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

D.1 Properties

The Pauli matrices have several interesting properties. They are Hermitian and unitary, and they obey the following commutation and anti-commutation relations:

$$[\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k, \quad \{\sigma_i, \sigma_j\} = 2\delta_{ij}\mathbb{I}_{2\times 2},$$

where $i, j, k \in \{1, 2, 3\}$, εijk is the Levi-Civita symbol, δ_{ij} is the Kronecker delta, and I is the 2×2 identity matrix.

E Tensor Product of Matrices

The tensor product of two matrices, denoted as $A \otimes B$, is a fundamental operation in linear algebra. If A is an $m \times n$ matrix and B is a $p \times q$ matrix, then the tensor product $A \otimes B$ results in an $mp \times nq$ block matrix.

The elements of $A \otimes B$ are formed by multiplying each element of A by the entire matrix B, resulting in blocks of B scaled by the elements of A.

Formally:

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}$$

Example E.1

Consider two matrices:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

$$B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$$

The tensor product $A \otimes B$ will be:

$$A \otimes B = \begin{bmatrix} 1 \times B & 2 \times B \\ 3 \times B & 4 \times B \end{bmatrix} = \begin{bmatrix} 5 & 6 & 10 & 12 \\ 7 & 8 & 14 & 16 \\ 15 & 18 & 20 & 24 \\ 21 & 24 & 28 & 32 \end{bmatrix}$$