DEPARTMENT OF PHYSICS
UNIVERSITY OF CAMBRIDGE

PART II EXPERIMENTAL AND
THEORETICAL PHYSICS

Theoretical Physics 2

Lecture Notes and Examples

B R Webber and N R Cooper

Lent Term 2004

Preface

In this course, we cover the necessary mathematical tools that underpin modern theoretical physics. We examine topics in quantum mechanics (with which you have some familiarity from previous courses) and apply the mathematical tools learnt in the IB Mathematics course (complex analysis, differential equations, matrix methods, special functions etc.) to topics like perturbation theory, scattering theory, etc. A course outline is provided below. Items indicated by a * are non-examinable material. They are there to illustrate the application of the course material to topics that you will come across in the PartII/Part III Theoretical Physics options. While we have tried to make the notes as self-contained as possible, you are encouraged to read the relevant sections of the recommended texts listed below. Throughout the notes, there are "mathematical interlude" sections reminding you of the the maths you are supposed to have mastered in the IB course. The "worked examples" are used to illustrate the concepts and you are strongly encouraged to work through every step, to ensure that you master these concepts and the mathematical techniques.

We are most grateful to Dr Guna Rajagopal for preparing the lecture notes of which these are an updated version.

Course Outline

- **Operator Methods in Quantum Mechanics (2 lectures):** Mathematical foundations of non-relativistic quantum mechanics; vector spaces; operator methods for discrete and continuous eigenspectra; generalized form of the uncertainty principle; simple harmonic oscillator; delta-function potential; introduction to second quantization.
- **Angular Momentum (2 lectures):** Eigenvalues/eigenvectors of the angular momentum operators (orbital/spin); spherical harmonics and their applications; Pauli matrices and spinors; addition of angular momenta.
- **Approximation Methods for Bound States (2 lectures):** Variational methods and their application to problems of interest; perturbation theory (time-independent and time depen-

dent) including degenerate and non-degenerate cases; the JWKB method and its application to barrier penetration and radioactive decay.

- **Scattering Theory (2 lectures):** Scattering amplitudes and differential cross-section; partial wave analysis; the optical theorem; Green functions; weak scattering and the Born approximation; *relation between Born approximation and partial wave expansions; *beyond the Born approximation.
- **Identical Particles in Quantum Mechanics (2 lectures):** Wave functions for non-interacting systems; symmetry of many-particle wave functions; the Pauli exclusion principle; fermions and bosons; exchange forces; the hydrogen molecule; scattering of identical particles; *second quantization method for many-particle systems; *pair correlation functions for bosons and fermions;
- **Density Matrices (2 lectures):** Pure and mixed states; the density operator and its properties; position and momentum representation of the density operator; applications in statistical mechanics.

Problem Sets

The problem sets (integrated within the lecture notes) are a vital and integral part of the course. The problems have been designed to reinforce key concepts and mathematical skills that you will need to master if you are serious about doing theoretical physics. Many of them will involve significant algebraic manipulations and it is vital that you gain the ability to do these long calculations without making careless mistakes! They come with helpful hints to guide you to their solution. Problems that you may choose to skip on a first reading are indicated by †.

Books

There is no single book that covers all of material in this course to the conceptual level or mathematical rigour required. Below are some books that come close. Liboff is at the right level for this course and it is particularly strong on applications. Sakurai is more demanding mathematically although he makes a lot of effort to explain the concepts clearly. This book is a recommended text in many graduate schools. Reed and Simon show what is involved in a mathematically rigorous treatment.

At about the level of the course: Liboff, **Quantum Mechanics,** 3^{rd} **Ed.**, Addison-Wesley.

At a more advanced level: Sakurai, **Quantum Mechanics,** 2^{nd} **Ed.**, Addison-Wesley;
Reed and Simon, **Methods of Modern Mathematical Physics**, Academic Press.

Contents

Chapter 1

Operator Methods In Quantum Mechanics

1.1 Introduction

The purpose of the first two lectures is twofold. First, to review the mathematical formalism of elementary non-relativistic quantum mechanics, especially the terminology. The second purpose is to present the basic tools of operator methods, commutation relations, shift operators, etc. and apply them to familiar problems such as the harmonic oscillator. Before we get down to the operator formalism, let's remind ourselves of the fundamental postulates of quantum mechanics as covered in earlier courses. They are:

- **Postulate 1:** The state of a quantum-mechanical system is completely specified by a function $\Psi(\mathbf{r}, t)$ (which in general can be complex) that depends on the coordinates of the particles (collectively denoted by $\mathbf{r}$) and on the time. This function, called the *wave function* or the state function, has the important property that $\Psi^*(\mathbf{r}, t)\Psi(\mathbf{r}, t)\, d\mathbf{r}$ is the probability that the system will be found in the volume element $d\mathbf{r}$, located at $\mathbf{r}$, at the time t.

- **Postulate 2:** To every observable A in classical mechanics, there corresponds a *linear Hermitian operator* $\hat{A}$ in quantum mechanics.

- **Postulate 3:** In any measurement of the observable A, the only values that can be obtained are the *eigenvalues* $\{a\}$ of the associated operator $\hat{A}$, which satisfy the eigenvalue equation

$$\hat{A}\Psi_a = a\Psi_a$$

where Ψ_a is the *eigenfunction* of $\hat{A}$ corresponding to the eigenvalue a.

- **Postulate 4:** If a system is in a state described by a normalised wavefunction Ψ, and the eigenfunctions $\{\Psi_a\}$ of $\hat{A}$ are also normalised, then the probability of obtaining the value a in a measurement of the observable A is given by

  $$P(a) = \left| \int_{-\infty}^{\infty} \Psi_a^* \Psi \, d\mathbf{r} \right|^2$$

  (Recall that a function $\Phi(\mathbf{r})$ such that

  $$\int_{-\infty}^{\infty} \Phi^* \Phi \, d\mathbf{r} = 1$$

  is said to be *normalised*.)

- **Postulate 5:** As a result of a measurement of the observable A in which the value a is obtained, the wave function of the system becomes the corresponding eigenfunction Ψ_a. (This is sometimes called the *collapse of the wave function*.)

- **Postulate 6:** Between measurements, the wave function evolves in time according to the *time-dependent Schrödinger equation*

  $$\frac{\partial \Psi}{\partial t} = -\frac{i}{\hbar} \hat{H} \Psi$$

  where $\hat{H}$ is the Hamiltonian operator of the system.

The justification for the above postulates ultimately rests with experiment. Just as in geometry one sets up axioms and then logically deduces the consequences, one does the same with the postulates of QM. To date, there has been no contradiction between experimental results and the outcomes predicted by applying the above postulates to a wide variety of systems.

We now explore the mathematical structure underpinning quantum mechanics.

1.1.1 Mathematical foundations

In the standard formulation of quantum theory, the state of a physical system is described by a vector in a Hilbert space H over the complex numbers. The observables and dynamical variables of the system are represented by linear operators which transform each state vector into another (possibly the same) state vector. Throughout this course (unless stated otherwise) we will adopt Dirac's notation: thus a state vector is denoted by a *ket* $|\Psi\rangle$. This ket provides a complete description of the physical state. In the next section we will explore the mathematical properties of the Hilbert space and learn why it plays such a central role in the mathematical formulation of quantum mechanics.

1.1.2 Hilbert space

A Hilbert space H,

$$H = \{|a\rangle, |b\rangle, |c\rangle, \ldots\}, \tag{1.1}$$

is a *linear vector space* over the field of complex number $\mathbf{C}$ i.e. it is an abstract set of elements (called vectors) with the following properties

1. $\forall\ |a\rangle, |b\rangle \in H$ we have
   - $|a\rangle + |b\rangle \in H$ (closure property)
   - $|a\rangle + |b\rangle = |b\rangle + |a\rangle$ (commutative law)
   - $(|a\rangle + |b\rangle) + |c\rangle = |a\rangle + (|b\rangle) + |c\rangle)$ (associative law)
   - $\exists$ a *null vector*, $|\text{null}\rangle \in H$ with the property

     $$|a\rangle + |\text{null}\rangle = |a\rangle \tag{1.2}$$

   - $\forall\ |a\rangle \in H\ \exists\ |-a\rangle \in H$ such that

     $$|a\rangle + |-a\rangle = |\text{null}\rangle \tag{1.3}$$

   - $\forall\ \alpha, \beta \in \mathbf{C}$

     $$\begin{aligned} \alpha(|a\rangle + |b\rangle) &= \alpha|a\rangle + \alpha|b\rangle && (1.4)\\ (\alpha+\beta)|a\rangle &= \alpha|a\rangle + \beta|a\rangle && (1.5)\\ (\alpha\beta)|a\rangle &= \alpha(\beta|a\rangle) && (1.6)\\ 1|a\rangle &= |a\rangle && (1.7) \end{aligned}$$

2. A *scalar product* is defined in H. It is denoted by $(|a\rangle, |b\rangle)$ or $\langle a|b\rangle$, yielding a complex number. The scalar product has the following properties

   $$\begin{aligned} (|a\rangle, \lambda|b\rangle) &= \lambda(|a\rangle, |b\rangle) && (1.8)\\ (|a\rangle, |b\rangle + |c\rangle) &= (|a\rangle, |b\rangle) + (|a\rangle, |c\rangle) && (1.9)\\ (|a\rangle, |b\rangle) &= (|b\rangle, |a\rangle)^* && (1.10) \end{aligned}$$

The last equation can also be written as

$$\langle a|b\rangle = \langle b|a\rangle^* \tag{1.11}$$

From the above, we can deduce that

$$\begin{aligned} (\lambda|a\rangle, |b\rangle) &= \lambda^*(|a\rangle, |b\rangle) && (1.12) \\ &= \lambda^*\langle a|b\rangle && (1.13) \end{aligned}$$

and

$$\begin{aligned} (|a_1\rangle + |a_2\rangle, |b\rangle) &= (|a_1\rangle, |b\rangle) + (|a_2\rangle, |b\rangle) && (1.14) \\ &= \langle a_1|b\rangle + \langle a_2|b\rangle && (1.15) \end{aligned}$$

The *norm* of a vector is defined by

$$\|a\| = \sqrt{\langle a|a\rangle} \tag{1.16}$$

and corresponds to the "length" of a vector. Note that the norm of a vector is a real number ≥ 0. (This follows from (1.11)).

1.1.3 The Schwartz inequality

Given any $|a\rangle, |b\rangle \in H$ we have

$$\|a\| \;\; \|b\| \geq |\langle a|b\rangle| \tag{1.17}$$

with the equality only being valid for the case

$$|a\rangle = \lambda|b\rangle \tag{1.18}$$

(with λ a complex number) i.e. when one vector is proportional to the other.

Proof

Define a $|c\rangle$ such that

$$|c\rangle = |a\rangle + \lambda|b\rangle \tag{1.19}$$

where λ is an arbitrary complex number. Whatever λ may be:

$$\begin{aligned} \langle c|c\rangle &= \langle a|a\rangle + \lambda\langle a|b\rangle + \lambda^*\langle b|a\rangle + \lambda\lambda^*\langle b|b\rangle && (1.20) \\ &\geq 0 && (1.21) \end{aligned}$$

Choose for λ the value

$$\lambda = -\frac{\langle b|a\rangle}{\langle b|b\rangle} \tag{1.22}$$

and substitute into the above equation, which reduces to

$$\langle a|a\rangle - \frac{\langle a|b\rangle\langle b|a\rangle}{\langle b|b\rangle} \geq 0 \tag{1.23}$$

Since $\langle b|b\rangle$ is positive, multiply the above inequality by $\langle b|b\rangle$ to get

$$\langle a|a\rangle\langle b|b\rangle \geq \langle a|b\rangle\langle b|a\rangle \tag{1.24}$$

$$\geq |\langle a|b\rangle|^2 \tag{1.25}$$

and finally taking square roots and using the definition of the *norm* we get the required result. (This result will be used when we prove the *generalised uncertainty principle*).

1.1.4 Some properties of vectors in a Hilbert space

$\forall \ |a\rangle \in H$, a sequence $\{|a_n\rangle\}$ of vectors exists, with the property that for every $\epsilon > 0$, there exists at least one vector $|a_n\rangle$ of the sequence with

$$\||a\rangle - |a_n\rangle\| \leq \epsilon \tag{1.26}$$

A sequence with this property is called *compact.*

The Hilbert space is *complete* i.e. every $|a\rangle \in H$ can be arbitrarily closely approximated by a sequence $\{|a_n\rangle\}$, in the sense that

$$\lim_{n\to\infty} \||a\rangle - |a_n\rangle\| = 0 \tag{1.27}$$

Then the sequence $\{|a_n\rangle\}$ has a unique limiting value $|a\rangle$.

The above properties are necessary for vector spaces of infinite dimension that occur in QM.

1.1.5 Orthonormal systems

Orthogonality of vectors. $|a\rangle$, $|b\rangle \in H$ are said to be **orthogonal** if

$$\langle a|b\rangle = 0 \tag{1.28}$$

Orthonormal system. The set $\{|a_n\rangle\}$ of vectors is an **orthonormal** system if the vectors are orthogonal and normalised, i.e.

$$\langle a_n|a_m\rangle = \delta_{n,m} \tag{1.29}$$

where

$$\delta_{n,m} = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}$$

Complete orthonormal system. The orthonormal system $\{|a_n\rangle\}$ is **complete** in H if an arbitrary vector $|a\rangle \in H$ can be expressed as

$$|a\rangle = \sum_n \alpha_n |a_n\rangle \tag{1.30}$$

where in general α_n are complex numbers whose values are

$$\alpha_m = \langle a_m|a\rangle \tag{1.31}$$

Proof

$$\begin{aligned} \langle a_m|a\rangle &= \langle a_m|\Big(\sum_n \alpha_n |a_n\rangle\Big) \\ &= \sum_n \alpha_n \langle a_m|a_n\rangle \\ &= \sum_n \alpha_n \delta_{m,n} \\ &= \alpha_m \end{aligned} \tag{1.32}$$

Thus we can write

$$|a\rangle = \sum_n |a_n\rangle\langle a_n|a\rangle \tag{1.33}$$

Note that this implies

$$\hat{I} = \sum_n |a_n\rangle\langle a_n| \tag{1.34}$$

called the "**resolution of the identity operator**" or the **closure relation**. The complex numbers α_n are called the a_n−representation of $|a\rangle$, i.e. they are the components of the vector $|a\rangle$ in the basis $\{|a_n\rangle\}$.

1.1.6 Operators on Hilbert space

A linear operator $\hat{A}$ induces a mapping of H onto itself or onto a subspace of H. (What this means is that if $\hat{A}$ acts on some arbitrary vector $\in H$ the result is another vector $\in H$ or in some subset of H. Hence

$$\hat{A}(\alpha|a\rangle + \beta|b\rangle) = \alpha\hat{A}|a\rangle + \beta\hat{A}|b\rangle \tag{1.35}$$

The operator $\hat{A}$ is *bounded* if

$$\|\hat{A}|a\rangle\| \leq C\||a\rangle\| \tag{1.36}$$

$\forall\ |a\rangle \in H$, and C is a real positive constant ($< \infty$).

Bounded linear operators are *continuous*, i.e. if

$$|a_n\rangle \to |a\rangle \tag{1.37}$$

then it follows that

$$\hat{A}|a_n\rangle \to \hat{A}|a\rangle \tag{1.38}$$

Two operators $\hat{A}$ and $\hat{B}$ are *equal* ($\hat{A} = \hat{B}$) if, $\forall |a\rangle \in H$,

$$\hat{A}|a\rangle = \hat{B}|a\rangle \tag{1.39}$$

The following definitions are valid $\forall\ \ |a\rangle \in H$:

Unit operator, $\hat{I}$

$$\hat{I}|a\rangle = |a\rangle \tag{1.40}$$

Zero operator, $\hat{0}$

$$\hat{0}|a\rangle = |\text{null}\rangle \tag{1.41}$$

Sum operator, $\hat{A} + \hat{B}$

$$(\hat{A} + \hat{B})|a\rangle = \hat{A}|a\rangle + \hat{B}|a\rangle \tag{1.42}$$

Product operator, $\hat{A}\hat{B}$

$$(\hat{A}\hat{B})|a\rangle = \hat{A}(\hat{B}|a\rangle) \tag{1.43}$$

Adjoint operator, $\hat{A}^\dagger$: Given $\hat{A}$, an *adjoint operator*, $\hat{A}^\dagger$, exists if $\forall\ |a\rangle, |b\rangle \in H$

$$(|b\rangle, \hat{A}|a\rangle) = (\hat{A}^\dagger|b\rangle, |a\rangle) \tag{1.44}$$

or

$$\langle b|\hat{A}|a\rangle = \langle a|\hat{A}^\dagger|b\rangle^* \tag{1.45}$$

The adjoint of an operator has the following properties:

$$(\alpha\hat{A})^\dagger = \alpha^*\hat{A}^\dagger \tag{1.46}$$

$$(\hat{A} + \hat{B})^\dagger = \hat{A}^\dagger + \hat{B}^\dagger \tag{1.47}$$

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger \tag{1.48}$$

$$(\hat{A}^\dagger)^\dagger = \hat{A} \tag{1.49}$$

If $\hat{A}$ is Hermitian, then

$$\begin{aligned} \hat{A} &= \hat{A}^\dagger \\ \langle b|\hat{A}|b\rangle &= \langle b|\hat{A}^\dagger|b\rangle \\ &= \langle b|\hat{A}^\dagger|b\rangle^* \\ &= \langle b|\hat{A}|b\rangle^* \\ &= \text{real} \end{aligned} \tag{1.50}$$

Unitary operator, U : The operator $\hat{U}$ is called *unitary* if

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I} \tag{1.51}$$

Projection operator, $|a\rangle\langle a|$: Given any normalised vector $|a\rangle$, a *projection operator* $\hat{P}$ can be defined as the operator that projects any vector into its component along $|a\rangle$

$$\hat{P}|b\rangle = \langle a|b\rangle|a\rangle = |a\rangle\langle a|b\rangle \tag{1.52}$$

We write this symbolically as

$$\hat{P} = |a\rangle\langle a| \tag{1.53}$$

Note that a projection operator is *idempotent*: its square (or any power) is equal to itself

$$\hat{P}^2 = |a\rangle\langle a|a\rangle\langle a| = |a\rangle\langle a| \tag{1.54}$$

since $|a\rangle$ is normalised. Note that the resolution of the identity (1.34) is a sum of projection operators.

Commutator, $[\hat{A},\hat{B}]$

$$[\hat{A},\hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \tag{1.55}$$

Note that in general

$$\hat{A}\hat{B} \neq \hat{B}\hat{A} \tag{1.56}$$

Properties of commutators:

$$[\hat{A},\hat{B}] = -[\hat{B},\hat{A}] \tag{1.57}$$

$$[\hat{A},(\hat{B}+\hat{C})] = [\hat{A},\hat{B}] + [\hat{A},\hat{C}] \tag{1.58}$$

$$[\hat{A},\hat{B}\hat{C}] = [\hat{A},\hat{B}]\hat{C} + \hat{B}[\hat{A},\hat{C}] \tag{1.59}$$

$$[\hat{A},[\hat{B},\hat{C}]] + [\hat{B},[\hat{C},\hat{A}]] + [\hat{C},[\hat{A},\hat{B}]] = \hat{0} \tag{1.60}$$

$$[\hat{A},\hat{B}]^\dagger = [\hat{B}^\dagger,\hat{A}^\dagger] \tag{1.61}$$

EXAMPLE

Suppose the operators $\hat{P}$ and $\hat{Q}$ satisfy the commutation relation

$$[\hat{P}, \hat{Q}] = a\hat{I}$$

where a is a constant (real) number.

- Reduce the commutator $[\hat{P}, \hat{Q}^n]$ to its simplest possible form.

  Answer: Let

  $$\hat{R}_n = [\hat{P}, \hat{Q}^n] \quad n = 1, 2, \cdots$$

  Then $\hat{R}_1 = [\hat{P}, \hat{Q}] = a\hat{I}$ and

  $$\hat{R}_{n+1} = [\hat{P}, \hat{Q}^{n+1}] = [\hat{P}, \hat{Q}^n\hat{Q}] = [\hat{P}, \hat{Q}^n]\hat{Q} + \hat{Q}^n[\hat{P}, \hat{Q}]$$

  (We have used $[\hat{A}, \hat{B}\hat{C}] = \hat{B}[\hat{A}, \hat{C}] + [\hat{A}, \hat{B}]\hat{C}$). Therefore,

  $$\hat{R}_{n+1} = \hat{R}_n\hat{Q} + \hat{Q}^n(a\hat{I}) = \hat{R}_n\hat{Q} + a\hat{Q}^n$$

  which gives $\hat{R}_2 = 2a\hat{Q}$, $\hat{R}_3 = 3a\hat{Q}^2$ etc. This implies that

  $$\hat{R}_n = [\hat{P}, \hat{Q}^n] = na\hat{Q}^{n-1}$$

  Note that in general,

  $$[\hat{P}, f(\hat{Q})] = a\frac{\partial f}{\partial \hat{Q}}$$

- Reduce the commutator

  $$[\hat{P}, e^{i\hat{Q}}]$$

  to its simplest form.

  Answer: Use results above to get

  $$[\hat{P}, e^{i\hat{Q}}] = iae^{i\hat{Q}}$$

Problem 1: Two operators, $\hat{A}$ and $\hat{B}$ satisfy the equations

$$\begin{aligned} \hat{A} &= \hat{B}^\dagger \hat{B} + 3 \\ \hat{A} &= \hat{B}\hat{B}^\dagger + 1 \end{aligned} \tag{1.62}$$

- Show that $\hat{A}$ is self-adjoint
- Find the commutator $[\hat{B}^\dagger, \hat{B}]$

  Answer: $-2\hat{I}$
- Find the commutator $[\hat{A}, \hat{B}]$

  Answer: $-2\hat{B}$

1.1.7 Eigenvectors and eigenvalues

If

$$\hat{A}|a\rangle = a|a\rangle \tag{1.63}$$

then $|a\rangle$ is an **eigenvector** of the operator $\hat{A}$ with **eigenvalue** a (which in general is a complex number). The set of all eigenvalues of a operator is called its **spectrum**, which can take discrete or continuous values (or both). For the case of Hermitian operators the following is true:

- The eigenvalues are real
- The eigenvectors corresponding to different eigenvalues are orthogonal i.e

  $$\hat{A}|a\rangle = a|a\rangle \tag{1.64}$$

  $$\hat{A}|a'\rangle = a'|a'\rangle \tag{1.65}$$

  and if $a \neq a'$, then

  $$\langle a|a'\rangle = 0 \tag{1.66}$$
- In addition, the normalised eigenvectors of a **bounded** Hermitian operator give rise to a *countable, complete orthonormal system.* The eigenvalues form a discrete spectrum.

Problem 2: Prove that if $\hat{H}$ is a Hermitian operator, then its eigenvalues are real and its eigenvectors (corresponding to different eigenvalues) are orthogonal.

Answer: To be discussed in class.

From above, we deduce that an arbitrary $|\psi\rangle \in H$ can be expanded in terms of the complete, orthonormal eigenstates $\{|a\rangle\}$ of a Hermitian operator $\hat{A}$:

$$|\psi\rangle = \sum_a |a\rangle\langle a|\psi\rangle \tag{1.67}$$

where the infinite set of complex numbers $\{\langle a|\psi\rangle\}$ are called the A *representation* of $|\psi\rangle$.

Problem 3: The operator $\hat{Q}$ satisfies the equations

$$\hat{Q}^\dagger\hat{Q}^\dagger = 0$$

$$\hat{Q}\hat{Q}^\dagger + \hat{Q}^\dagger\hat{Q} = \hat{I} \tag{1.68}$$

The Hamiltonian for the system is given by

$$\hat{H} = \alpha\hat{Q}\hat{Q}^\dagger$$

where α is a real constant.

- Show that $\hat{H}$ is self-adjoint
- Find an expression for $\hat{H}^2$ in terms of $\hat{H}$

  Answer: Use the anti-commutator property of $\hat{Q}$ to get $\hat{H}^2 = \alpha\hat{H}$.
- Deduce the eigenvalues of $\hat{H}$ using the results obtained above.

  *Answer:*The eigenvalues are 0 and α.

Problem 4 : Manipulating Operators

- **Show that** if $|a\rangle$ is an eigenvector of $\hat{A}$ with eigenvalue a, then it is an eigenvector of $f(\hat{A})$ with eigenvalue $f(a)$.
- **Show that**
$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger \tag{1.69}$$
and in general
$$(\hat{A}\hat{B}\hat{C}\ldots)^\dagger = \ldots\hat{C}^\dagger\hat{B}^\dagger\hat{A}^\dagger \tag{1.70}$$
- **Show that** $\hat{A}\hat{A}^\dagger$ is Hermitian even if $\hat{A}$ is not.
- **Show that** if $\hat{A}$ is Hermitian, then the expectation value of $\hat{A}^2$ are non-negative, and the eigenvalues of $\hat{A}^2$ are non-negative.
- Suppose there exists a linear operator $\hat{A}$ that has an eigenvector $|\psi\rangle$ with eigenvalue a. If there also exists an operator $\hat{B}$ such that
$$[\hat{A},\hat{B}] = \hat{B} + 2\hat{B}\hat{A}^2 \tag{1.71}$$
then **show that** $\hat{B}|\psi\rangle$ is an eigenvector of $\hat{A}$ and find the eigenvalue.

  Answer: Eigenvalue is $1 + a + 2a^2$.

EXAMPLE

- (a) Suppose the operators $\hat{A}$ and $\hat{B}$ commute with their commutator, i.e. $[\hat{B},[\hat{A},\hat{B}]] = [\hat{A},[\hat{A},\hat{B}]] = 0$. Show that $[\hat{A},\hat{B}^n] = n\hat{B}^{n-1}[\hat{A},\hat{B}]$ and $[\hat{A}^n,\hat{B}] = n\hat{A}^{n-1}[\hat{A},\hat{B}]$.

  Answer: To show this, consider the following steps:

$$\begin{aligned}
[\hat{A},\hat{B}^n] &= \hat{A}\hat{B}^n - \hat{B}^n\hat{A} \qquad (1.72)\\
&= \hat{A}\hat{B}\hat{B}^{n-1} - \hat{B}\hat{A}\hat{B}^{n-1} + \hat{B}(\hat{A}\hat{B})\hat{B}^{n-2} - \hat{B}(\hat{B}\hat{A})\hat{B}^{n-3} + \cdots \hat{B}^{n-1}\hat{A}\hat{B} - \hat{B}^{n-1}\hat{B}\hat{A}\\
&= [\hat{A},\hat{B}]\hat{B}^{n-1} + \hat{B}[\hat{A},\hat{B}]\hat{B}^{n-2} + \cdots + \hat{B}^{n-1}[\hat{A},\hat{B}]
\end{aligned}$$

Since $\hat{B}$ commutes with $[\hat{A},\hat{B}]$, we obtain

$$[\hat{A},\hat{B}^n]=\hat{B}^{n-1}[\hat{A},\hat{B}]+\hat{B}^{n-1}[\hat{A},\hat{B}]+\cdots+\hat{B}^{n-1}[\hat{A},\hat{B}]=n\hat{B}^{n-1}[\hat{A},\hat{B}]$$

as required. In the same way, since $[\hat{A}^n,B]=-[\hat{B},\hat{A}^n]$ and using the above steps, we obtain

$$[\hat{A}^n,\hat{B}]=n\hat{A}^{n-1}[\hat{A},\hat{B}]$$

as required.

- (b) Just as in (a), show that for any analytic function, $f(x)$, we have $[\hat{A},f(\hat{B})]=[\hat{A},\hat{B}]f'(\hat{B})$, where $f'(x)$ denotes the derivative of $f(x)$.

  Answer: We use the results from (a). Since $f(x)$ is analytic, we can expand it in a power series $\sum_n a_n x^n$. Then

$$\begin{aligned}[\hat{A},f(\hat{B})] &= [\hat{A},\sum_n a_n\hat{B}^n] \\ &= \sum_n a_n[\hat{A},\hat{B}^n] \\ &= [\hat{A},\hat{B}]\sum_n n\ a_n\ \hat{B}^{n-1} \\ &= [\hat{A},\hat{B}]f^{'}(\hat{B})\end{aligned}\tag{1.73}$$

- (c) Just as in (a), show that $e^{\hat{A}}e^{\hat{B}}=e^{\hat{A}+\hat{B}}\ e^{\frac{1}{2}[\hat{A},\hat{B}]}$.

  Answer: Consider an operator $\hat{F}(s)$ which depends on a real parameter s:

$$\hat{F}(s)=e^{s\hat{A}}\ e^{s\hat{B}}$$

  Its derivative with respect to s is:

$$\begin{aligned}\frac{d\hat{F}}{ds} &= \left(\frac{d}{ds}e^{s\hat{A}}\right)e^{s\hat{B}}+e^{s\hat{A}}\left(\frac{d}{ds}e^{s\hat{B}}\right) \\ &= \hat{A}e^{s\hat{A}}e^{s\hat{B}}+e^{s\hat{A}}\hat{B}e^{s\hat{B}} \\ &= \hat{A}e^{s\hat{A}}e^{s\hat{B}}+e^{s\hat{A}}\hat{B}e^{-s\hat{A}}e^{s\hat{A}}e^{s\hat{B}} \\ &= \left[\hat{A}+e^{s\hat{A}}\hat{B}e^{-s\hat{A}}\right]\hat{F}(s)\end{aligned}\tag{1.74}$$

  Using part (a), we can write

$$[e^{s\hat{A}},\hat{B}]=-[\hat{B},e^{s\hat{A}}]=-s[\hat{B},\hat{A}]e^{s\hat{A}}=s[\hat{A},\hat{B}]e^{s\hat{A}}$$

This means that $e^{s\hat{A}}\hat{B} = \hat{B}e^{-s\hat{A}} + s[\hat{A},\hat{B}]e^{s\hat{A}}$ and $e^{s\hat{A}}\hat{B}e^{-s\hat{A}} = \hat{B} + s[\hat{A},\hat{B}]$. Substituting this into the equation above, we get

$$\frac{d\hat{F}}{ds} = \left[\hat{A} + \hat{B} + s[\hat{A},\hat{B}]\right] \hat{F}(s)$$

Since $\hat{A} + \hat{B}$ and $[\hat{A},\hat{B}]$ commute, we can integrate this differential equation. This yields

$$\hat{F}(s) = \hat{F}(0)\, e^{(\hat{A}+\hat{B})s+\frac{1}{2}[\hat{A},\hat{B}]s^2}$$

Setting $s = 0$, we obtain $\hat{F}(0) = \hat{I}$. Finally substituting $\hat{F}(0)$ and $s = 1$, we obtain the required result.

- (d) Prove the following identity for *any* two operators $\hat{A}$ and $\hat{B}$:

$$e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + [\hat{A},\hat{B}] + \frac{1}{2!}[\hat{A},[\hat{A},\hat{B}]] + \frac{1}{3!}[\hat{A},[\hat{A},[\hat{A},\hat{B}]]] + \cdots \tag{1.75}$$

Answer: To show this, define

$$f(\lambda) = e^{\lambda\hat{A}}\hat{B}e^{-\lambda\hat{A}}$$

where λ is a real parameter. Then,

$$\begin{aligned} f(0) &= \hat{B} \qquad (1.76)\\ f(1) &= e^{\hat{A}}\hat{B}e^{-\hat{A}} \\ f'(\lambda) &= e^{\lambda\hat{A}}[\hat{A},\hat{B}]e^{-\lambda\hat{A}} \\ f'(0) &= [\hat{A},\hat{B}] \\ f''(\lambda) &= e^{\lambda\hat{A}}[\hat{A},[\hat{A},\hat{B}]]e^{-\lambda\hat{A}} \\ f''(0) &= [\hat{A},[\hat{A},\hat{B}]] \end{aligned}$$

The Taylor expansion of $f(\lambda)$ is given by

$$f(\lambda) = f(0) + \lambda f'(0) + \frac{1}{2!}\lambda^2 f''(0) + \cdots$$

This implies

$$e^{\lambda\hat{A}}\hat{B}e^{-\lambda\hat{A}} = \hat{B} + \lambda[\hat{A},\hat{B}] + \frac{1}{2!}\lambda^2[\hat{A},[\hat{A},\hat{B}]] + \cdots$$

Now setting $\lambda = 1$, we get the required result.

1.1.8 Observables

A Hermitian operator $\hat{A}$ is an **observable** if its eigenvectors $|\psi_n\rangle$ are a basis in the Hilbert space: that is, if an arbitrary state vector can be written as

$$|\psi\rangle = \sum_{n=1}^{D} |\psi_n\rangle\langle\psi_n|\psi\rangle \tag{1.77}$$

(If D, the dimensionality of the Hilbert space is finite, then all Hermitian operators are observables; if D is infinite, this is not necessarily so.)

In quantum mechanics, it is a postulate that every measurable physical quantity is described by an observable and that the only possible result of the measurement of a physical quantity is one of the eigenvalues of the corresponding observable. Immediately after an observation of $\hat{A}$ which yields the eigenvalue a_n, the system is in the corresponding state $|\psi_n\rangle$. It is also a postulate that the probability of obtaining the result a_n when observing $\hat{A}$ on a system in the normalised state $|\psi\rangle$, is

$$P(a_n) = |\langle\psi_n|\psi\rangle|^2 \tag{1.78}$$

(The probability is determined empirically by making a large number of separate observations of $\hat{A}$, each observation being made on a copy of the system in the state $|\psi\rangle$.) The normalisation of $|\psi\rangle$ and the closure relation ensure that

$$\sum_{n=1}^{D} P(a_n) = 1 \tag{1.79}$$

For an observable, by using the closure relation, one can deduce that

$$\hat{A} = \sum_n a_n |\psi_n\rangle\langle\psi_n| \tag{1.80}$$

which is the **spectral decomposition** of $\hat{A}$.

The **expectation value** $\langle\hat{A}\rangle$ of an observable $\hat{A}$, when the state vector is $|\psi\rangle$, is defined as the average value obtained in the limit of a large number of separate observations of $\hat{A}$, each made on a copy of the system in the state $|\psi\rangle$. From equations (1.78) and (1.80), we have

$$\begin{aligned}\langle\hat{A}\rangle &= \sum_n a_n P(a_n) &= \sum_n a_n |\langle\psi_n|\psi\rangle|^2 \\ &= \sum_n a_n \langle\psi|\psi_n\rangle\langle\psi_n|\psi\rangle &= \langle\psi|\hat{A}|\psi\rangle\end{aligned} \tag{1.81}$$

Let $\hat{A}$ and $\hat{B}$ be two observables and suppose that rapid successive measurements yield the results a_n and b_n respectively. If immediate repetition of the observations always yields the same results for all possible values of a_n and b_n, then $\hat{A}$ and $\hat{B}$ are **compatible** (or non-interfering) observables.

Problem 5: A system described by the Hamiltonian $\hat{H}_0$ has just two orthogonal energy eigenstates, $|1\rangle$ and $|2\rangle$ with

$$\begin{aligned} \langle 1|1\rangle &= 1 \\ \langle 1|2\rangle &= 0 \\ \langle 2|2\rangle &= 1 \end{aligned} \tag{1.82}$$

The two eigenstates have the same eigenvalues E_0:

$$\hat{H}_0|i\rangle = E_0|i\rangle$$

for $i = 1, 2$. Suppose the Hamiltonian for the system is changed by the addition of the term $\hat{V}$, giving

$$\hat{H} = \hat{H}_0 + \hat{V}$$

The matrix elements of $\hat{V}$ are

$$\begin{aligned} \langle 1|\hat{V}|1\rangle &= 0 \\ \langle 1|\hat{V}|2\rangle &= V_{12} \\ \langle 2|\hat{V}|2\rangle &= 0 \end{aligned} \tag{1.83}$$

- Find the eigenvalues of $\hat{H}$
- Find the normalised eigenstates of $\hat{H}$ in terms of $|1\rangle$ and $|2\rangle$.

Answer: This will be done in class.

1.1.9 Generalised uncertainty principle

Suppose $\hat{A}$ and $\hat{B}$ are any two non-commuting operators i.e.

$$[\hat{A}, \hat{B}] = i\hat{C} \tag{1.84}$$

(where $\hat{C}$ is Hermitian). It can be shown that

$$\Delta A\,\Delta B \geq \frac{1}{2}\left|\langle\hat{C}\rangle\right| \tag{1.85}$$

where

$$\Delta A = \left[\langle(\hat{A} - \langle\hat{A}\rangle)^2\rangle\right]^{\frac{1}{2}} \tag{1.86}$$

and similarly for ΔB. The expectation value is over some arbitrary state vector. This is the **generalised uncertainty principle**, which implies that it is not possible for two non-commuting observables to possess a complete set of simultaneous eigenstates. In particular if $\hat{C}$ is a non-zero real number (times the unit operator), then $\hat{A}$ and $\hat{B}$ cannot possess any simultaneous eigenstates.

Problem 6: Prove (1.85).

If the eigenvalues of $\hat{A}$ are non-degenerate, the normalised eigenvectors $|\psi_n\rangle$ are unique to within a phase factor i.e. the kets $|\psi_n\rangle$ and $e^{i\theta}|\psi_n\rangle$, where θ is any real number yield the same physical results. Hence a well defined physical state can be obtained by measuring $\hat{A}$. If the eigenvalues of $\hat{A}$ are degenerate we can in principle identify additional observables $\hat{B}, \hat{C}, \ldots$ which commute with $\hat{A}$ and each other (but not functions of $\hat{A}$ or each other), until we have a set of commuting observables for which there is no degeneracy. Then the simultaneous eigenvectors $|a_n, b_p, c_q, \ldots\rangle$ are unique to within a phase factor; they are a basis for which the orthonormality relations are

$$\langle a_{n'}, b_{p'}, c_{q'}, \ldots | a_n, b_p, c_q, \ldots\rangle = \delta_{n'n}\delta_{p'p}\delta_{q'q}\ldots \tag{1.87}$$

The observables $\hat{A}, \hat{B}, \hat{C}, \ldots$ constitute a **complete set of commuting observables (CSCO)**. A well defined initial state can be obtained by an observation of a CSCO.

Problem 7: Given a set of observables $\hat{A}, \hat{B}, \ldots$ prove that any one of the following conditions proves the other two:

- $\hat{A}, \hat{B}, \ldots$ commute with each other,
- $\hat{A}, \hat{B}, \ldots$ are compatible,
- $\hat{A}, \hat{B}, \ldots$ possess a complete orthonormal set of simultaneous eigenvectors (assuming no degeneracy).

1.1.10 Basis transformations

Suppose $\{|\psi_n\rangle\}$ and $\{|\phi_n\rangle\}$ respectively are the eigenvectors of the non-commuting observables $\hat{A}$ and $\hat{B}$ of a system. This means that we can use either $\{|\psi_n\rangle\}$ or $\{|\phi_n\rangle\}$ as **basis** kets for the Hilbert space. The two bases are related by the transformation

$$|\phi_n\rangle = \hat{U}|\psi_n\rangle \tag{1.88}$$

where

$$\hat{U} = \sum_i |\phi_i\rangle\langle\psi_i| \tag{1.89}$$

Orthonormality of $\{|\phi_n\rangle\}$ and the closure relation for $\{|\psi_n\rangle\}$ ensure that $\hat{U}$ is a unitary operator (i.e. $\hat{U}^\dagger\hat{U} = \hat{I}$).

Problem 8:

- Prove that $\hat{U}$ as defined above is unitary.
- Starting from the eigenvalue equation:

  $$\hat{A}|\psi_n\rangle = a_n|\psi_n\rangle \tag{1.90}$$

  show that the operator

  $$\hat{A}' = \hat{U}\hat{A}\hat{U}^\dagger \tag{1.91}$$

  has $\hat{U}|\psi_n\rangle$ as its eigenvector with eigenvalue a_n.
- Show also that the inner product, $\langle\Psi|\Phi\rangle$ is preserved under a unitary transformation.
- If $\hat{U}$ is unitary and $\hat{A}$ is Hermitian, then show that $\hat{U}\hat{A}\hat{U}^\dagger$ is also Hermitian.
- Show that the form of the operator equation $\hat{G} = \hat{A}\hat{B}$ is preserved under a unitary transformation.

The problem above shows that a unitary transformation preserves the form of the eigenvalue equation. In addition, since the eigenvalues of an operator corresponding to an observable are physically measurable quantities, these values should not be affected by a transformation of basis in Hilbert space. It therefore follows that the eigenvalues and the Hermiticity of an observable are preserved in a unitary transformation.

1.1.11 Matrix representation of operators

From the closure relation (or resolution of the identity) it is possible to express any operator as

$$\hat{A} = \hat{I}\hat{A}\hat{I} = \sum_n \sum_{n'} |n'\rangle\langle n'|\hat{A}|n\rangle\langle n| \tag{1.92}$$

where the set $\{|n\rangle\}$ are a set of basis vectors in the Hilbert space and the complex numbers $\langle n'|\hat{A}|n\rangle$ are a **matrix representation** of $\hat{A}$. (Note that the matrix representation of $\hat{A}^\dagger$ is obtained by transposing the matrix representation of $\hat{A}$ and taking the complex conjugate of each element.) The table below lists various matrix properties:

Matrix	Definition	Matrix Elements
Symmetric	$A = A^T$	$A_{pq} = A_{qp}$
Antisymmetric	$A = -A^T$	$A_{pq} = -A_{qp}$
Orthogonal	$A = (A^T)^{-1}$	$(A^T A)_{pq} = \delta_{pq}$
Real	$A = A^*$	$A_{pq} = A^*_{pq}$
Pure Imaginary	$A = -A^*$	$A_{pq} = -A^*_{pq}$
Hermitian	$A = A^\dagger$	$A_{pq} = A^*_{qp}$
Anti-Hermitian	$A = -A^\dagger$	$A_{pq} = -A^*_{qp}$
Unitary	$A = (A^\dagger)^{-1}$	$(A^\dagger A)_{pq} = \delta_{pq}$
Singular	$\|A\| = 0$	

where T denotes the transpose of a matrix and $|A|$ denotes the determinant of matrix A.

Problem 9:

- If A, B, C are 3 $n \times n$ square matrices, show that $Tr(ABC) = Tr(CAB) = Tr(BCA)$, where Tr denotes the trace of a matrix, i.e. the sum of its diagonal elements.
- Show that the trace of a matrix remains the same (i.e. invariant) under a unitary transformation.
- Let A be an $n \times n$ square matrix with eigenvalues $a_1, a_2, \ldots, a_n$. Show that $|A| = a_1 a_2 \ldots a_n$ and hence that the determinant of A is another invariant property.
- Show that if A is Hermitian, then $U = (A + iI)(A - iI)^{-1}$ is unitary. (I here is the identity matrix.)

- Show that $|I + \epsilon A| = I + \epsilon\, TrA + O(\epsilon^2)$ where A is an $n \times n$ square matrix.
- Show that $|e^A| = e^{TrA}$ where A is a $n \times n$ square matrix.

1.1.12 Mathematical interlude: Dirac delta function

Definition

The Dirac delta function $\delta(x)$ is defined as follows

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

Its integral properties are

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\delta(x)dx &= f(0) \\ \int_{-\infty}^{\infty} \delta(x)dx &= 1 \\ \int_{-\infty}^{\infty} f(x')\delta(x-x')dx' &= f(x) \\ \int_{-\infty}^{\infty} \delta(x-x')dx' &= 1 \end{aligned} \tag{1.93}$$

Note that

$$\int_a^b f(x)\delta(x)dx = \begin{cases} f(0) & 0 \in [a,b] \\ 0 & \text{Otherwise} \end{cases}$$

In three dimensions, the above definition is generalised as follows

$$\int_{all\ space} f(\mathbf{r})\delta(\mathbf{r}-\mathbf{a})d\mathbf{r} = f(\mathbf{a}) \tag{1.94}$$

In mathematics, an object such as $\delta(x)$, which is defined in terms of its integral properties, is called a **distribution**.

Some useful properties

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

$$
\begin{aligned}
\delta^{'}(x) &= -\delta^{'}(-x) \\
x\;\delta(x) &= 0 \\
\delta(a\;x) &= \frac{1}{|a|}\;\delta(x) \\
\delta(x^2 - a^2) &= \frac{1}{|2a|}\Big[\delta(x-a) - \delta(x+a)\Big] \\
\int_{-\infty}^{\infty} \delta(a-x)\;\delta(x-b)dx &= \delta(a-b) \\
f(x)\;\delta(x-a) &= f(a)\;\delta(x-a) \\
x\;\delta^{'}(x) &= -\delta(x) \\
\int g(x)\;\delta\Big[f(x)-a\Big]\;dx &= \frac{g(x)}{|df/dx|}\Big|_{x=x_0,\;f(x_0)=a} \qquad (1.95)
\end{aligned}
$$

These relations can easily be verified by using some arbitrary function. For example, to prove

$$x\;\delta^{'}(x) = -\delta(x)$$

we proceed as follows

$$
\begin{aligned}
\int_{-\infty}^{\infty} f(x)x\;\delta^{'}(x)dx &= \int_{-\infty}^{\infty}\frac{d}{dx}\Big(f\;x\;\delta\Big)dx - \int_{-\infty}^{\infty}\delta\;\frac{d}{dx}\Big(f\;x\;\Big)dx \\
&= -\int_{-\infty}^{\infty}\delta(x)\Big(x\frac{df}{dx}+f\Big)dx \\
&= -\int_{-\infty}^{\infty}\delta(x)f(x)dx \qquad (1.96)
\end{aligned}
$$

where we have used integration by parts.

1.1.13 Operators with continuous or mixed (discrete-continuous) spectra

There exist operators which do not have a purely discrete spectra, but either have a continuous or mixed (discrete-continuous) spectrum. An example is the Hamiltonian for the hydrogen atom. In general, **all** Hamiltonians for atoms and nuclei have both discrete and continuous spectral ranges. Usually the discrete spectrum is connected with bound states while the continuous spectrum is connected with free (unbound) states. The representation related to such operators cause difficulties because eigenstates with continuous spectra are not normalizable to unity. (A rigorous discussion is too difficult so we will just state the results.)

An observable $\hat{A}$ has a continuous spectrum if its eigenvalues $\{a\}$

$$\hat{A}|a\rangle = a|a\rangle$$

are a continuous set of real numbers. The eigenstates $\{|a\rangle\}$ can no longer be normalised to unity but must be normalised to Dirac delta functions:

$$\langle a|a'\rangle = \delta(a-a') \tag{1.97}$$

The resolution of the identity (or closure relation) becomes

$$\int da\, |a\rangle\langle a| = \hat{I} \tag{1.98}$$

and an arbitrary state can $|\psi\rangle$ be expanded in terms of the complete set $\{|a\rangle\}$ via

$$|\psi\rangle = \int da'|a'\rangle\langle a'|\psi\rangle \tag{1.99}$$

with $\langle a'|\psi\rangle$ denoting $|\psi\rangle$ in the A representation. The inner product for two state vectors $|\psi\rangle$ and $|\phi\rangle$ is defined as

$$\begin{aligned} \langle\psi|\phi\rangle &= \int da'\langle\psi|a'\rangle\langle a'|\phi\rangle \\ &= \int \psi^*(a')\phi(a')da' \end{aligned} \tag{1.100}$$

If the spectrum is mixed, then the expansion of $|\psi\rangle$ is

$$|\psi\rangle = \sum_{a'} |a'\rangle\langle a'|\psi\rangle + \int |a'\rangle\langle a'|\psi\rangle da' \tag{1.101}$$

where the sum is over the discrete eigenvectors and the integral is over the continuous eigenvectors $|a\rangle$.

Position and momentum representations for free particles

In one dimension, the eigenvalue equations for $\hat{x}$ and $\hat{p}$ read

$$\begin{aligned} \hat{x}|x'\rangle &= x'|x'\rangle \\ \hat{p}|p'\rangle &= p'|p'\rangle \\ \langle x|x'\rangle &= \delta(x-x') \\ \langle p|p'\rangle &= \delta(p-p') \end{aligned} \tag{1.102}$$

These definitions, the fundamental commutator

$$[\hat{x}, \hat{p}] = i\hbar \tag{1.103}$$

and the above properties of the Dirac delta function can be used to determine the following matrix elements:

$$\begin{aligned}
\langle x'|\hat{p}|x''\rangle &= \frac{\hbar}{i}\frac{\partial}{\partial x'}\delta(x'-x'') \\
\langle p'|\hat{x}|p''\rangle &= -\frac{\hbar}{i}\frac{\partial}{\partial p'}\delta(p'-p'') \\
\langle x'|\hat{p}^2|x''\rangle &= \left(-i\hbar\frac{\partial}{\partial x'}\right)^2\delta(x'-x'') \\
\langle p'|\hat{x}^2|p''\rangle &= \left(i\hbar\frac{\partial}{\partial p'}\right)^2\delta(p'-p'')
\end{aligned} \tag{1.104}$$

Problem 10†[1]: Verify the formulae (1.104)

Now consider the eigenvalue problem for the momentum operator in the position representation. If

$$\hat{p}|p'\rangle = p'|p'\rangle$$

then we have

$$\begin{aligned}
\langle x'|\hat{p}|p'\rangle &= \int dx''\langle x'|\hat{p}|x''\rangle\langle x''|p'\rangle \\
&= \int dx''\left(-i\hbar\frac{\partial}{\partial x'}\delta(x'-x'')\right)\langle x''|p'\rangle \\
&= -i\hbar\frac{\partial}{\partial x'}\int dx''\delta(x'-x'')\langle x''|p'\rangle \\
&= -i\hbar\frac{\partial}{\partial x'}\langle x'|p'\rangle
\end{aligned} \tag{1.105}$$

On the other hand, we also have

$$\langle x'|\hat{p}|p'\rangle = p'\langle x'|p'\rangle$$

---

[1] Problems that you may choose to skip on a first reading are indicated by †.

Therefore

$$-i\hbar\frac{\partial}{\partial x'}\langle x'|p'\rangle = p'\langle x'|p'\rangle \tag{1.106}$$

which implies

$$\langle x'|p'\rangle = \frac{1}{\sqrt{2\pi\hbar}}\exp\left(\frac{ip'x'}{\hbar}\right) \tag{1.107}$$

where we have chosen the normalisation such that

$$\begin{aligned}
\langle p''|p'\rangle &= \int dx'\langle p''|x'\rangle\langle x'|p'\rangle \\
&= \int dx'\langle x'|p''\rangle^*\langle x'|p'\rangle \\
&= \frac{1}{(2\pi\hbar)}\int dx'\exp\left(\frac{i(p'-p'')x'}{\hbar}\right) \\
&= \delta(p''-p') \qquad (1.108)
\end{aligned}$$

These results can be generalised to three-dimensions. We have

$$\begin{aligned}
|\mathbf{r}\rangle &= |x,y,z\rangle \\
\hat{\mathbf{r}}|\mathbf{r}\rangle &= \mathbf{r}|\mathbf{r}\rangle \\
\langle\mathbf{r}'|\mathbf{r}''\rangle &= \delta(\mathbf{r}'-\mathbf{r}'') \\
|\mathbf{p}\rangle &= |p_x,p_y,p_z\rangle \\
\hat{\mathbf{p}}|\mathbf{p}\rangle &= \mathbf{p}|\mathbf{p}\rangle \\
\langle\mathbf{p}'|\mathbf{p}''\rangle &= \delta(\mathbf{p}'-\mathbf{p}'') \\
\langle\mathbf{r}'|\hat{\mathbf{p}}|\mathbf{r}''\rangle &= -i\hbar\nabla_{\mathbf{r}'}\delta(\mathbf{r}'-\mathbf{r}'') \\
\langle\mathbf{p}'|\hat{\mathbf{r}}|\mathbf{p}''\rangle &= i\hbar\nabla_{\mathbf{p}'}\delta(\mathbf{p}'-\mathbf{p}'') \\
\langle\mathbf{r}|\mathbf{p}\rangle &= \frac{1}{(2\pi\hbar)^{3/2}}\exp\Big(i\mathbf{r}\cdot\mathbf{p}/\hbar\Big) \qquad (1.109)
\end{aligned}$$

1.2 Applications

We apply the formalism developed in the previous section to some familiar (and not so familiar!) examples. Foremost is the quantum mechanical treatment of the simple harmonic oscillator (SHO). The SHO is ubiquitous in the quantum mechanical treatment of real phenomena where one is considering the vibrations of a system after a small displacement from its equilibrium position.

1.2.1 Harmonic oscillator

The Hamiltonian for a one-dimensional quantum-mechanical oscillator is

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + \frac{1}{2}m\omega^2\hat{x}^2 \tag{1.110}$$

where $\hat{x}$ and $\hat{p}$ are the one-dimensional position and momentum operators: they are linear Hermitian operators satisfying the canonical commutation relation

$$[\hat{x}, \hat{p}] = i\hbar \tag{1.111}$$

Introduce a dimensionless operator

$$\hat{a} = \frac{1}{(2m\hbar\omega)^{\frac{1}{2}}}(i\hat{p} + m\omega\hat{x}) \tag{1.112}$$

Then we have

$$\hat{a}^\dagger = \frac{1}{(2m\hbar\omega)^{\frac{1}{2}}}(-i\hat{p} + m\omega\hat{x}) \tag{1.113}$$

Note that $\hat{a}^\dagger \neq \hat{a}$ so that $\hat{a}$ is NOT Hermitian.

Problem 11: Prove the following properties of the operators $\hat{a}$ and $\hat{a}^\dagger$:

- $[\hat{a}, \hat{a}^\dagger] = \hat{I}$
- We define a dimensionless operator $\hat{N}$ where $\hat{N} = \hat{a}^\dagger\hat{a}$. Show that $\hat{N}$ is Hermitian and that it satisfies the following commutation relations: $[\hat{N}, \hat{a}] = -\hat{a}$ and $[\hat{N}, \widehat{a^\dagger}] = \hat{a}^\dagger$.

The observables $\hat{x}, \hat{p}$ and $\hat{H}$ are given in terms of $\hat{a}, \hat{a}^\dagger, \hat{N}$ by

$$\hat{x} = \left(\frac{\hbar}{2m\omega}\right)^{\frac{1}{2}}(\hat{a} + \hat{a}^\dagger) \tag{1.114}$$

$$\hat{p} = i\left(\frac{1}{2}m\hbar\omega\right)^{\frac{1}{2}}(\hat{a}^\dagger - \hat{a}) \tag{1.115}$$

$$\hat{H} = \left(\hat{a}\hat{a}^\dagger - \frac{1}{2}\right)\hbar\omega = \left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)\hbar\omega = \left(\hat{N} + \frac{1}{2}\right)\hbar\omega \tag{1.116}$$

Problem 12: Show that the above expression for $\hat{H}$ is correct.

Let $|n\rangle$ denote an eigenvector of $\hat{N}$ (we assume that such an eigenvector exists) with eigenvalue n:

$$\hat{N}|n\rangle = n|n\rangle \tag{1.117}$$

Then we find from the above expression for $\hat{H}$ that

$$\begin{aligned} \hat{H}|n\rangle &= \left(\hat{N} + \frac{1}{2}\right)\hbar\omega|n\rangle \\ &= \left(n + \frac{1}{2}\right)\hbar\omega|n\rangle \end{aligned}$$

This means that $|n\rangle$ is an energy eigenvector with eigenvalue

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \tag{1.118}$$

(Note that you would expect this since from the definition of $\hat{N}$ and $\hat{H}$ above, it is clear that $[\hat{H}, \hat{N}] = [\hat{N}, \hat{H}] = \hat{0}$ and so are compatible.) All that we know about n is that it must be a real number (Why?). To determine n, consider the effect of $\hat{a}$ and $\hat{a}^\dagger$ on $|n\rangle$.

Problem 13: Prove the following results:

$$\begin{aligned} \hat{N}(\hat{a}|n\rangle) &= (n-1)(\hat{a}|n\rangle) \\ \hat{N}(\hat{a}^\dagger|n\rangle) &= (n+1)(\hat{a}^\dagger|n\rangle) \end{aligned}$$

This indicates that $\hat{a}$ and $\hat{a}^\dagger$ act as lowering and raising operators for the quantum number n:

$$\hat{a}|n\rangle = \alpha^-|n-1\rangle \tag{1.119}$$

$$\hat{a}^\dagger|n\rangle = \alpha^+|n+1\rangle \tag{1.120}$$

where $\alpha^{\pm}$ are numbers to be determined. For this reason $\hat{a}^{\dagger}$ and $\hat{a}$ are often called *creation* and *annihilation* operators, respectively: they create or annihilate one quantum of excitation of the oscillator (not the oscillator itself!).

Problem 14: Show that for normalised kets $|n\rangle$ then $\alpha^{+} = (n+1)^{\frac{1}{2}}$ and $\alpha^{-} = n^{\frac{1}{2}}$

Therefore we have

$$\hat{a}|n\rangle = n^{\frac{1}{2}}|n-1\rangle \tag{1.121}$$

$$\hat{a}^{\dagger}|n\rangle = (n+1)^{\frac{1}{2}}|n+1\rangle \tag{1.122}$$

Problem 15: Prove that the possible values n can take are $n = 0, 1, 2, \ldots$ i.e. the non-negative integers.

The ground state, denoted by the ket $|0\rangle$ has energy

$$E_0 = \frac{1}{2}\hbar\omega \tag{1.123}$$

(i.e. the zero-point energy) and has the property

$$\hat{a}|0\rangle = 0 \tag{1.124}$$

All other eigenvectors can be determined from the ground state eigenvector by the repeated application of the raising operator $\hat{a}^{\dagger}$:

$$|n\rangle = \frac{1}{(n!)^{\frac{1}{2}}}(\hat{a}^{\dagger})^{n}|0\rangle \tag{1.125}$$

The kets $|n\rangle$ defined above are orthonormal (since non-degenerate eigenvectors of a Hermitian operator corresponding to different eigenvalues are orthogonal). They are also complete:

$$\sum_{n=0}^{\infty}|n\rangle\langle n| = \hat{I} \tag{1.126}$$

The set $\{|n\rangle\}$ is a basis for the Hilbert space i.e any state $|\psi\rangle$ can be expressed as

$$|\psi\rangle = \sum_{n=0}^{\infty}|n\rangle\langle n|\psi\rangle \tag{1.127}$$

Similarly, any operator $\hat{A}$ acting on this space can be written

$$\hat{A} = \sum_{n=0}^{\infty}\sum_{n'=0}^{\infty}|n'\rangle\langle n'|\hat{A}|n\rangle\langle n| \tag{1.128}$$

Wave functions for the harmonic oscillator

First we determine the ground state wave function using the identity

$$\hat{a}|0\rangle = 0$$

we have

$$\begin{aligned}
\langle x|\hat{a}|0\rangle &= 0 \\
\langle x|\hat{a}\Big(\int |x'\rangle\langle x'|dx'\Big)|0\rangle &= 0 \\
\int dx'\Big(\langle x|\hat{a}|x'\rangle\langle x'|0\rangle\Big) &= 0 \\
\int dx'\Big(\langle x|\hat{a}|x'\rangle\psi_0(x')\Big) &= 0 \qquad (1.129)
\end{aligned}$$

where we have defined $\psi_0(x') = \langle x'|0\rangle$. To evaluate $\langle x|\hat{a}|x'\rangle$ we use (1.112):

$$\begin{aligned}
\langle x|\hat{a}|x'\rangle &= \frac{1}{(2m\hbar\omega)^{\frac{1}{2}}}\langle x|(i\hat{p} + m\omega\hat{x})|x'\rangle \\
&= \frac{1}{(2m\hbar\omega)^{\frac{1}{2}}}\Big[i\langle x|\hat{p}|x'\rangle + m\omega\langle x|\hat{x}|x'\rangle\Big] \qquad (1.130)
\end{aligned}$$

Using properties of the operators $\hat{x}, \hat{p}$ in the position representation :

$$\begin{aligned}
\hat{x}|x'\rangle &= x'|x'\rangle \\
\langle x|\hat{x}|x'\rangle &= x'\langle x|x'\rangle \\
&= x'\delta(x - x') \\
\langle x|\hat{p}|x'\rangle &= -i\hbar\frac{\partial}{\partial x}\delta(x - x') \qquad (1.131)
\end{aligned}$$

we get

$$\langle x|\hat{a}|x'\rangle = \frac{1}{(2m\hbar\omega)^{\frac{1}{2}}}\Big[\hbar\frac{\partial}{\partial x}\delta(x - x') + m\omega x'\delta(x - x')\Big] \qquad (1.132)$$

Substituting (1.132) into (1.129) we get

$$\begin{aligned}
\int dx'\Big[\hbar\frac{\partial}{\partial x}\delta(x - x') + m\omega x'\delta(x - x')\Big]\psi_0(x') &= 0 \\
\Big(\hbar\frac{\partial}{\partial x} + m\omega x\Big)\psi_0(x) &= 0 \qquad (1.133)
\end{aligned}$$

which has the normalised solution

$$\begin{aligned}\psi_0(x) &= \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} \exp\Big(-\frac{m\omega}{2\hbar}x^2\Big) \\ \int_{-\infty}^{\infty}|\psi_0(x)|^2 &= 1 \qquad (1.134)\end{aligned}$$

To determine the excited state wave functions we proceed as follows: Since

$$\hat{a}^\dagger|n\rangle = (n+1)^{\frac{1}{2}}|n+1\rangle$$

we have

$$\langle x|\hat{a}^\dagger|n\rangle = (n+1)^{\frac{1}{2}}\langle x|n+1\rangle \qquad (1.135)$$

Defining $\psi_n(x) = \langle x|n\rangle$, (1.135) becomes

$$\begin{aligned}\langle x|\hat{a}^\dagger|n\rangle &= \int dx'\langle x|\hat{a}^\dagger|x'\rangle\psi_n(x') \\ &= (n+1)^{\frac{1}{2}}\psi_{n+1}(x) \qquad (1.136)\end{aligned}$$

We then have

$$\begin{aligned}\langle x|\hat{a}^\dagger|x'\rangle &= \frac{1}{(2m\hbar\omega)^{\frac{1}{2}}}\langle x|(-i\hat{p}+m\omega\hat{x})|x'\rangle \\ &= \frac{1}{(2m\hbar\omega)^{\frac{1}{2}}}\Big[-i\langle x|\hat{p}|x'\rangle + m\omega\langle x|\hat{x}|x'\rangle\Big] \\ &= \frac{1}{(2m\hbar\omega)^{\frac{1}{2}}}\Big[-\hbar\frac{\partial}{\partial x}\delta(x-x') + m\omega x'\delta(x-x')\Big] \qquad (1.137)\end{aligned}$$

and (1.136) reduces to

$$\left(\frac{m\omega}{2\hbar}\right)^{\frac{1}{2}}\Big(-\frac{\hbar}{m\omega}\frac{\partial}{\partial x}+x\Big)\psi_n(x) = (n+1)^{\frac{1}{2}}\psi_{n+1}(x) \qquad (1.138)$$

which gives the following recursion formula

$$\psi_{n+1}(x) = \frac{1}{(n+1)^{\frac{1}{2}}}\left(\frac{m\omega}{2\hbar}\right)^{\frac{1}{2}}\Big(-\frac{\hbar}{m\omega}\frac{\partial}{\partial x}+x\Big)\psi_n(x) \qquad (1.139)$$

Problem 16†: Work out the details leading to the formula (1.139) starting from (1.137).

Since we already know $\psi_0(x)$, (1.139) enables us to generate all the excited wave functions (although they still have to be normalised). This is a “neater” way of solving for the eigenstates of the quantum harmonic oscillator when compared with the power series method you have come across before. In fact, this is an example of the so called *factorisation method*, a very powerful method for solving second-order differential equations by purely algebraic means. (We will come across this technique again when we determine the eigenfunctions of the orbital angular momentum operator.)

The theory outlined in the above paragraph provides a complete description of the quantum-mechanical properties of a one-dimensional harmonic oscillator. For example, matrix elements of certain functions of the position and momentum operators can easily be evaluated.

Problem 17†: Prove the following results:

- $$\langle n'|\hat{x}|n\rangle = \left(\frac{\hbar}{2m\omega}\right)^{\frac{1}{2}}\left[n^{\frac{1}{2}}\delta_{n',n-1} + (n+1)^{\frac{1}{2}}\delta_{n',n+1}\right] \tag{1.140}$$
- $$\langle n'|\hat{p}|n\rangle = i\left(\frac{1}{2}m\hbar\omega\right)^{\frac{1}{2}}\left[-n^{\frac{1}{2}}\delta_{n',n-1} + (n+1)^{\frac{1}{2}}\delta_{n',n+1}\right] \tag{1.141}$$
- $$\langle n'|\widehat{x^2}|n\rangle = \frac{\hbar}{2m\omega}\left\{\left[n(n-1)\right]^{\frac{1}{2}}\delta_{n',n-2} + (2n+1)\delta_{n,n'} + \left[(n+1)(n+2)\right]^{\frac{1}{2}}\delta_{n',n+2}\right\} \tag{1.142}$$
- $$\langle n'|\widehat{p^2}|n\rangle = \frac{1}{2}m\hbar\omega\left\{-\left[n(n-1)\right]^{\frac{1}{2}}\delta_{n',n-2} + (2n+1)\delta_{n,n'} - \left[(n+1)(n+2)\right]^{\frac{1}{2}}\delta_{n',n+2}\right\} \tag{1.143}$$

Problem 18:

- Show that if $f(\hat{a}^\dagger)$ is any polynomial in $\hat{a}^\dagger$, then
$$\hat{a}f(\hat{a}^\dagger)|0\rangle = \frac{df(\hat{a}^\dagger)}{d\hat{a}^\dagger}|0\rangle \tag{1.144}$$
- Show that:
$$e^{\lambda\hat{a}}f(\hat{a}^\dagger)|0\rangle = f(\hat{a}^\dagger + \lambda)|0\rangle \tag{1.145}$$

- Show that:

$$e^{\lambda \hat{a}} f(\hat{a}^\dagger) e^{-\lambda \hat{a}} = f(\hat{a}^\dagger + \lambda) \tag{1.146}$$

- Use the preceding relation to prove that

$$e^{\alpha \hat{a} + \beta \hat{a}^\dagger} = e^{\alpha \hat{a}} e^{\beta \hat{a}^\dagger} e^{-\frac{1}{2}\alpha\beta} \tag{1.147}$$

- Show that:

$$e^{\lambda \hat{a}^\dagger} f(\hat{a}) e^{-\lambda \hat{a}^\dagger} = f(\hat{a} - \lambda) \tag{1.148}$$

  From this, show that

$$e^{\alpha \hat{a} + \beta \hat{a}^\dagger} = e^{\beta \hat{a}^\dagger} e^{\alpha \hat{a}} e^{\frac{1}{2}\alpha\beta} \tag{1.149}$$

- A state $|\alpha\rangle$ that obeys the equation

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle \tag{1.150}$$

  is called a *coherent state*. Show that the state $|\alpha\rangle$ may be written as

$$|\alpha\rangle = C e^{\alpha \hat{a}^\dagger}|0\rangle \tag{1.151}$$

  and determine the value of C. By expanding the state $|\alpha\rangle$ in the basis $\{|n\rangle\}$ find the probability that a coherent state contains n quanta. This probability distribution is called a *Poisson distribution*. Finally calculate the value of $\langle\alpha|\hat{N}|\alpha\rangle$ which corresponds to the average number of quanta in the coherent state. (This example is relevant in quantum optics.)

1.2.2 Delta-function potential well

As an example of a system with a mixed (discrete-continuous) spectrum, consider a finite potential well of width a and depth V_0:

$$\begin{aligned} V(x) &= -V_0 \quad \text{for } |x| < \tfrac{1}{2}a \\ V(x) &= 0 \quad \text{elsewhere} \end{aligned} \tag{1.152}$$

In the limit that the well becomes very deep and narrow, such that $V_0 \to \infty$ and $a \to 0$ while $aV_0 \equiv \mathcal{V}$ remains fixed, we may approximate the potential by a Dirac delta function:

$$V(x) = -\mathcal{V}\,\delta(x) \tag{1.153}$$

(This will also give us some practice at handling the delta function.)

Negative-energy eigenstates of this system correspond to bound states, which will be normalisable and have a discrete spectrum. The wave function must fall off exponentially outside the well and hence must take the form

$$\psi(x) = Ae^{-\kappa|x|} \tag{1.154}$$

where

$$\kappa = \frac{\sqrt{-2mE}}{\hbar} \tag{1.155}$$

and (for normalisation) $A = \sqrt{\kappa}$. Integrating the Schrödinger equation

$$E\psi = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} - \mathcal{V}\,\delta(x)\psi \tag{1.156}$$

between the limits $x = -\epsilon$ and $x = +\epsilon$ gives

$$E\int_{-\epsilon}^{+\epsilon}\psi\,dx = -\frac{\hbar^2}{2m}\left[\left(\frac{\partial\psi}{\partial x}\right)_{x=+\epsilon} - \left(\frac{\partial\psi}{\partial x}\right)_{x=-\epsilon}\right] - \mathcal{V}\psi(0) \tag{1.157}$$

Now taking the limit $\epsilon \to 0$ will make the integral on the left-hand side vanish, since ψ must be finite and continuous at $x = 0$. Therefore ψ must have a discontinuity in its derivative at the origin, such that

$$\lim_{\epsilon\to 0}\left[\left(\frac{\partial\psi}{\partial x}\right)_{x=+\epsilon} - \left(\frac{\partial\psi}{\partial x}\right)_{x=-\epsilon}\right] = -\frac{2m\mathcal{V}}{\hbar^2}\psi(0) \tag{1.158}$$

Inserting the form (1.154) for the solution, we find that

$$\kappa = \frac{m\mathcal{V}}{\hbar^2} \tag{1.159}$$

and hence

$$E = -\frac{1}{2}m\left(\frac{\mathcal{V}}{\hbar}\right)^2 \tag{1.160}$$

Thus for $E < 0$ there is a unique solution for E, and hence a single bound state.

For $E > 0$, on the other hand, we can obtain plane-wave solutions with wave number $k = \sqrt{2mE}/\hbar$ for any value of E. Since the potential is an even function of x, we can classify the eigenstates according to parity. Those with odd parity must vanish at the origin and then equation (1.158) tells us there is no change in the derivative at the origin, just as if the potential well were not there. Thus the odd-parity eigenfunctions are simply of the form

$$\psi(x) = C\sin kx \tag{1.161}$$

and any odd-parity wave function of the system can be Fourier decomposed into sine waves, just as for a free particle. For the usual delta function normalisation $\langle k|k'\rangle = \delta(k - k')$, we require $C = 1/\sqrt{\pi}$.

The even-parity eigenstates, on the other hand, need not vanish at the origin, and hence they feel the presence of the potential well. For $E > 0$ we can write them in the general form

$$\psi(x) = C\cos(k|x| + \phi) \tag{1.162}$$

where, from equation (1.158), the phase shift ϕ satisfies the equation

$$\tan\phi = \frac{m\mathcal{V}}{\hbar^2 k} = \frac{\kappa}{k} \tag{1.163}$$

The presence of the phase shift guarantees that the positive-energy even-parity eigenstates (1.162) are orthogonal to the bound state (1.154). To see this, consider the overlap integral

$$\begin{aligned}\int_{-\infty}^{\infty}\cos(k|x| + \phi)e^{-\kappa|x|}dx &= 2\int_0^{\infty}\cos(kx + \phi)e^{-\kappa x}dx \\ &= \int_0^{\infty}\left(e^{ikx+i\phi-\kappa x} + e^{-ikx-i\phi-\kappa x}\right)dx \\ &= \frac{e^{i\phi}}{\kappa - ik} + \frac{e^{-i\phi}}{\kappa + ik} \\ &= \frac{2}{\kappa^2 + k^2}(\kappa\cos\phi - k\sin\phi) \\ &= 0 \end{aligned} \tag{1.164}$$

on account of equation (1.163). Hence any unbound state of the system can be Fourier decomposed into sine waves and the "kinked cosine" waves (1.162). Conversely, the square modulus of the overlap between any normalised state and the bound state (1.154) gives the probability that the particle is bound in the potential well.

Problem 19: Find the probability P_b that a particle with wave function

$$\psi(x) = \sqrt{c}\, e^{-c|x|} \tag{1.165}$$

will be found in the bound state of the delta-function potential well. Confirm that $P_b \leq 1$. Find also the probability $P(k)dk$ that the particle will be found to be unbound, with wave number between k and $k + dk$.

Answer:

$$P_b = \frac{4c\kappa}{(c+\kappa)^2}\,, \qquad P(k) = \frac{4ck^2(c-\kappa)^2}{\pi(\kappa^2 + k^2)(c^2 + k^2)^2}$$

Problem 20†: Confirm that

$$\int_0^{\infty} P(k)\,dk = 1 - P_b \tag{1.166}$$

1.3 Introduction to second quantisation

Historically, the development of quantum mechanics focused on the physics of entities which in the classical limit are particles. To these particles was attributed a wave nature (with Schrödinger's equation being the wave equation), which shows up in experiments like electron diffraction for example. This approach may be called *first quantisation.* This wave-particle duality forms the foundation on which quantum mechanics was built. Similarly, it is only natural to expect that waves in classical physics such as electromagnetic radiation could have particle-like attributes, a conjecture spectacularly confirmed by the Compton effect and Einstein's explanation of the photoelectric effect. In the case of the electromagnetic field, it is appropriate to define it not in terms of its field amplitudes (**E** or **B**) but by how many "photons" or radiation quanta are present in a given mode, and this is especially important when one tries to understand the interaction of EM radiation with matter, where the number of photons can vary. The procedure whereby one constructs a particle description of classical waves is known as *second quantisation.* Note however that this procedure is not limited to classical waves. It can be applied to the Schrödinger equation (a wave equation for the probability amplitude of an electron for example). In the second quantisation procedure, wave fields are "quantised" to describe the problem in terms of "quanta" in a given mode of vibration. Examples include photons (quanta of EM field), phonons (quanta of vibrations in solids), plasmons (quanta of collective electron vibrations) etc. The physical states of this system are characterised by *particle occupation numbers*, i.e. the number of quanta in a particular mode. Below, we will use the vibrations of a stretched string (covered in IB Physics!) to bring out the essence of the second quantisation approach.

1.3.1 Vibrating string

We first review the classical treatment of the vibrating string. Consider a stretched string of uniform linear density ρ and length L with fixed ends. The transverse displacement as a function of distance x along the string is denoted by $y(x,t)$. The equation of the vibrating string is given by

$$\frac{\partial^2 y}{\partial t^2} = v^2 \frac{\partial^2 y}{\partial x^2} \tag{1.167}$$

with the boundary conditions $y(0,t) = y(L,t) = 0$ for all t. Note that $v = \sqrt{\frac{\tau}{\rho}}$ where τ denotes the tension in the string. We now use the method of separation of variables to solve equation (1.167) subject to the stated boundary conditions. Let

$$y(x,t) = X(x)\ \psi(t) \tag{1.168}$$

Substituting this into (1.167) and dividing throughout by $X(x)\psi(t)$ gives

$$v^2 \frac{1}{X}\frac{d^2X}{dx^2} = \frac{1}{\psi}\frac{d^2\psi}{dt^2} \tag{1.169}$$

As the left-hand side depends only x and the right-hand side depends only on t, each must be a constant: let it be $-k^2$. We then have the ordinary differential equations

$$\begin{aligned} \frac{d^2X}{dx^2} + k^2X &= 0 \\ \frac{d^2\psi}{dt^2} + v^2k^2\psi &= 0 \end{aligned} \tag{1.170}$$

whose boundary conditions become $X(0) = X(L) = 0$ and corresponding solutions are respectively

$$\begin{aligned} X(x) &= c_1\cos(kx) + c_2\sin(kx) \\ \psi(t) &= b_1\cos(vkt) + b_2\sin(vkt) \end{aligned} \tag{1.171}$$

Inserting the boundary condition gives

$$c_1 = 0, \qquad c_2\sin(kL) = 0$$

As c_1 and c_2 cannot both be zero (otherwise $y(x,t)$ would be identically zero), we have to choose $\sin(kL) = 0$ or

$$kL = n\pi, \qquad n = 1, 2, 3\cdots$$

Thus the allowed general solutions are

$$\begin{aligned} y(x,t) &= \sum_{n=1}^{\infty}\psi_n(t)\sin\left(\frac{n\pi x}{L}\right) \\ \psi_n(t) &= \left[A_n\cos\left(\frac{n\pi vt}{L}\right) + B_n\sin\left(\frac{n\pi vt}{L}\right)\right]. \end{aligned} \tag{1.172}$$

(We have replaced the integration constants b_1c_2 by A_n etc.) Each term in the above general solution is an allowed solution corresponding to an allowed *mode* of vibration. The Fourier expansion coefficients $\psi_n(t)$ are called the *normal coordinates* and for an arbitrary $y(x,t)$ can be determined by

$$\psi_n(t) = \frac{2}{L}\int_0^L y(x,t)\,\sin\left(\frac{n\pi x}{L}\right)\,dx \tag{1.173}$$

A knowledge of all the normal coordinates completely determines $y(x,t)$. Equation (1.172) represents the superposition of independent modes of vibration i.e. the *normal modes*, each with its characteristic frequency given by $\omega_n = n\pi v/L$. An important property of normal coordinates is that the kinetic and potential energies are expressible as uncoupled sums of the normal coordinates and their first derivatives. In the case of the kinetic energy, we have

$$\begin{aligned} T &= \int_0^L \frac{1}{2}\rho \left(\frac{\partial y}{\partial t}\right)^2 dx \\ \frac{\partial y}{\partial t} &= \sum_{n=1}^{\infty} \dot{\psi}_n \sin\left(\frac{n\pi x}{L}\right) \end{aligned} \tag{1.174}$$

which implies

$$\begin{aligned} T &= \frac{1}{2}\rho \int_0^L \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \dot{\psi}_n \, \dot{\psi}_m \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi x}{L}\right) dx \\ &= \sum_{n=1}^{\infty} \frac{1}{4}\rho L \, \dot{\psi}_n^2. \end{aligned} \tag{1.175}$$

Similarly, for the potential energy, we have

$$\begin{aligned} V &= \int_0^L \frac{1}{2}\tau \left(\frac{\partial y}{\partial x}\right)^2 dx \\ &= \sum_{n=1}^{\infty} \frac{1}{4}\frac{\tau}{L} n^2 \pi^2 \psi_n^2 \end{aligned} \tag{1.176}$$

To derive the equation of motion for the normal coordinates, we construct the Lagrangian $\mathcal{L}$ where

$$\mathcal{L} = T - V = \sum_{n=1}^{\infty} \left[\frac{1}{4}\rho L \dot{\psi}_n^2 - \frac{1}{4}\frac{\tau}{L} n^2 \pi^2 \psi_n^2\right] \tag{1.177}$$

and using Lagrange's equation

$$\frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\psi}_n}\right) - \frac{\partial \mathcal{L}}{\partial \psi_n} = 0 \tag{1.178}$$

we get

$$\ddot{\psi}_n + \frac{n^2 \pi^2 \tau}{L^2 \rho} \, \psi_n = 0 \tag{1.179}$$

This is just the equation for a simple harmonic oscillator!

We can equally well describe the vibrating string problem in the Hamiltonian formalism. Recall that the *generalised momentum* corresponding to the *generalised coordinate* ψ_n is defined by

$$\Pi_n = \frac{\partial \mathcal{L}}{\partial \dot{\psi}_n} = \frac{1}{2}\rho L \dot{\psi}_n \tag{1.180}$$

The Hamiltonian function, H (where $H(p,q) = \sum pq - \mathcal{L}$) is given by

$$\begin{aligned} H &= \sum_{n=1}^{\infty} \Pi_n \, \dot{\psi}_n - \mathcal{L} \\ &= \sum_{n=1}^{\infty} \left[\frac{\Pi_n^2}{\rho L} + \frac{n^2 \pi^2 \tau}{4L} \psi_n^2\right] \end{aligned} \tag{1.181}$$

which is an infinite sum of harmonic oscillator Hamiltonian functions. The above equation serves as the starting point for *quantizing* the vibrations of the string.

1.3.2 Quantisation of vibrating string

To **quantize** a system, we make the transition from dynamical variables to operators. Define the quantum mechanical Hamiltonian operator by making the replacement

$$\Pi_n \rightarrow -i\hbar \frac{\partial}{\partial \psi_n} \tag{1.182}$$

Equation (1.181) then becomes

$$\begin{aligned} \hat{H} &= \sum_{n=1}^{\infty} \left[-\frac{\hbar^2}{\rho L} \frac{\partial^2}{\partial \psi_n^2} + \frac{\tau}{4L} n^2 \pi^2 \psi_n^2 \right] \\ &= \sum_{n=1}^{\infty} \hat{H}_n \end{aligned} \tag{1.183}$$

i.e. the Hamiltonian reduces to a sum of terms, each of which is dependent on a single variable (ψ_n). Therefore the total wave function for this system is given by

$$\Psi(\psi_n, t) = \prod_{n=1}^{\infty} \Phi_{N_n}(\psi_n, t) \tag{1.184}$$

where Φ_{N_n} are the eigenfunctions of the SHO problem satisfying

$$\left[-\frac{\hbar^2}{\rho L} \frac{\partial^2}{\partial \psi_n^2} + \frac{\tau}{4L} n^2 \pi^2 \psi_n^2 \right] \Phi_{N_n}(\psi_n, t) = i\hbar \frac{\partial}{\partial t} \Phi_{N_n}(\psi_n, t) \tag{1.185}$$

The solutions of the above equation have the generic form

$$\Phi_{N_n}(\psi_n, t) = U_{N_n}(\psi_n)\, e^{-iE_n(N_n)t/\hbar} \tag{1.186}$$

where the U_{N_n} satisfy

$$-\frac{\hbar^2}{\rho L} \frac{d^2 U_{N_n}}{d\psi_n^2} + \frac{\tau}{4L} n^2 \pi^2 \psi_n^2\, U_{N_n} = E(N_n) U_{N_n} \tag{1.187}$$

This is just the differential equation for the SHO with the following correspondence

$$\begin{aligned} \text{“mass” } m &\rightarrow \frac{1}{2}\rho L \\ \text{“force constant” } k &\rightarrow \frac{\tau}{2L} n^2 \pi^2 \\ \text{“frequency” } \omega_n = \sqrt{\frac{k}{m}} &= \frac{n\pi}{L} \sqrt{\frac{\tau}{\rho}} \end{aligned} \tag{1.188}$$

The solutions U_{N_n} are the Hermite polynomials discussed earlier in the chapter. So we can draw the following conclusions:

- The energy in any given normal mode of frequency ω_n is
$$\begin{aligned} E_n(N_n) &= \left(N_n + \frac{1}{2}\right)\hbar\omega_n \\ &= \left(N_n + \frac{1}{2}\right)\frac{n\pi\hbar}{L}\sqrt{\frac{\tau}{\rho}} \qquad (1.189) \end{aligned}$$
The number $N_n = 0, 1, 2, \ldots$ is the *occupation number* of the mode.

- We can define annihilation and creation operators for each mode as before:
$$\begin{aligned} \hat{a}_n &= \frac{1}{(\rho L\hbar\omega_n)^{\frac{1}{2}}}\left(\frac{1}{2}\rho L\omega_n\hat{\psi}_n + i\hat{\Pi}_n\right) \\ \hat{a}_n^\dagger &= \frac{1}{(\rho L\hbar\omega_n)^{\frac{1}{2}}}\left(\frac{1}{2}\rho L\omega_n\hat{\psi}_n - i\hat{\Pi}_n\right) \\ \hat{\psi}_n &= \left(\frac{\hbar}{\rho L\omega_n}\right)^{\frac{1}{2}}\left(\hat{a}_n + \hat{a}_n^\dagger\right) \\ \hat{\Pi}_n &= \frac{1}{2}i\left(\rho L\hbar\omega_n\right)^{\frac{1}{2}}\left(\hat{a}_n^\dagger - \hat{a}_n\right) \qquad (1.190) \end{aligned}$$
so that the number operator that counts the number of quanta in a particular mode is given by
$$\hat{N}_n = \hat{a}_n^\dagger\hat{a}_n \qquad (1.191)$$

- The state $|N_n\rangle$ indicates that the system has N_n quanta of excitation of the n-th mode, each with energy $\hbar\omega_n$. The effect of the creation and annihilation operators on a state $|N_n\rangle$ is:
$$\begin{aligned} \hat{a}_n^\dagger|N_n\rangle &= \sqrt{(N_n+1)}|N_n+1\rangle \\ \hat{a}_n|N_n\rangle &= \sqrt{N_n}|N_n-1\rangle \qquad (1.192) \end{aligned}$$

- In each mode, the lowest energy state does not have zero energy but has energy $\frac{1}{2}\hbar\omega_n$. This represents the "vacuum" or ground state of the mode.

1.3.3 General second quantisation procedure

The procedure for quantising the vibrating string described in the previous section can in principle be applied to any wave equation. The end result is that one is able to "second-quantise" the wave or field amplitude. Instead of describing the allowed modes in terms of the amplitude, we now describe them by the "quanta" or "particle" number in the mode. This procedure can be summarised as follows:

1. Given a field or wave equation, represent it in normal coordinates.
2. Find the Lagrangian and Hamiltonian functions corresponding to the field equations with appropriately defined generalised coordinates and momenta.
3. Treat the normal coordinate Hamiltonian obtained above as a quantum mechanical harmonic oscillator so that
    - States are defined by the number of quanta (particles) in each allowed mode
    - Creation and annihilation operators as in the SHO case allow particles/quanta to be added to or removed from the physical state
    - A vacuum state with a fixed energy represents the lowest energy state.

We will extend the concept of second quantisation to interacting many-particle systems, in the process second-quantising the Schrödinger equation, in Chapter 5. We will see that the second quantisation procedure makes the treatment of interacting many-particle systems more manageable.

Chapter 2

Angular Momentum

2.1 Introduction

As in classical mechanics, the angular momentum (AM) of an isolated quantum mechanical system is conserved, and this provides a useful general classification of the states of the system in terms of their angular momenta. However, the mathematics of AM is a good deal more complicated in QM, because the values of the AM are quantised and because its different components are not compatible observables.

The conservation of AM is a consequence of the fact that the properties of an isolated physical system are invariant under rotations of the system in space (or equivalently, under reorientation of a Cartesian coordinate system used to describe the system). This is an example of a general connection between invariances and conservation laws, which exists in classical mechanics but is especially clear in QM.

In the following sections we develop the mathematical treatment of AM and explain its connection with rotational invariance.

2.2 Orbital angular momentum

In classical mechanics, the AM of a particle (with respect to the origin) is given by the formula:

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \tag{2.1}$$

so that,

$$L_x = yp_z - zp_y \tag{2.2}$$

$$L_y = zp_x - xp_z \tag{2.3}$$

$$L_z = xp_y - yp_x \tag{2.4}$$

where $\mathbf{r} = (x, y, z)$ and $\mathbf{p} = (p_x, p_y, p_z)$ are the position and linear momentum vectors, respectively.

The mathematics of AM comes about through the quantum rule of replacing the linear momentum $\mathbf{p}$ of a classical point particle, located at point $\mathbf{r}$, by $\hat{\mathbf{p}} = -i\hbar\nabla$, thus replacing the classical angular momentum of a point particle (about the origin of some chosen Cartesian coordinate system) by the AM operator:

$$\hat{\mathbf{L}} = -i\hbar\mathbf{r} \times \nabla \tag{2.5}$$

This implies:

$$\begin{aligned}
\hat{L}_x &= y\hat{p}_z - z\hat{p}_y \\
&= \frac{\hbar}{i}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) \\
\hat{L}_y &= z\hat{p}_x - x\hat{p}_z \\
&= \frac{\hbar}{i}\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) \\
\hat{L}_z &= x\hat{p}_y - y\hat{p}_x \\
&= \frac{\hbar}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)
\end{aligned} \tag{2.6}$$

Equation (2.6) can be written more compactly as

$$\begin{aligned}
\hat{L}_i &= \epsilon_{ijk}\hat{r}_j\hat{p}_k \\
i, j, k &= 1, 2, 3
\end{aligned} \tag{2.7}$$

where $1 \rightarrow x-$component, $2 \rightarrow y-$component, $3 \rightarrow z-$component and we have used the Einstein summation convention (i.e. the expression is summed over repeated indices). ϵ_{ijk} is the completely antisymmetric tensor of the third rank (also called the *Levi-Civita tensor*)

$$\epsilon_{ijk} = \begin{cases} +1 & \text{for even permutations of (123)} \\ -1 & \text{for odd permutations of (123)} \\ 0 & \text{otherwise} \end{cases}$$

Some useful identities that the ϵ_{ijk} satisfy are

$$\epsilon_{ijk}\,\epsilon_{abc} = \begin{vmatrix} \delta_{ia} & \delta_{ib} & \delta_{ic} \\ \delta_{ja} & \delta_{jb} & \delta_{jc} \\ \delta_{ka} & \delta_{kb} & \delta_{kc} \end{vmatrix} \tag{2.8}$$

The antisymmetric nature of the ϵ_{ijk} means that

$$\epsilon_{ijk} = -\epsilon_{jik} = -\epsilon_{kji}$$

The product of two of these tensors, when summed over a pair of indices (i.e. the indices are *contracted*), has the following useful properties

$$\epsilon_{ijk}\,\epsilon_{ibc} = \delta_{jb}\delta_{kc} - \delta_{kb}\delta_{jc} \tag{2.9}$$

$$\epsilon_{ijk}\,\epsilon_{ijc} = 2\delta_{kc} \tag{2.10}$$

$$\epsilon_{ijk}\,\epsilon_{ijk} = 6 \tag{2.11}$$

which can be proved by expanding out the terms. We will use these results later to prove some AM operator identities, but familiarity with them will be useful in your future study of special and general relativity, classical and quantum electrodynamics, and quantum field theory.

The quantal AM properties of a simple one-particle system are then to be inferred from the properties of these operators and their actions in the associated Hilbert space. In later sections, we will deduce the eigenvalues and eigenfunctions of these operators. The components of the quantum mechanical AM operator satisfy the following commutation relations:

$$\begin{aligned} [\hat{L}_x, \hat{L}_y] &= i\hbar\hat{L}_z \\ [\hat{L}_y, \hat{L}_z] &= i\hbar\hat{L}_x \\ [\hat{L}_z, \hat{L}_x] &= i\hbar\hat{L}_y \end{aligned} \tag{2.12}$$

or

$$[\hat{L}_i, \hat{L}_j] = i\hbar\epsilon_{ijk}\hat{L}_k \tag{2.13}$$

Further identities include

$$[\hat{L}_i, \hat{r}_j] = i\hbar\epsilon_{ijk}\hat{r}_k \tag{2.14}$$

$$[\hat{L}_i, \hat{p}_j] = i\hbar\epsilon_{ijk}\hat{p}_k \tag{2.15}$$

Problem 1: Verify (2.13)-(2.15) by direct calculation.

Partial solution: Here is the proof for (2.13):

$$[\hat{L}_i, \hat{L}_j] = \epsilon_{i\alpha\beta}\epsilon_{j\mu\nu}[\hat{r}_\alpha\hat{p}_\beta, \hat{r}_\mu\hat{p}_\nu] \tag{2.16}$$

We use the following operator identity to expand out the right hand side of (2.16)

$$\begin{aligned}
[\hat{A}\hat{B}, \hat{C}\hat{D}] &= \hat{A}[\hat{B}, \hat{C}\hat{D}] + [\hat{A}, \hat{C}\hat{D}]\hat{B} \\
&= \hat{A}\Big([\hat{B}, \hat{C}]\hat{D} + \hat{C}[\hat{B}, \hat{D}]\Big) + \Big([\hat{A}, \hat{C}]\hat{D} + \hat{C}[\hat{A}, \hat{D}]\Big)\hat{B} \\
&= \hat{A}[\hat{B}, \hat{C}]\hat{D} + \hat{A}\hat{C}[\hat{B}, \hat{D}] + [\hat{A}, \hat{C}]\hat{D}\hat{B} + \hat{C}[\hat{A}, \hat{D}]\hat{B} \qquad (2.17)
\end{aligned}$$

Making the substitutions

$$\begin{aligned}
\hat{A} &\rightarrow \hat{r}_\alpha \\
\hat{B} &\rightarrow \hat{p}_\beta \\
\hat{C} &\rightarrow \hat{r}_\mu \\
\hat{D} &\rightarrow \hat{p}_\nu \qquad (2.18)
\end{aligned}$$

into (2.17) we have

$$\begin{aligned}
[\hat{L}_i, \hat{L}_j] &= \epsilon_{i\alpha\beta}\epsilon_{j\mu\nu}[\hat{r}_\alpha\hat{p}_\beta, \hat{r}_\mu\hat{p}_\nu] \\
&= \epsilon_{i\alpha\beta}\epsilon_{j\mu\nu}\Big\{\hat{r}_\alpha[\hat{p}_\beta, \hat{r}_\mu]\hat{p}_\nu + \hat{r}_\alpha\hat{r}_\mu[\hat{p}_\beta, \hat{p}_\nu] + [\hat{r}_\alpha, \hat{r}_\mu]\hat{p}_\nu\hat{p}_\beta + \hat{r}_\mu[\hat{r}_\alpha, \hat{p}_\nu]\hat{p}_\beta\Big\} \qquad (2.19)
\end{aligned}$$

We now make use of the following operator identities

$$[\hat{r}_i, \hat{r}_j] = 0 \tag{2.20}$$

$$[\hat{p}_i, \hat{p}_j] = 0 \tag{2.21}$$

$$[\hat{r}_i, \hat{p}_j] = i\hbar\delta_{ij} \tag{2.22}$$

so that (2.16) reduces to

$$\begin{aligned}
[\hat{L}_i, \hat{L}_j] &= \epsilon_{i\alpha\beta}\epsilon_{j\mu\nu}\Big\{-i\hbar\hat{r}_\alpha\hat{p}_\nu\delta_{\beta\mu} + i\hbar\hat{r}_\mu\hat{p}_\beta\delta_{\alpha\nu}\Big\} \\
&= -i\hbar\epsilon_{i\alpha\mu}\epsilon_{j\mu\nu}\hat{r}_\alpha\hat{p}_\nu + i\hbar\epsilon_{i\nu\beta}\epsilon_{j\mu\nu}\hat{r}_\mu\hat{p}_\beta \qquad (2.23)
\end{aligned}$$

Since $\epsilon_{j\mu\nu} = -\epsilon_{j\nu\mu}$ etc. we find that

$$\begin{aligned}\epsilon_{i\alpha\mu}\epsilon_{j\mu\nu} &= -\epsilon_{i\alpha\mu}\epsilon_{j\nu\mu} \\ &= \delta_{i\nu}\delta_{\alpha j} - \delta_{ij}\delta_{\alpha\nu} \\ \epsilon_{i\nu\beta}\epsilon_{j\mu\nu} &= \delta_{i\mu}\delta_{\beta j} - \delta_{ij}\delta_{\beta\nu} \qquad (2.24)\end{aligned}$$

Substituting (2.19) into (2.18) we have

$$\begin{aligned}[\hat{L}_i, \hat{L}_j] &= -i\hbar\Big[\delta_{i\nu}\delta_{\alpha j} - \delta_{ij}\delta_{\alpha\nu}\Big]\hat{r}_\alpha\hat{p}_\nu + i\hbar\Big[\delta_{i\mu}\delta_{\beta j} - \delta_{ij}\delta_{\beta\nu}\Big]\hat{r}_\mu\hat{p}_\beta \\ &= -i\hbar\Big[\delta_{i\nu}\delta_{\alpha j} - \delta_{ij}\delta_{\alpha\nu}\Big]\hat{r}_\alpha\hat{p}_\nu + i\hbar\Big[\delta_{i\alpha}\delta_{\nu j} - \delta_{ij}\delta_{\nu\alpha}\Big]\hat{r}_\alpha\hat{p}_\nu \qquad (2.25)\end{aligned}$$

Note that in the last term of the above equation we have relabelled the indices $\mu \to \alpha$ and $\beta \to \nu$ since these are dummy indices (remember, due to the summation convention, summing over repeated indices α is the same as summing over repeated indices ν). Equation (2.20) reduces to

$$\begin{aligned}[\hat{L}_i, \hat{L}_j] &= i\hbar\Big[\delta_{i\alpha}\delta_{\nu j} - \delta_{i\nu}\delta_{\alpha j}\Big]\hat{r}_\alpha\hat{p}_\nu \\ &= i\hbar\epsilon_{kij}\epsilon_{k\alpha\nu}\hat{r}_\alpha\hat{p}_\nu \qquad (2.26)\end{aligned}$$

Since $\epsilon_{kij} = -\epsilon_{ikj} = \epsilon_{ijk}$ and $\hat{L}_k = \epsilon_{k\alpha\nu}\hat{r}_\alpha\hat{p}_\nu$ we finally have the result we want to prove i.e.

$$[\hat{L}_i, \hat{L}_j] = i\hbar\epsilon_{ijk}\hat{L}_k \qquad (2.27)$$

Actually we have done the hardest part. The proof for the other identities is much easier!

From these fundamental commutation relations, the entire theory of AM can be deduced. Evidently $\hat{L}_x, \hat{L}_y, \hat{L}_z$ are incompatible observables so it will be futile to look for states that are simultaneous eigenfunctions of $\hat{L}_x$ and $\hat{L}_y$. From the generalised uncertainty principle we deduce that:

$$\Delta L_x \Delta L_y \geq \frac{\hbar}{2}|\langle\hat{L_z}\rangle| \qquad (2.28)$$

On the other hand, the square of the total angular momentum, $\hat{L}^2$,

$$\hat{L}^2 \equiv \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 \qquad (2.29)$$

does commute with $\hat{L}_x, \hat{L}_y$ and $\hat{L}_z$.

Problem 2: Show that:

$$[\hat{L}^2, \hat{L}_x] = 0 \tag{2.30}$$

$$[\hat{L}^2, \hat{L}_y] = 0 \tag{2.31}$$

$$[\hat{L}^2, \hat{L}_z] = 0 \tag{2.32}$$

or more compactly,

$$[\hat{L^2}, \hat{\mathbf{L}}] = 0 \tag{2.33}$$

Problem 3:

- Show that $\hat{L}^2 = \hat{\mathbf{r}}^2\hat{\mathbf{p}}^2 - (\hat{\mathbf{r}} \cdot \hat{\mathbf{p}})^2 + i\hbar(\hat{\mathbf{r}} \cdot \hat{\mathbf{p}})$

  Solution: We use the tensor method.

$$\begin{aligned}
\hat{L}^2 &= \epsilon_{ijk}\hat{r}_j\hat{p}_k\epsilon_{i\alpha\beta}\hat{r}_\alpha\hat{p}_\beta \\
&= (\delta_{j\alpha}\delta_{k\beta} - \delta_{k\alpha}\delta_{j\beta})\hat{r}_j\hat{p}_k\hat{r}_\alpha\hat{p}_\beta \\
&= \hat{r}_\alpha\hat{p}_\beta\hat{r}_\alpha\hat{p}_\beta - \hat{r}_\beta\hat{p}_\alpha\hat{r}_\alpha\hat{p}_\beta \\
&= \hat{r}_\alpha\hat{r}_\alpha\hat{p}_\beta\hat{p}_\beta - i\hbar\delta_{\alpha\beta}\hat{r}_\alpha\hat{p}_\beta - \hat{r}_\beta(\hat{r}_\alpha\hat{p}_\alpha - 3i\hbar\hat{I})\hat{p}_\beta \\
&= (\hat{r} \cdot \hat{r})(\hat{p} \cdot \hat{p}) - i\hbar(\hat{r} \cdot \hat{p}) - \hat{r}_\beta\hat{r}_\alpha\hat{p}_\alpha\hat{p}_\beta + 3i\hbar(\hat{r} \cdot \hat{p}) \\
&= (\hat{r} \cdot \hat{r})(\hat{p} \cdot \hat{p}) - i\hbar(\hat{r} \cdot \hat{p}) - \hat{r}_\beta\hat{r}_\alpha\hat{p}_\beta\hat{p}_\alpha + 3i\hbar(\hat{r} \cdot \hat{p})
\end{aligned} \tag{2.34}$$

  Using

$$\begin{aligned}
\hat{r}_\beta\hat{r}_\alpha\hat{p}_\beta\hat{p}_\alpha &= \hat{r}_\beta(\hat{p}_\beta\hat{r}_\alpha + i\hbar\delta_{\alpha\beta})\hat{p}_\alpha \\
&= \hat{r}_\beta\hat{p}_\beta\hat{r}_\alpha\hat{p}_\alpha + i\hbar\hat{r}_\alpha\hat{p}_\alpha \\
&= (\hat{r} \cdot \hat{p})(\hat{r} \cdot \hat{p}) + i\hbar(\hat{r} \cdot \hat{p})
\end{aligned} \tag{2.35}$$

  Putting this into the above equation, we get the desired result. Note that we have used

$$\begin{aligned}
\epsilon_{ijk}\epsilon_{i\alpha\beta} &= \delta_{j\alpha}\delta_{k\beta} - \delta_{k\alpha}\delta_{j\beta} \\
\hat{p}_\alpha\hat{r}_\alpha &= \hat{r}_\alpha\hat{p}_\alpha - 3i\hbar \\
\hat{p}_k\hat{r}_j &= \hat{r}_j\hat{p}_k - i\hbar\delta_{jk}
\end{aligned} \tag{2.36}$$

- Show that $[\hat{L}^2, \hat{\mathbf{r}}] = -2i\hbar\hat{\mathbf{L}} \times \hat{\mathbf{r}} - 2\hbar^2\hat{\mathbf{r}}$
- Show that $[\hat{L}^2, \hat{\mathbf{p}}] = -2i\hbar\hat{\mathbf{L}} \times \hat{\mathbf{p}} - 2\hbar^2\hat{\mathbf{p}}$

Problem 4:

- Evaluate the commutators $[\hat{L}_z, r^2]$ and $[\hat{L}_z, p^2]$.

  Answer: $[\hat{L}_z, r^2] = 0, \quad [\hat{L}_z, p^2] = 0$
- Show that the Hamiltonian $\hat{H} = (p^2/2m) + V$ commutes with all three components of $\hat{\mathbf{L}}$, provided that V depends only on r. (Thus $\hat{H}, \hat{L}^2$, and $\hat{L}_z$ form a CSCO.)

2.2.1 Eigenvalues of orbital angular momentum

The results of the previous section imply that $\hat{L}^2$ is compatible with each component of $\hat{\mathbf{L}}$ and so it is possible to find simultaneous eigenstates of $\hat{L}^2$ and $\hat{L}_z$ for example (although note that there is nothing special about the choosing the z-component. We could just as well have chosen the x or y component). We then have:

$$\hat{L}^2 Y(x, y, z) = \lambda Y(x, y, z) \tag{2.37}$$

and

$$\hat{L}_z Y(x, y, z) = \mu Y(x, y, z) \tag{2.38}$$

where $Y(x, y, z)$ is an eigenfunction of both $\hat{L}^2$ and $\hat{L}_z$.

To determine the eigenvalues, we use the "ladder operator" technique, similar to the one used to study the harmonic oscillator. We define:

$$\hat{L}_\pm \equiv \hat{L}_x \pm i\hat{L}_y \tag{2.39}$$

which satisfy the following commutation relation

$$[\hat{L}_z, \hat{L}_\pm] = \pm\hbar\hat{L}_\pm \tag{2.40}$$

and

$$[\hat{L}^2, \hat{L}_\pm] = 0 \tag{2.41}$$

From the result above, we have

$$\begin{aligned} \hat{L}^2(\hat{L}_\pm Y) &= \hat{L}_\pm(\hat{L}^2 Y) \\ &= \hat{L}_\pm(\lambda Y) \\ &= \lambda(\hat{L}_\pm Y) \end{aligned} \tag{2.42}$$

So if Y is an eigenfunction of $\hat{L}^2$ then so is $\hat{L}_\pm Y$, with the *same* eigenvalue of $\hat{L}^2$. Now consider the following:

$$\begin{aligned} \hat{L}_z(\hat{L}_\pm Y) &= (\hat{L}_z\hat{L}_\pm - \hat{L}_\pm\hat{L}_z)Y + \hat{L}_\pm\hat{L}_z Y \\ &= \pm\hbar\hat{L}_\pm Y + \hat{L}_\pm(\mu Y) \\ &= (\mu \pm \hbar)(\hat{L}_\pm Y) \end{aligned} \tag{2.43}$$

This means that $(\hat{L}_\pm Y)$ is also an eigenfunction of $\hat{L}_z$, but with a *different* eigenvalue, $\mu \pm \hbar$. $\hat{L}_+$ is called a "raising" operator because it increases the value of $\hat{L}_z$ by $\hbar$, whereas $\hat{L}_-$ is called a "lowering " operator since it decreases the eigenvalue of $\hat{L}_z$ by $\hbar$. Therefore starting from a given value of λ one obtains a "ladder" of states with each "rung" separated from its neighbours by one unit of $\hbar$ in the eigenvalue of $\hat{L}_z$. Then to go up the ladder, one applies $\hat{L}_+$ while to go down, $\hat{L}_-$. This process cannot go on forever though: Eventually we are going to reach a state for which the z-component of its AM exceeds the *total* AM of that state and this cannot be!

Problem 5: Prove that if Y is simultaneously an eigenfunction of $\hat{L}^2$ and $\hat{L}_z$ then the square of the eigenvalue of $\hat{L}_z$ cannot exceed the eigenvalue of $\hat{L}^2$. (Hint: Examine the expectation value of $\hat{L}^2$).

The above problem indicates that there must exist a "top rung", Y_{top}, such that

$$\hat{L}_+ Y_{top} = 0 \tag{2.44}$$

Let the eigenvalue of $\hat{L}_z$ for this "top" state to be $\hbar l$. Then we have

$$\begin{aligned} \hat{L}_z Y_{top} &= \hbar l Y_{top} \\ \hat{L}^2 Y_{top} &= \lambda Y_{top} \end{aligned} \tag{2.45}$$

Problem 6: Prove that

$$\hat{L}^2 = \hat{L}_\pm \hat{L}_\mp + \hat{L}_z^2 \mp \hbar \hat{L}_z \tag{2.46}$$

Using the above result it follows that

$$\begin{aligned}
\hat{L}^2 Y_{top} &= (\hat{L}_- \hat{L}_+ + \hat{L}_z^2 + \hbar \hat{L}_z) Y_{top} \\
&= (0 + \hbar^2 l^2 + \hbar^2 l) Y_{top} \\
&= \hbar^2 l(l+1) Y_{top}
\end{aligned} \tag{2.47}$$

and hence

$$\lambda = \hbar^2 l(l+1) \tag{2.48}$$

which tells us the eigenvalue of $\hat{L}^2$ in terms of the maximum eigenvalue of $\hat{L}_z$. Following the same reasoning, there is also a "bottom" rung, Y_{bottom}, such that

$$\hat{L}_- Y_{bottom} = 0 \tag{2.49}$$

with $\hbar \bar{l}$ the eigenvalue of $\hat{L}_z$. Following the same method above, we deduce that

$$\lambda = \hbar^2 \bar{l}(\bar{l} - 1) \tag{2.50}$$

Problem 7: Verify (2.50). Hence deduce that the only possible value of $\bar{l}$ is

$$\bar{l} = -l \tag{2.51}$$

Evidently the eigenvalues of $\hat{L}_z$ are $m\hbar$, where m is an integer going from $+l$ to $-l$ in N integer steps. This implies that $l = \frac{N}{2}$ so that l must be an integer or half-integer. The eigenfunctions of $\hat{L}^2$ and $\hat{L}_z$ are characterised by the numbers l and m, where:

$$\begin{aligned}
\hat{L}^2 Y_{l,m} &= \hbar^2 l(l+1) Y_{l,m} \\
l &= 0, \frac{1}{2}, 1, \frac{3}{2}, \cdots
\end{aligned} \tag{2.52}$$

$$\begin{aligned} \hat{L}_z Y_{l,m} &= \hbar m Y_{l,m} \\ m &= -l, -l+1, \ldots, l-1, l \end{aligned} \tag{2.53}$$

Note that for a given value of l, there are $2l+1$ different values for m.

Problem 8: The raising and lowering operators change the value of m by one unit, i.e.

$$\hat{L}_\pm Y_{l,m} = A^\pm_{l,m} Y_{l,m\pm 1} \tag{2.54}$$

What is the value of $A^\pm_{l,m}$, if the eigenfunctions $Y_{l,m}$ are to be normalised?

Answer: $A^\pm_{l,m}$ can take the values:

$$\begin{aligned} A^\pm_{l,m} &= \hbar\sqrt{l(l+1) - m(m\pm 1)} \\ &= \hbar\sqrt{(l \mp m)(l \pm m + 1)} \end{aligned} \tag{2.55}$$

This result will be of use later when we study the addition of two angular momenta.

2.2.2 Eigenfunctions of orbital angular momentum

We now determine the eigenfunctions for the operators $\hat{L}^2$ and $\hat{L}_z$. The conventional way is to write out these operators in spherical polar coordinates:

$$\begin{aligned} x &= r\sin\theta\cos\varphi \\ y &= r\sin\theta\sin\varphi \\ z &= r\cos\theta \end{aligned} \tag{2.56}$$

from which we can derive:

$$\begin{aligned} \frac{\partial}{\partial x} &= \sin\theta\cos\varphi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\cos\varphi\frac{\partial}{\partial\theta} - \frac{\sin\varphi}{r\sin\theta}\frac{\partial}{\partial\varphi} \\ \frac{\partial}{\partial y} &= \sin\theta\sin\varphi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\sin\varphi\frac{\partial}{\partial\theta} + \frac{\cos\varphi}{r\sin\theta}\frac{\partial}{\partial\varphi} \\ \frac{\partial}{\partial z} &= \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta} \end{aligned} \tag{2.57}$$

Then the $\hat{L}_z$ operator in spherical polar coordinates is:

$$\hat{L}_z = -i\hbar\frac{\partial}{\partial\varphi} \tag{2.58}$$

while the raising/lowering operators are:

$$\hat{L}_\pm = \hbar e^{\pm i\varphi}(\pm\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\varphi}) \tag{2.59}$$

Problem 9: Verify this. From these results determine the value of $\hat{L}^2$ in spherical polar coordinates. (Hint: You could use the result proven earlier, i.e. $\hat{L}^2 = \hat{L}_+\hat{L}_- + \hat{L}_z^2 - \hbar\hat{L}_z$).

Since we have already shown that $Y_{lm}(\theta,\varphi)$ is a simultaneous eigenfunction of $\hat{L}_z$ and $\hat{L}^2$, from the above results it is clear that the easiest way to determine $Y_{lm}(\theta,\varphi)$ is by using the method of separation of variables:

$$Y_{lm}(\theta,\varphi) = \Theta_{lm}(\theta)\Phi_m(\varphi) \tag{2.60}$$

so that

$$\hat{L}_z Y_{lm}(\theta,\varphi) = m\hbar Y_{lm}(\theta,\varphi) \tag{2.61}$$

implies

$$\frac{d\Phi_m(\varphi)}{d\varphi} = im\Phi_m(\varphi) \tag{2.62}$$

which is a first-order differential equation in φ. The normalised solutions are:

$$\begin{aligned} \Phi_m(\varphi) &= \frac{1}{\sqrt{(2\pi)}}e^{im\varphi} \\ \int_0^{2\pi} d\varphi\, |\Phi_m|^2 &= 1 \end{aligned} \tag{2.63}$$

Therefore

$$Y_{lm}(\theta,\varphi) = \Theta_{lm}(\theta)e^{im\varphi} \tag{2.64}$$

To solve for $\Theta_{lm}(\theta)$ we proceed as follows. We use the fact that the maximum value of m is l such that:

$$\hat{L}_+ Y_{ll} = 0 \tag{2.65}$$

Expanding this out gives the equation

$$\hbar e^{i\varphi}\left(\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\varphi}\right)\Theta_{ll}e^{il\varphi} = \hbar e^{i(l+1)\varphi}\left(\frac{\partial}{\partial\theta} - l\cot\theta\right)\Theta_{ll}(\theta) = 0 \tag{2.66}$$

whose solution is

$$\Theta_{ll}(\theta) = (\sin\theta)^l \tag{2.67}$$

Starting from this, we can determine all the other states by applying the lowering operator. This means that an arbitrary state can be represented by:

$$Y_{lm}(\theta,\varphi) = C\hat{L}_-^{l-m}(\sin\theta)^l e^{il\varphi} \tag{2.68}$$

where C is a constant.

Problem 10: We investigate the action of $\hat{L}_-$ on the "top" state. First verify the following identity:

$$\left(\frac{\partial}{\partial\theta} + l\cot\theta\right)f(\theta) = \frac{1}{(\sin\theta)^l}\frac{d}{d\theta}\Big[(\sin\theta)^l f(\theta)\Big] \tag{2.69}$$

Then show that by acting on Y_{ll} with $\hat{L}_-$ we get:

$$\begin{aligned} Y_{l,l-1} &= C'\frac{e^{i(l-1)\varphi}}{(\sin\theta)^l}\Big(-\frac{d}{d\theta}\Big)\Big[(\sin\theta)^l(\sin\theta)^l\Big] \\ Y_{l,l-2} &= C''\frac{e^{i(l-2)\varphi}}{(\sin\theta)^{(l-1)}}\frac{d}{d\theta}\Big[\frac{1}{\sin\theta}\frac{d}{d\theta}(\sin\theta)^{2l}\Big] \end{aligned} \tag{2.70}$$

where C', C'' are constants.

If we define a new variable $u = \cos\theta$ such that $-\frac{1}{\sin\theta}\frac{d}{d\theta} = \frac{d}{du}$ then the general form for $Y_{lm}(\theta,\varphi)$ is

$$Y_{lm}(\theta,\varphi) = C\frac{e^{im\varphi}}{(\sin\theta)^m}\left(\frac{d}{du}\right)^{l-m}[(1-u^2)^l] \tag{2.71}$$

These eigenfunctions (the **spherical harmonics**) are to be normalised over the unit sphere where the range of integration is $0 \le \varphi \le 2\pi$ and $0 \le \theta \le \pi$. This means

$$\begin{aligned} \int d\Omega\,|Y_{lm}(\theta,\varphi)|^2 &= \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta\sin\theta\,|Y_{lm}(\theta,\varphi)|^2 \\ &= \int_0^{2\pi} d\varphi \int_{-1}^{1} du\,|Y_{lm}(\theta,\varphi)|^2 \\ &= 1 \end{aligned} \tag{2.72}$$

2.2.3 Mathematical interlude: Legendre polynomials and spherical harmonics

Legendre polynomials

The Legendre polynomials $P_l(u)$ arise when we solve Laplace's equation in spherical polar coordinates for a problem possessing azimuthal symmetry. They are real functions (polynomials of order l) defined in the interval $[-1, +1]$ and have the following properties:

The generating function is

$$\frac{1}{(1 - 2hu + h^2)^{\frac{1}{2}}} = \sum_{l=0}^{\infty} P_l(u) h^l \tag{2.73}$$

$$P_l(u) = \frac{1}{2^l l!} \frac{d^l}{du^l} (u^2 - 1)^l \tag{2.74}$$

where $-1 \leq u \leq 1$, and l is a non-negative integer. The $P_l(u)$ arise from a power-series solution of the 2^{nd} order differential equation, the *Legendre equation*:

$$(1 - u^2) P_l^{''}(u) - 2u P_l^{'}(u) + l(l+1) P_l(u) = 0 \tag{2.75}$$

and satisfy the following recurrence relations

$$\begin{aligned} (l+1) P_{l+1}(u) &= (2l+1) u P_l(u) - l P_{l-1}(u) \\ (1 - u^2) P_l^{'}(u) &= -l u P_l(u) + l P_{l-1}(u) \end{aligned} \tag{2.76}$$

They are normalised as follows:

$$\int_{-1}^{1} P_l(u) P_m(u) du = \frac{2}{2l+1} \delta_{lm} \tag{2.77}$$

The first few polynomials are

$$\begin{aligned} P_0(u) &= 1 \\ P_1(u) &= u \\ P_2(u) &= \frac{1}{2}(3u^2 - 1) \\ P_3(u) &= \frac{1}{2}(5u^3 - 3u) \end{aligned} \tag{2.78}$$

Note that

$$\begin{aligned} P_l(u) &= (-1)^l P_l(-u) \\ P_l(1) &= 1 \end{aligned} \tag{2.79}$$

Spherical harmonics

The **spherical harmonics**, $Y_{lm}(\theta,\varphi)$, that are conventionally used are:

$$Y_{lm}(\theta,\varphi) = (-1)^m \left[\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}\right]^{\frac{1}{2}} P_l^m(\cos\theta)e^{im\varphi} \tag{2.80}$$

which is valid for $m \geq 0$, while the eigenfunctions corresponding to negative values of m are obtained from:

$$Y_{l,-m} = (-1)^m Y_{lm}^* \tag{2.81}$$

The functions P_l^m are the associated Legendre polynomials defined by:

$$P_l^m(u) = (-1)^{l+m}\frac{(l+m)!}{(l-m)!}\frac{(1-u^2)^{-m/2}}{2^l l!}\left(\frac{d}{du}\right)^{l-m}(1-u^2)^l \tag{2.82}$$

which is valid for $m \geq 0$, and the values for negative m given by:

$$P_l^{-m}(u) = (-1)^m \frac{(l-m)!}{(l+m)!}P_l^m(u) \tag{2.83}$$

(Remember $u = \cos\theta$).

Problem 11†: From the above definitions of $Y_{lm}(\theta,\varphi)$ and $P_l^m(\cos\theta)$, determine the values of $Y_{00}, Y_{10}, Y_{11}, Y_{20}, Y_{21}, Y_{22}$. Write out the quantities $x, y, z, xy, yz, zx, x^2-y^2, 2z^2-x^2-y^2, y^2-z^2$ in terms of the above $Y_{lm}(\theta,\varphi)$ (and their complex conjugates). (For example, $z = r\sqrt{\frac{4\pi}{3}}Y_{1,0}$ where $r^2 = x^2+y^2+z^2$.) Hence show that $\frac{x}{r}$ is an eigenfunction of $\hat{L}_x$ and $\frac{y}{r}$ is an eigenfunction of $\hat{L}_y$, and find their respective eigenvalues.

When $m = 0$, (2.80) simplifies to a value independent of the azimuthal angle φ

$$Y_{l0}(\theta,\varphi) \to Y_{l0}(\theta,0) = \sqrt{\frac{2l+1}{4\pi}}P_l(\cos\theta) \tag{2.84}$$

where $P_l(\cos\theta)$ are the Legendre polynomials.

At the poles $\theta = 0$ and $\theta = \pi$ the azimuthal angle φ are indistinguishable so $Y_{lm}(\theta,\varphi)$ cannot depend on φ at these angles. Therefore the only non-zero spherical harmonics for these two cases has to have $m = 0$ i.e.

$$Y_{lm}(0,\varphi) = \sqrt{\frac{2l+1}{4\pi}}\delta_{m0} \tag{2.85}$$

and

$$Y_{lm}(\pi,\varphi) = (-1)^l\sqrt{\frac{2l+1}{4\pi}}\delta_{m0} \tag{2.86}$$

Consider two arbitrary directions in space defined respectively by the angles (θ, φ) and (θ', φ') and call the angle between them Θ. The **spherical harmonic addition theorem** states that

$$P_l(\cos\Theta) = \frac{4\pi}{2l+1} \sum_{m=-l}^{m=+l} Y_{lm}(\theta, \varphi)\, Y_{lm}^*(\theta', \varphi') \tag{2.87}$$

Problem 12: Let $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$ be the distance between two points in space so that the $\mathbf{r}_1$ direction is (θ_1, φ_1) and $\mathbf{r}_2$ direction is (θ_2, φ_2). Prove that

$$\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} = \sum_{l=0}^{\infty} \sum_{m=-l}^{m=+l} \frac{4\pi}{(2l+1)} \frac{r_<^l}{r_>^{l+1}} Y_{lm}(\theta_1, \varphi_1) Y_{lm}^*(\theta_2, \varphi_2) \tag{2.88}$$

where $r_<$ stands for the smaller of the two distances $|\mathbf{r}_1|$ and $|\mathbf{r}_2|$, and $r_>$ is the larger of the two distances.

This result is very useful and we will apply it when we study scattering theory as well as in variational calculations for the ground state of the helium atom.

Note that the $Y_{lm}(\theta, \varphi)$ form a complete set of orthonormal functions, i.e.

$$\int d\Omega\; Y_{lm}(\theta, \varphi)\; Y_{l'm'}^*(\theta, \varphi) = \delta_{ll'}\delta_{mm'} \tag{2.89}$$

and

$$\sum_{l=0}^{\infty} \sum_{m=-l}^{m=+l} Y_{lm}(\theta, \varphi) Y_{lm}^*(\theta', \varphi') = \frac{1}{\sin\theta}\delta(\theta - \theta')\delta(\varphi - \varphi') \tag{2.90}$$

This means that any function of θ and φ can be expanded as follows:

$$f(\theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{m=+l} C_{lm} Y_{lm}(\theta, \varphi) \tag{2.91}$$

where

$$C_{lm} = \int d\Omega\; Y_{lm}^*(\theta, \varphi) f(\theta, \varphi) \tag{2.92}$$

Furthermore, if $f(\theta, \varphi)$ is the angular wavefunction of some state, normalised such that

$$\int d\Omega |f(\theta, \varphi)|^2 = 1 \tag{2.93}$$

then the $|C_{lm}|^2$ are the probabilities that simultaneous measurement of $\hat{L}^2$ and $\hat{L}_z$ on the state described by $f(\theta, \varphi)$ yields $l(l+1)\hbar^2$ and $m\hbar$ respectively.

In the Dirac notation, an arbitrary state $|\psi\rangle$ can be expanded as:

$$|\psi\rangle = \sum_{lm} C_{lm}|lm\rangle \tag{2.94}$$

where with the help of the orthonormality condition

$$\langle l'm'|lm\rangle = \delta_{ll'}\delta_{mm'} \tag{2.95}$$

we get

$$C_{lm} = \langle lm|\psi\rangle \tag{2.96}$$

These eigenstates are also complete so:

$$\sum_{l=0}^{\infty}\sum_{m=-l}^{m=+l} |lm\rangle\langle lm| = \hat{I} \tag{2.97}$$

Problem 13: Using $\langle lm'|\hat{L_z}|lm\rangle = \hbar m \delta_{m'm}$ and

$$\langle lm'|\hat{L_\pm}|lm\rangle = \hbar\Big[l(l+1) - m(m\pm 1)\Big]^{\frac{1}{2}}\delta_{m',m\pm 1}$$

calculate the matrix representation of $\hat{L}_x$, $\hat{L}_y$ and $\hat{L}_z$ for the AM state 3/2. Check that the commutation relations hold.

Problem 14: The Hamiltonian of a *rigid rotator* with principal moments of inertia I_1, I_2, I_3 is

$$\hat{H} = \frac{1}{2I_1}\hat{L}_x^2 + \frac{1}{2I_2}\hat{L}_y^2 + \frac{1}{2I_3}\hat{L}_z^2$$

Find the eigenvalues of $\hat{H}$ if the AM of the system is 1.

Answer: The eigenvalues are $\hbar^2(\alpha+\beta)$, $\hbar^2(\beta+\gamma)$, and $\hbar^2(\alpha+\gamma)$, where $\alpha = \frac{1}{2I_1}$, $\beta = \frac{1}{2I_2}$, $\gamma = \frac{1}{2I_3}$.

2.2.4 Angular momentum and rotational invariance

The fact that the Hamiltonian commutes with the components of the AM is a consequence of the invariance of the Hamiltonian under reorientation of the coordinate axes used to describe the system. For example, a rotation of the axes through an angle α around the z-axis changes the coordinate system from S to S' where

$$\begin{aligned} x' &= x\cos\alpha + y\sin\alpha \\ y' &= -x\sin\alpha + y\cos\alpha \\ z' &= z \end{aligned} \tag{2.98}$$

or in terms of spherical polar coordinates

$$r' = r\ , \quad \theta' = \theta\ , \quad \phi' = \phi - \alpha \tag{2.99}$$

Notice that by Taylor expansion we can write

$$\begin{aligned} \psi(r,\theta,\phi-\alpha) &= \sum_{n=0}^{\infty} \frac{1}{n!}\left(-\alpha\frac{\partial}{\partial\phi}\right)^n \psi(r,\theta,\phi) \\ &= \exp\left(-\alpha\frac{\partial}{\partial\phi}\right)\psi(r,\theta,\phi) \\ &= \exp\left(-\frac{i}{\hbar}\alpha\hat{L}_z\right)\psi(r,\theta,\phi) \end{aligned} \tag{2.100}$$

The operator

$$\hat{U} = \exp\left(-\frac{i}{\hbar}\alpha\hat{L}_z\right) \tag{2.101}$$

is an example of a *rotation operator*, in this case a rotation of the axes through an angle α, or equivalently a rotation of the system through an angle $-\alpha$, about the z-axis. More generally, a rotation of the system through an angle ω around the unit vector $\mathbf{n}$ will be achieved by the operator

$$\hat{U} = \exp\left(\frac{i}{\hbar}\omega\hat{\mathbf{L}}\cdot\mathbf{n}\right) \tag{2.102}$$

Note that a rotation operator is unitary:

$$\hat{U}^\dagger = \exp\left(-\frac{i}{\hbar}\omega\hat{\mathbf{L}}\cdot\mathbf{n}\right) = \hat{U}^{-1} \tag{2.103}$$

The fact that the components of $\hat{\mathbf{L}}$ commute with the Hamiltonian ensures that any rotation operator also commutes with $\hat{H}$. It follows that, if ψ is a possible state of the system with energy E, then so is $\hat{U}\psi$, with the same energy: if $\hat{H}\psi = E\psi$ and $\hat{U}\hat{H} = \hat{H}\hat{U}$, then

$$\hat{H}(\hat{U}\psi) = \hat{U}\hat{H}\psi = \hat{U}E\psi = E(\hat{U}\psi) \tag{2.104}$$

Conversely, if every rotated state $\hat{U}\psi$ is an eigenstate of $\hat{H}$ with the same energy as ψ, then all components of $\hat{\mathbf{L}}$ must commute with $\hat{H}$, which implies that $\mathbf{L}$ is a conserved quantity.

2.3 Spin angular momentum

The electron possesses an internal angular momentum called the **spin** which can assume only the values $+\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$ in some arbitrarily chosen direction. In fact, all elementary particles have a spin degree of freedom. **Fermions** possess half-integral spin while **bosons** have integral spin (including zero). In what follows, we will develop the theory for spin-$\frac{1}{2}$ fermions.

Let the spin operator be $\hat{\mathbf{S}} = (\hat{S}_x, \hat{S}_y, \hat{S}_z)$. If $\mathbf{n}$ is a unit vector (pointing in some arbitrary direction), then the Stern-Gerlach experiment indicates that the eigenvalue of the operator $\hat{\mathbf{S}} \cdot \mathbf{n}$ has only two values (which turns out to be $\pm\frac{\hbar}{2}$) i.e.

$$\hat{\mathbf{S}} \cdot \mathbf{n}|\mathbf{n}, \pm\rangle = \pm\frac{\hbar}{2}|\mathbf{n}, \pm\rangle \tag{2.105}$$

Without loss of generality, one can choose $\mathbf{n}$ to point in the $z-$direction. (Then $\hat{\mathbf{S}} \cdot \mathbf{n} = \hat{S}_z$). The eigenvalue equation then takes the form:

$$\begin{aligned} \hat{S}_z|\uparrow\rangle &= \frac{\hbar}{2}|\uparrow\rangle \\ \hat{S}_z|\downarrow\rangle &= -\frac{\hbar}{2}|\downarrow\rangle \end{aligned} \tag{2.106}$$

where $|\uparrow\rangle$ corresponds to the spin "pointing" in the positive z-axis direction (i.e a *spin-up* state) and $|\downarrow\rangle$ corresponds to the spin "pointing" in the negative z-axis direction (i.e. a *spin-down* state). Since spin is a physical observable, $\hat{S}_z$ is Hermitian and the states belonging to distinct eigenvalues are orthogonal, that is:

$$\langle\uparrow|\downarrow\rangle = \langle\downarrow|\uparrow\rangle = 0 \tag{2.107}$$

We further normalise them to unity:

$$\langle\uparrow|\uparrow\rangle = \langle\downarrow|\downarrow\rangle = 1 \tag{2.108}$$

The spin operators satisfy the AM commutation relations:

$$\begin{aligned} \left[\hat{S}_i, \hat{S}_j\right] &= i\hbar\epsilon_{ijk}\hat{S}_k \\ \left[\hat{S}_z, \hat{S}_\pm\right] &= \pm\hbar\hat{S}_\pm \\ \left[\hat{S}_+, \hat{S}_-\right] &= 2\hbar\hat{S}_z \end{aligned} \tag{2.109}$$

where

$$\hat{S}_\pm = \hat{S}_x \pm i\hat{S}_y$$

$$\begin{aligned}\hat{S}_x &= \frac{1}{2}(\hat{S}_+ + \hat{S}_-)\\ \hat{S}_y &= -i\frac{1}{2}(\hat{S}_+ - \hat{S}_-)\end{aligned} \tag{2.110}$$

For spin $S = \frac{1}{2}$, $\hat{S}^2$ has the eigenvalue $\frac{3}{4}\hbar^2$:

$$\begin{aligned}\hat{S}^2|\uparrow\rangle = \frac{3}{4}\hbar^2|\uparrow\rangle\\ \hat{S}^2|\downarrow\rangle = \frac{3}{4}\hbar^2|\downarrow\rangle\end{aligned} \tag{2.111}$$

Problem 15: Show that:

$$\begin{aligned}\hat{S}_+|\uparrow\rangle &= 0\\ \hat{S}_-|\uparrow\rangle &= \hbar|\downarrow\rangle\\ \hat{S}_+|\downarrow\rangle &= \hbar|\uparrow\rangle\\ \hat{S}_-|\downarrow\rangle &= 0\end{aligned} \tag{2.112}$$

We can now represent the spin operators in the basis states $|\uparrow\rangle$ and $|\downarrow\rangle$ by the spin matrices

$$\hat{S}_i \rightarrow (S_i) = \begin{pmatrix} \langle\uparrow|\hat{S}_i|\uparrow\rangle & \langle\uparrow|\hat{S}_i|\downarrow\rangle \\ \langle\downarrow|\hat{S}_i|\uparrow\rangle & \langle\downarrow|\hat{S}_i|\downarrow\rangle \end{pmatrix} \tag{2.113}$$

Problem 16: Show that:

$$\hat{S}_+ = \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{2.114}$$

$$\hat{S}_- = \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{2.115}$$

$$\hat{S}_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.116}$$

Introducing the **Pauli spin matrices** by

$$\mathbf{S} = \frac{\hbar}{2}\sigma \tag{2.117}$$

we obtain for them

$$\begin{aligned}
\sigma_x &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
\sigma_y &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\
\sigma_z &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}
\end{aligned} \tag{2.118}$$

Problem 17: Show that:

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I \tag{2.119}$$

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k \tag{2.120}$$

$$\sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij} \tag{2.121}$$

$$Tr(\sigma_i) = 0 \tag{2.122}$$

$$|\sigma_i| = -1 \tag{2.123}$$

where I is the 2×2 unit matrix. Show that the above results can be summarised compactly by the identity

$$\sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k \tag{2.124}$$

Hence, given arbitrary vectors $\mathbf{a}, \mathbf{b}$ show that

$$(\sigma \cdot \mathbf{a})(\sigma \cdot \mathbf{b}) = I\mathbf{a} \cdot \mathbf{b} + i\sigma \cdot (\mathbf{a} \times \mathbf{b}) \tag{2.125}$$

Problem 18:

Spin state space is a two-dimensional Hilbert space H_S, spanned by the orthonormal basis vectors $\{|\uparrow\rangle, |\downarrow\rangle\}$. In matrix notation, these basis vectors are defined as:

$$\begin{aligned}
|\uparrow\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
|\downarrow\rangle &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}
\end{aligned} \tag{2.126}$$

These basis vectors are the eigenvectors of the $\hat{S}_z$, i.e.

$$\begin{aligned} \hat{S}_z|\uparrow\rangle &= +\frac{\hbar}{2}|\uparrow\rangle \\ \hat{S}_z|\downarrow\rangle &= -\frac{\hbar}{2}|\downarrow\rangle \end{aligned} \tag{2.127}$$

with

$$\begin{aligned} \langle\uparrow|\uparrow\rangle &= 1 \\ \langle\downarrow|\downarrow\rangle &= 1 \\ \langle\uparrow|\downarrow\rangle &= 0 \\ \langle\downarrow|\uparrow\rangle &= 0 \end{aligned} \tag{2.128}$$

This basis is complete:

$$|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow| = \hat{I} \tag{2.129}$$

In general, $\forall|\psi\rangle \in H_S$, we have:

$$\begin{aligned} &|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle \\ &|\alpha|^2 + |\beta|^2 = 1 \end{aligned} \tag{2.130}$$

(Note that in general α and β are complex numbers.) In the $\{|\uparrow\rangle, |\downarrow\rangle\}$ basis, the matrix representation of $\hat{S}_z$ is diagonal:

$$(S_z) = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.131}$$

while the $x-$ and $y-$components of the spin operator are:

$$\begin{aligned} (S_x) &= \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ (S_y) &= \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \end{aligned} \tag{2.132}$$

- **Show that** the normalised eigenvectors of $\hat{S}_x$ (denoted by $|\pm\rangle_x$) and $\hat{S}_y$ (denoted by $|\pm\rangle_y$) when expanded in the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$ are:

$$\begin{aligned} |\pm\rangle_x &= \frac{1}{\sqrt{2}}[|\uparrow\rangle \pm |\downarrow\rangle] \\ |\pm\rangle_y &= \frac{1}{\sqrt{2}}[|\uparrow\rangle \pm i|\downarrow\rangle] \end{aligned} \tag{2.133}$$

- Let **u** be a unit vector characterised by polar angles ϑ and φ with Cartesian components:

  $$\mathbf{u} = (\sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi, \cos\vartheta) \tag{2.134}$$

  Show that if $S_u = \mathbf{S}\cdot\mathbf{u}$ then:

  $$(S_u) = \frac{\hbar}{2}\begin{pmatrix} \cos\vartheta & \sin\vartheta e^{-i\varphi} \\ \sin\vartheta e^{i\varphi} & -\cos\vartheta \end{pmatrix} \tag{2.135}$$

  Then **show that** in the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$, the eigenvectors of S_u are:

  $$|\uparrow\rangle_u = \cos(\frac{\vartheta}{2})e^{-i\frac{\varphi}{2}}|\uparrow\rangle + \sin(\frac{\vartheta}{2})e^{i\frac{\varphi}{2}}|\downarrow\rangle \tag{2.136}$$

  $$|\downarrow\rangle_u = -\sin(\frac{\vartheta}{2})e^{-i\frac{\varphi}{2}}|\uparrow\rangle + \cos(\frac{\vartheta}{2})e^{i\frac{\varphi}{2}}|\downarrow\rangle \tag{2.137}$$

2.3.1 Spinors

In the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$, a general spin state $|\chi\rangle$ can be written as

$$|\chi\rangle = a_+|\uparrow\rangle + a_-|\downarrow\rangle \tag{2.138}$$

with complex coefficients $a_\pm$. Normalisation requires that

$$|a_+|^2 + |a_-|^2 = 1 \tag{2.139}$$

The state $|\chi\rangle$ defined above can also be represented by a two-component column vector called a **spinor** whose components are given by the projections onto the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$:

$$|\chi\rangle = \begin{pmatrix} a_+ \\ a_- \end{pmatrix} = \begin{pmatrix} \langle\uparrow|\chi\rangle \\ \langle\downarrow|\chi\rangle \end{pmatrix} \tag{2.140}$$

The basis spinors are

$$\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2.141}$$

and the completeness relation in this matrix representation is

$$\chi_+\chi_+^\dagger + \chi_-\chi_-^\dagger = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.142}$$

Problem 19: An electron is in the spin state

$$|\chi\rangle = \frac{1}{3}|\uparrow\rangle + \frac{2\sqrt{2}}{3}|\downarrow\rangle$$

- What is the probability that a measurement of the $z-$component of the spin will result in $\frac{1}{2}\hbar$? In $-\frac{1}{2}\hbar$?

  Answer: $\frac{1}{9}$ and $\frac{8}{9}$.

- What is the expectation value of $\hat{S}_z$?

  Answer: $-\frac{7}{18}\hbar$

- What is the RMS uncertaintity in $\hat{S}_z$?

  Answer: $\frac{2\sqrt{2}}{9}\hbar$.

Spin is an additional degree of freedom *independent* of the spatial degrees of freedom. Spin and position (or momentum) can assume precise values simultaneously and independently of one another i.e.

$$[\hat{\mathbf{S}}, \hat{\mathbf{r}}] = [\hat{\mathbf{S}}, \hat{\mathbf{p}}] = [\hat{\mathbf{S}}, \hat{\mathbf{L}}] = 0 \tag{2.143}$$

The total quantum state of a particle is constructed from the **direct product** of the position and spin eigenstates. The states $|\mathbf{r}\rangle \otimes |\uparrow\rangle$ and $|\mathbf{r}\rangle \otimes |\downarrow\rangle$ forms a basis for the Hilbert space (that takes into account the spatial and spin degrees of freedom). A general state $|\Psi\rangle$ in this Hilbert space is then given by

$$|\Psi\rangle = \int d\mathbf{r}' \Big\{ \psi_+(\mathbf{r}')\ |\mathbf{r}'\rangle \otimes |\uparrow\rangle + \psi_-(\mathbf{r}')\ |\mathbf{r}'\rangle \otimes |\downarrow\rangle \Big\} \tag{2.144}$$

The projection onto position and spin eigenstates are

$$\begin{aligned} \langle \mathbf{r}|\Psi\rangle &= \int d\mathbf{r}' \Big\{ \psi_+(\mathbf{r}')\ \langle \mathbf{r}|\mathbf{r}'\rangle |\uparrow\rangle + \psi_-(\mathbf{r}')\ \langle \mathbf{r}|\mathbf{r}'\rangle |\downarrow\rangle \Big\} \\ &= \psi_+(\mathbf{r})|\uparrow\rangle + \psi_-(\mathbf{r})|\downarrow\rangle \end{aligned} \tag{2.145}$$

since $\langle \mathbf{r}|\mathbf{r}'\rangle = \delta(\mathbf{r} - \mathbf{r}')$ and

$$\begin{aligned} \langle \mathbf{r}| \Big(|\mathbf{r}'\rangle \otimes |\uparrow\rangle \Big) &= \langle \mathbf{r}|\mathbf{r}'\rangle |\uparrow\rangle \\ \langle \uparrow | \Big(\langle \mathbf{r}|\Psi\rangle \Big) &= \psi_+(\mathbf{r}) \\ \langle \downarrow | \Big(\langle \mathbf{r}|\Psi\rangle \Big) &= \psi_-(\mathbf{r}) \end{aligned} \tag{2.146}$$

The quantities $|\psi_\pm(\mathbf{r})|^2$ express the probability of finding the particle at position $\mathbf{r}$ with $z-$component of spin of $\pm\frac{\hbar}{2}$. The normalisation condition is

$$\langle \Psi|\Psi\rangle = \int d\mathbf{r}' \Big\{ \langle \mathbf{r}'| \otimes \langle \uparrow | \psi_+^*(\mathbf{r}') + \langle \mathbf{r}'| \otimes \langle \downarrow | \psi_-^*(\mathbf{r}') \Big\} \Big\{ |\mathbf{r}'\rangle \otimes |\uparrow\rangle \psi_+(\mathbf{r}') + |\mathbf{r}'\rangle \otimes |\downarrow\rangle \psi_-(\mathbf{r}') \Big\}$$

$$
\begin{aligned}
&= \int d\mathbf{r}' \Big[|\psi_+(\mathbf{r}')|^2 + |\psi_-(\mathbf{r}')|^2\Big] \\
&= 1
\end{aligned}
\tag{2.147}
$$

2.4 Addition of angular momenta

In general, if we have two angular momentum operators, $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$ (where the $\hat{\mathbf{J}}_i$ could correspond to the sum of the orbital AM and the spin AM of one particle, or the spin AM of two electrons, etc.), we would like to know the possible values of the total AM $\hat{\mathbf{J}}$ can take where

$$\hat{\mathbf{J}} = \hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2 \tag{2.148}$$

Assuming that $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$ correspond to distinct degrees of freedom, they commute with each other:

$$[\hat{\mathbf{J}}_1, \hat{\mathbf{J}}_2] = 0 \tag{2.149}$$

Together with the AM commutation relations for $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$ individually, the commutation relations for $\hat{\mathbf{J}}$ are given by:

$$[\hat{J}_i, \hat{J}_j] = i\hbar\epsilon_{ijk}\hat{J}_k \tag{2.150}$$

So all the properties of AM and their eigenstates discussed in the sections above also hold for the total AM.

We start with states $|j_1, m_1\rangle$ and $|j_2, m_2\rangle$ where the two quantum numbers j_1 and j_2 are fixed and the m_i take the values $-j_i, \cdots, +j_i$. The corresponding eigenvalue equations are:

$$
\begin{aligned}
\hat{J}_1^2|j_1, m_1\rangle &= \hbar^2 j_1(j_1+1)|j_1, m_1\rangle \\
\hat{J}_{1z}|j_1, m_1\rangle &= \hbar m_1|j_1, m_1\rangle
\end{aligned}
\tag{2.151}
$$

$$
\begin{aligned}
\hat{J}_2^2|j_2, m_2\rangle &= \hbar^2 j_2(j_2+1)|j_2, m_2\rangle \\
\hat{J}_{2z}|j_2, m_2\rangle &= \hbar m_2|j_2, m_2\rangle
\end{aligned}
\tag{2.152}
$$

From these states we construct the product states

$$|j_1, m_1, j_2, m_2\rangle = |j_1, m_1\rangle|j_2, m_2\rangle \tag{2.153}$$

which are eigenfunctions of $\hat{J}_z$ with eigenvalue $\hbar(m_1+m_2)$. However, these are NOT eigenfunctions of $\hat{J}^2$ since

$$\hat{J}^2 = \hat{J}_1^2 + \hat{J}_2^2 + 2\hat{\mathbf{J}}_1 \cdot \hat{\mathbf{J}}_2 \tag{2.154}$$

and the last term does not commute with $\hat{J}_{1z}$ or $\hat{J}_{2z}$:

$$[\hat{\mathbf{J}}_1 \cdot \hat{\mathbf{J}}_2, \hat{J}_{1z}] = [\hat{J}_{1x}, \hat{J}_{1z}]\hat{J}_{2x} + [\hat{J}_{1y}, \hat{J}_{1z}]\hat{J}_{2y} = i\hbar(-\hat{J}_{1y}\hat{J}_{2x} + \hat{J}_{1x}\hat{J}_{2y}) \tag{2.155}$$

and similarly

$$[\hat{\mathbf{J}}_1 \cdot \hat{\mathbf{J}}_2, \hat{J}_{2z}] = [\hat{J}_{1x}, \hat{J}_{1z}]\hat{J}_{2x} + [\hat{J}_{1y}, \hat{J}_{1z}]\hat{J}_{2y} = i\hbar(-\hat{J}_{1x}\hat{J}_{2y} + \hat{J}_{1y}\hat{J}_{2x}) \tag{2.156}$$

Thus we cannot simultaneously specify the quantum numbers m_1 and m_2 together with the eigenvalue of $\hat{J}^2$.

The product states (2.153) are eigenstates of the operators

$$\hat{J}_1^2, \hat{J}_{1z}, \hat{J}_2^2, \hat{J}_{2z} \tag{2.157}$$

What we need to do is search for states in which $\hat{J}^2$ is also diagonal; that is, we seek eigenfunctions

$$|j, m_j, j_1, j_2\rangle \tag{2.158}$$

of the four mutually commuting operators

$$\hat{J}^2, \hat{J}_z, \hat{J}_1^2, \hat{J}_2^2 \tag{2.159}$$

with eigenvalues $\hbar^2 j(j+1), \hbar m_j, \hbar^2 j_1(j_1+1), \hbar^2 j_2(j_2+1)$. At the same time we have to find the values taken by j (the corresponding m_j are then $-j, \cdots, j$), and we have to represent $|j, m_j, j_1, j_2\rangle$ as a linear combination of the product states above. Before considering the general case, we first look at the addition of two spins, and then at the addition of an orbital AM and a spin.

2.4.1 Addition of spin-$\frac{1}{2}$ operators

Let $\hat{\mathbf{S}}_1$ and $\hat{\mathbf{S}}_2$ be two spin-$\frac{1}{2}$ operators whose total spin AM is given by

$$\hat{\mathbf{S}} = \hat{\mathbf{S}}_1 + \hat{\mathbf{S}}_2 \tag{2.160}$$

There are four possible product states:

$$\begin{aligned} |++\rangle &= |\uparrow\rangle|\uparrow\rangle \\ |+-\rangle &= |\uparrow\rangle|\downarrow\rangle \\ |-+\rangle &= |\downarrow\rangle|\uparrow\rangle \\ |--\rangle &= |\downarrow\rangle|\downarrow\rangle \end{aligned} \tag{2.161}$$

in which the first (second) symbol refers to the first (second) spin. These product states are eigenstates of the operators $\hat{S}_1^2, \hat{S}_2^2, \hat{S}_{1z}, \hat{S}_{2z}$.

Problem 20:

- Evaluate the action of the operator $\hat{S}_z$ on each of the product states.
- Show that

  $$\begin{aligned} \hat{S}^2 &= \hat{S}_1^2 + \hat{S}_2^2 + 2\hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2 \\ &= \frac{3}{2}\hbar^2 + 2\hat{S}_{1z}\hat{S}_{2z} + \hat{S}_{1+}\hat{S}_{2-} + \hat{S}_{1-}\hat{S}_{2+} \end{aligned} \tag{2.162}$$

  Hence show that

  $$\begin{aligned} \hat{S}^2|--\rangle &= 2\hbar^2|--\rangle \\ \hat{S}^2|++\rangle &= 2\hbar^2|++\rangle \end{aligned} \tag{2.163}$$

  The states $|--\rangle$ and $|++\rangle$ therefore have total spin $S = 1$ and whose $z-$component of total spin takes the value of $-\hbar$ and $\hbar$ respectively.
- Given that $\hat{S}_- = \hat{S}_{1-} + \hat{S}_{2-}$ show that

  $$\frac{1}{\hbar\sqrt{2}}\hat{S}_-|++\rangle = \frac{1}{\sqrt{2}}(|+-\rangle + |-+\rangle) \tag{2.164}$$

  (The resulting state has been normalised to unity by inserting the factor $\frac{1}{\hbar\sqrt{2}}$. Show that the $z-$component of this state is zero. Using the notation $|S, m\rangle$ where S designates the total spin and m its $z-$component, we have

  $$\begin{aligned} |1,1\rangle &= |++\rangle \\ |1,0\rangle &= \frac{1}{\sqrt{2}}(|+-\rangle + |-+\rangle) \\ |1,-1\rangle &= |--\rangle \end{aligned} \tag{2.165}$$

- There is an additional state $|0,0\rangle$ which is orthogonal to those given above:

  $$|0,0\rangle = \frac{1}{\sqrt{2}}(|+-\rangle - |-+\rangle) \tag{2.166}$$

  By acting on this state with $\hat{S}^2$ and $\hat{S}_z$ show that this state has spin zero. So we have found all the eigenstates of $\hat{S}^2$ and $\hat{S}_z$. The states above corresponding to $S = 1$ are referred to as *triplet states* and that corresponding to $S = 0$ as a *singlet state*.

- Consider the following **projection operators**:

$$\hat{P}_{S=1} = \frac{3}{4} + \frac{1}{\hbar^2}\hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2$$

$$\hat{P}_{S=0} = \frac{1}{4} - \frac{1}{\hbar^2}\hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2 \tag{2.167}$$

Show that $\hat{P}_{S=1}$ projects onto triplet states while $\hat{P}_{S=0}$ projects onto singlet states.

2.4.2 Addition of spin-$\frac{1}{2}$ and orbital angular momentum

Starting with $\hat{\mathbf{L}}$ and $\hat{\mathbf{S}}$, define the total AM as:

$$\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}} \tag{2.168}$$

The eigenstates of the operators $\hat{L}^2, \hat{L}_z, \hat{S}^2, \hat{S}_z$ are given by:

$$\begin{aligned}
\hat{L}_z|l, m_l\rangle &= \hbar m_l|l, m_l\rangle \\
\hat{L}^2|l, m_l\rangle &= \hbar^2 l(l+1)|l, m_l\rangle \\
\hat{S}_z|\uparrow\rangle &= \frac{\hbar}{2}|\uparrow\rangle \\
\hat{S}_z|\downarrow\rangle &= -\frac{\hbar}{2}|\downarrow\rangle \\
\hat{S}^2|\uparrow\rangle &= \frac{3}{4}\hbar^2|\uparrow\rangle \\
\hat{S}^2|\downarrow\rangle &= \frac{3}{4}\hbar^2|\downarrow\rangle
\end{aligned} \tag{2.169}$$

where $m_l = -l, \cdots, l$. From these states one can form $2(2l+1)$ product states:

$$\begin{aligned}
&|l, m_l\rangle \otimes |\uparrow\rangle \\
&|l, m_l\rangle \otimes |\downarrow\rangle
\end{aligned} \tag{2.170}$$

However, these states are NOT eigenstates of the total AM, $\hat{J}^2$. We therefore seek eigenstates of $\hat{J}^2, \hat{J}_z, \hat{L}^2, \hat{S}^2$ to be obtained by forming linear combinations of the above product states. To this end, we will make use of the results proved earlier:

$$\hat{L}_\pm|l, m_l\rangle = \hbar\sqrt{(l \pm m_l + 1)(l \mp m_l)}|l, m_l \pm 1\rangle \tag{2.171}$$

and

$$\hat{J}^2 = \hat{L}^2 + \hat{S}^2 + 2\hat{L}_z\hat{S}_z + \hat{L}_+\hat{S}_- + \hat{L}_-\hat{S}_+ \tag{2.172}$$

Denote the eigenstates of $\hat{J}^2, \hat{J}_z, \hat{L}^2, \hat{S}^2$ by $|j, m_j, l\rangle$ (we could have added a fourth index $S = \frac{1}{2}$ but this is understood as we are concerned with electron-spin here). Presumably, j, the quantum number to $\hat{J}^2$ has the values:

$$j = l + \frac{1}{2} \ \ , \ \ l - \frac{1}{2} \tag{2.173}$$

(This would give the right number of states since $2(l + \frac{1}{2}) + 1 + 2(l - \frac{1}{2}) + 1 = 2(2l + 1)$). First consider the case $j = l + \frac{1}{2}$ and the eigenstate at the top of the ladder:

$$|j = l + \frac{1}{2}, m_j = l + \frac{1}{2}, l\rangle = |l, l\rangle \otimes |\uparrow\rangle \tag{2.174}$$

Applying $\hat{J}_z$ we have:

$$\begin{aligned} \hat{J}_z|l,l\rangle|\uparrow\rangle &= (\hat{L}_z + \hat{S}_z)|l,l\rangle \otimes |\uparrow\rangle \\ &= \hbar(l + \frac{1}{2})|l,l\rangle \otimes |\uparrow\rangle \\ &= \hbar m_j|l,l\rangle \otimes |\uparrow\rangle \end{aligned} \tag{2.175}$$

Applying $\hat{J}^2$ we have:

$$\begin{aligned} \hat{J}^2|l,l\rangle|\uparrow\rangle &= \hbar^2(l(l+1) + \frac{3}{4} + (2l)\frac{1}{2})|l,l\rangle \otimes |\uparrow\rangle \\ &= \hbar^2(l + \frac{1}{2})(l + \frac{3}{2})|l,l\rangle \otimes |\uparrow\rangle \\ &= \hbar^2 j(j+1)|l,l\rangle \otimes |\uparrow\rangle \end{aligned} \tag{2.176}$$

This shows that $|l,l\rangle|\uparrow\rangle$ is an eigenstate of $\hat{J}^2$ with the eigenvalues $j = l + \frac{1}{2}, m_j = l + \frac{1}{2}$. To obtain all other states just apply the lowering operator:

$$\hat{J}_- = \hat{L}_- + \hat{S}_- \tag{2.177}$$

repeatedly to $|l,l\rangle|\uparrow\rangle$.

Problem 21: Show that

$$\begin{aligned} \hat{J}_-|l,l\rangle|\uparrow\rangle &= \sqrt{(2l)}\hbar|l,l-1\rangle \otimes |\uparrow\rangle + \hbar|l,l\rangle \otimes |\downarrow\rangle \\ \hat{J}_-|j = l + \frac{1}{2}, m_j = l + \frac{1}{2}, l\rangle &= \sqrt{(2l+1)}|j = l + \frac{1}{2}, m_j = l - \frac{1}{2}, l\rangle \end{aligned} \tag{2.178}$$

Hence, we have:

$$|j = l + \frac{1}{2}, m_j = l - \frac{1}{2}, l\rangle = \sqrt{\frac{2l}{2l+1}}|l, l-1\rangle \otimes |\uparrow\rangle + \sqrt{\frac{1}{2l+1}}|l, l\rangle \otimes |\downarrow\rangle \tag{2.179}$$

By repeated application of $\hat{J}_-$ one obtains the general result (which can be verified by mathematical induction with respect to m_j):

$$|j = l + \frac{1}{2}, m_j, l\rangle = \sqrt{\frac{l + m_j + \frac{1}{2}}{2l+1}}|l, m_j - \frac{1}{2}\rangle \otimes |\uparrow\rangle + \sqrt{\frac{l - m_j + \frac{1}{2}}{2l+1}}|l, m_j + \frac{1}{2}\rangle \otimes |\downarrow\rangle \tag{2.180}$$

where m_j takes half integer values in the range $(l+\frac{1}{2}), \cdots, -(l+\frac{1}{2})$. The eigenstates corresponding to $j = l - \frac{1}{2}$ are orthogonal to all the states derived above and where $m_j = (l - \frac{1}{2}), \cdots, -(l - \frac{1}{2})$ are given by

$$|j = l - \frac{1}{2}, m_j, l\rangle = -\sqrt{\frac{l - m_j + \frac{1}{2}}{2l+1}}|l, m_j - \frac{1}{2}\rangle \otimes |\uparrow\rangle + \sqrt{\frac{l + m_j + \frac{1}{2}}{2l+1}}|l, m_j + \frac{1}{2}\rangle \otimes |\downarrow\rangle \tag{2.181}$$

(This can be verified using the same approach as above).

2.4.3 General case

Two arbitrary AM vectors, $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$ have AM eigenstates are given by $|j_1, m_1\rangle$ and $|j_2, m_2\rangle$, respectively. Form the total AM vector $\hat{\mathbf{J}} = \hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2$. We then have available two different complete sets of commuting AM operators: $\hat{J}_1^2, \hat{J}_{1z}, \hat{J}_2^2, \hat{J}_{2z}$, with eigenstates $|j_1, m_1\rangle|j_2, m_2\rangle = |j_1, m_1, j_2, m_2\rangle$, and $\hat{J}^2, \hat{J}_z, \hat{J}_1^2, \hat{J}_2^2$, with eigenstates $|j, m, j_1, j_2\rangle$. Both sets of eigenstates form a complete orthonormal system, which implies that they can be expanded in terms of each other. For example the expansion of $|j, m, j_1, j_2\rangle$ in the basis $\{|j_1, m_1, j_2, m_2\rangle\}$ is:

$$|j, m, j_1, j_2\rangle = \sum_{m = m_1 + m_2} |j_1, m_1, j_2, m_2\rangle\langle j_1, m_1, j_2, m_2|j, m, j_1, j_2\rangle \tag{2.182}$$

The (real) coefficients $\langle j_1, m_1, j_2, m_2|j, m, j_1, j_2\rangle$ are called **Clebsch-Gordan coefficients**. (Note that because the z-component of the AM is conserved, the only composite states that contribute to the above sum are those for which $m = m_1 + m_2$). In a commonly applied terminology, one refers to $|j, m, j_1, j_2\rangle$ as an eigenfunction in the *coupled* representation and to $|j_1, m_1, j_2, m_2\rangle$ as an eigenfunction in the the *uncoupled* representation. The allowed values that j can take (given j_1 and j_2) are

$$j = j_1 + j_2, j_1 + j_2 - 1, \ldots, |j_1 - j_2| \tag{2.183}$$

with the possible values of m being:

$$m = m_1 + m_2 = j, j-1, \ldots, -j \tag{2.184}$$

A general formula for the Clebsch-Gordan coefficients has been derived by Wigner (using group theory) and numerical tables of these coefficients can be found in books dealing with atomic spectroscopy.

Chapter 3

Approximation Methods For Bound States

3.1 Introduction

There are very few problems in QM that can be solved exactly. To use QM either to verify or predict the results of experiments, one needs to resort to approximation techniques such as

- Variational methods
- Perturbative methods
- The JWKB method

3.2 Variational methods

Variational methods (usually used to determine the ground state) involve using a parameterised trial wave function to represent the ground state wave function. The parameters are optimised to minimise the ground state energy. Consider a physical system whose Hamiltonian $\hat{H}$ is time independent. Assume that the entire spectrum of $\hat{H}$ is discrete and non-degenerate.

$$\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle \quad (n = 0, 1, 2, \ldots) \tag{3.1}$$

The energy spectrum is ordered as follows:

$$E_0 < E_1 < E_2 < \ldots \tag{3.2}$$

3.2.1 Variational theorem

If $|\Psi\rangle$ represents an arbitrary state of the system, then:

$$\langle \hat{H} \rangle = \frac{\langle \Psi|\hat{H}|\Psi\rangle}{\langle \Psi|\Psi\rangle} \geq E_0 \tag{3.3}$$

with the equality occurring if and only if $|\Psi\rangle$ is the ground state eigenvector of $\hat{H}$ with eigenvalue E_0 (the ground state energy).

Proof: Expand $|\Psi\rangle$ in the basis of the normalised eigenstates of $\hat{H}$:

$$|\Psi\rangle = \sum_n c_n |\psi_n\rangle \tag{3.4}$$

This implies:

$$\begin{aligned} \langle \Psi|\hat{H}|\Psi\rangle &= \sum_n \sum_m \langle \psi_m|\hat{H}|\psi_n\rangle \, c_m^* \, c_n \\ &= \sum_n |c_n|^2 E_n \end{aligned} \tag{3.5}$$

and

$$\begin{aligned} \langle \Psi|\Psi\rangle &= \sum_n \sum_m \langle \psi_m|\psi_n\rangle c_m^* \, c_n \\ &= \sum_n |c_n|^2 \end{aligned} \tag{3.6}$$

Therefore

$$\langle \hat{H} \rangle = \frac{\sum_n |c_n|^2 E_n}{\sum_n |c_n|^2} \tag{3.7}$$

Since $E_0 < E_1 < E_2 < \ldots$,

$$\sum_n |c_n|^2 E_n \geq E_0 \sum_n |c_n|^2 \tag{3.8}$$

so that

$$\langle \hat{H} \rangle \geq \frac{E_0 \sum_n |c_n|^2}{\sum_n |c_n|^2} = E_0 \tag{3.9}$$

The equality sign holds when $c_0 = 1$ and $c_n = 0 \; \forall \; n \neq 0$, i.e. when $|\Psi\rangle = |\psi_0\rangle$.

In actual applications to determine the ground state properties of bound systems, one first chooses a *trial* wave function $\psi_{trial}(\mathbf{r}, \{\alpha, \beta, \ldots\})$ which depends on the parameters $\alpha, \beta, \ldots$ etc. We then calculate the expectation value

$$E(\alpha, \beta, \ldots) = \frac{\langle \psi_{trial}|\hat{H}|\psi_{trial}\rangle}{\langle \psi_{trial}|\psi_{trial}\rangle} \tag{3.10}$$

which (from the variational theorem) is an upper bound to the ground state energy of the system. (The only restriction on $\psi_{trial}(\mathbf{r}, \{\alpha, \beta, \ldots\})$ is that it obeys the **same** boundary conditions as the eigenstates of the Hamiltonian, $\hat{H}$. Otherwise, Eq. (3.4) is not valid.) We of course choose trial wave functions that are appropriate for the problem at hand.

We then optimise the parameters $\alpha, \beta, \ldots$ by determining those values of $\alpha, \beta, \ldots$ which **minimise** $E(\alpha, \beta, \ldots)$ for that particular trial wave function. This means that we have to solve the following set of equations (linear or non-linear, depending on the functional form of $\psi_{trial}(\mathbf{r}, \{\alpha, \beta, \ldots\})$)

$$\begin{aligned} \frac{\partial}{\partial \alpha} E(\alpha, \beta, \ldots) &= 0 \\ \frac{\partial}{\partial \beta} E(\alpha, \beta, \ldots) &= 0 \qquad (3.11) \end{aligned}$$

etc. Suppose we denote the values of the parameters that minimise $E(\alpha, \beta, \ldots)$ by $(\bar{\alpha}, \bar{\beta}, \ldots)$. Then the minimum value, $E(\bar{\alpha}, \bar{\beta}, \ldots)$ constitutes an **upper bound** to the exact ground state energy, while the trial wave function for the optimal values, $\psi(\mathbf{r}, \{\bar{\alpha}, \bar{\beta}, \ldots\})$, is an approximation to the exact ground state wave function.

Problem 1: Consider

$$E = \langle \psi_{trial} | \hat{H} | \psi_{trial} \rangle$$

with an arbitrary normalised trial wave function, ψ_{trial}. Show that if ψ_{trial} differs from the correct ground state wave function, ψ_0, by terms of order ϵ, then E as defined above differs from the ground state energy by order ϵ^2.

Problem 2: Use the variational method to show that a one-dimensional attractive potential will always have a bound state. (Hint: Consider a square well that will fit inside the potential.)

Problem 3: Use the variational method to estimate the ground state energy for the anharmonic oscillator

$$\hat{H} = \frac{\hat{P}^2}{2m} + \lambda \hat{x}^4$$

Compare your result with the exact result

$$E_0 = 1.060 \lambda^{1/3} \left(\frac{\hbar^2}{2m} \right)^{2/3}$$

(Hint: Use a Gaussian trial wave function.)

Answer: Using

$$\psi_{trial}(x) = \sqrt{\frac{\beta}{\sqrt{\pi}}} e^{-\frac{1}{2}\beta^2 x^2}$$

the optimal value of $\beta = (6m\lambda/\hbar^2)^{\frac{1}{6}}$ giving $E_{min} = 1.083(\frac{\hbar^2}{2m})^{\frac{2}{3}}\lambda^{\frac{1}{3}}$.

Problem 4†: A particle of mass m (moving in three dimensions) is bound in the ground state of the exponential potential

$$V(r) = -\frac{4\hbar^2}{3ma^2}e^{-r/a}$$

Using

$$\psi_{trial}(r) = A\ e^{(-\beta r/2a)}$$

as a trial function with β as the variational parameter (A is determined by normalisation), obtain an upper bound for the ground state energy.

Answer: Optimal value is $\beta = 1$ and the upper bound for the ground state energy is $-\frac{\hbar^2}{24ma^2}$.

Problem 5†: Let E_1 and E_2 be the ground state energies of a particle of mass m moving in the attractive potentials $V_1(\mathbf{r})$ and $V_2(\mathbf{r})$ respectively. If $V_1(\mathbf{r}) \leq V_2(\mathbf{r})$ for all $\mathbf{r}$ one intuitively expects $E_1 \leq E_2$. Use a variational argument to derive this result.

3.2.2 Interlude : atomic units

In this section we deal with a system of units called *atomic units* that is widely used in atomic, molecular and solid-state physics. What it boils down to is a set of units such that the following fundamental quantities, $\hbar, m, |e|, 4\pi\epsilon_0$ are set to 1. Here m is the electron mass, $|e|$ the magnitude of its charge, and ϵ_0 the permittivity of free space. What this means is that the Hamiltonian for a hydrogen-like atom reduces from

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{4\pi\epsilon_0 r}$$

to

$$\hat{H} = -\frac{1}{2}\nabla^2 - \frac{Z}{r}$$

which is independent of any physical constants. It turns out that other physical quantities can be expressed in terms of the above four basic units (i.e. they are derived units). Most important of these are the atomic unit of length, called the *bohr*, defined by

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{me^2} = 1 \quad \text{(au)}$$

where a_0 is the Bohr radius of the hydrogen atom in its ground state. (Note that "au" is an abbreviation for *atomic units*). In the same way, a more "natural" unit for energy is the *hartree*, defined by

$$1\ \text{hartree} = \frac{me^4}{16\pi^2\epsilon_0^2\hbar^2} = 1 \quad \text{(au)}$$

such that the ground state energy of the hydrogen atom is $-\frac{1}{2}$ hartree.

Problem 6†: Given the following SI equivalent quantities

$$m = 9.1091 \times 10^{-31}\ \text{kg} \qquad |e| = 1.6021 \times 10^{-19}\ \text{C}$$

$$\hbar = 1.0545 \times 10^{-34}\ \text{Js} \qquad 4\pi\epsilon_0 = 1.1126 \times 10^{-10}\ \text{C}^2\text{J}^{-1}\text{m}^{-1}$$

show that

(a) The atomic unit of length, the *bohr* is 5.29167×10^{-11} m = 0.529167 Å.

(b) The atomic unit of energy, the *hartree* is 4.35944×10^{-18} J = 27.211 eV.

(c) Express one hartree in units of kilojoules per mole and wavenumbers. *Answer:* 2625 kJ mol^{-1}, $\nu = \frac{1}{\lambda} = 2.195 \times 10^5\ \text{cm}^{-1}$. Note that wavenumbers are used commonly in spectroscopy.

(d) Show that the speed of light in atomic units is approximately equal to 137 au.

3.2.3 Hydrogen molecular ion, H_2^+

We now apply the variational method to a "real" problem. The simplest molecule is the H_2^+ ion. It so happens that it is possible to solve the Schrödinger equation for this molecule exactly. (Within the Born-Oppenheimer approximation, the Schrödinger equation for an electron in the field of the two fixed protons is separable in elliptic coordinates.) Unfortunately, the solutions are quite complicated and not easy to use. Instead, we will "solve" this problem approximately via the variational method, where trial wavefunctions are chosen and the energy is minimised with respect to some variational parameter (which in this case is the inter-proton distance). The Hamiltonian for the problem, in atomic units, is

$$\hat{H} = -\frac{1}{2}\nabla^2 - \frac{1}{|\mathbf{r} - \mathbf{R_1}|} - \frac{1}{|\mathbf{r} - \mathbf{R_2}|} + \frac{1}{|\mathbf{R_1} - \mathbf{R_2}|} \tag{3.12}$$

The first term represents the kinetic energy of the electron, the second and third terms describe the attractive Coulomb interaction between the electron and protons 1 and 2 with position vectors $\mathbf{R_1}$ and $\mathbf{R_2}$ respectively. The last term is the repulsive interaction between the two protons. The vector $\mathbf{r}$ denotes the position of the electron and we define the relative vector $\mathbf{R} = \mathbf{R_1} - \mathbf{R_2}$. Its magnitude R is the distance between the two protons. For the trial wavefunction we use a linear combination of $1s$ hydrogen atom wavefunctions localised at the first (ψ_1) and second (ψ_2) proton. We have

$$\psi(\mathbf{r}) = \alpha\psi_1(\mathbf{r}) + \beta\psi_2(\mathbf{r}) \tag{3.13}$$

where

$$\psi_1 = \frac{1}{\sqrt{\pi}} e^{-|\mathbf{r}-\mathbf{R_1}|} \tag{3.14}$$

$$\psi_2 = \frac{1}{\sqrt{\pi}} e^{-|\mathbf{r}-\mathbf{R_2}|} \tag{3.15}$$

Note that there is symmetry under reflection with respect to a plane lying exactly between the two protons. The wavefunction therefore has the same symmetry and can be classified by its parity. For positive parity, $\alpha = \beta$ while for negative parity, $\alpha = -\beta$. This means that

$$\psi_+(\mathbf{r}) = N_+\left[\psi_1(\mathbf{r}) + \psi_2(\mathbf{r})\right] \tag{3.16}$$

$$\psi_-(\mathbf{r}) = N_-\left[\psi_1(\mathbf{r}) - \psi_2(\mathbf{r})\right] \tag{3.17}$$

where $N_\pm$ are normalisation constants which are determined by the condition $\langle\psi_\pm|\psi_\pm\rangle = 1$, i.e.

$$N_+^2[2 + 2\int d\mathbf{r}\,\psi_1^*(\mathbf{r})\psi_2(\mathbf{r})] = 1 \tag{3.18}$$

$$N_-^2[2 - 2\int \psi_1^*(\mathbf{r})\psi_2(\mathbf{r})] = 1 \tag{3.19}$$

By defining the *overlap integral* $U(R)$,

$$U(R) = \int d\mathbf{r}\,\psi_1^*(\mathbf{r})\psi_2(\mathbf{r}) \tag{3.20}$$

where $R = |\mathbf{R_1} - \mathbf{R_2}|$ is the inter-proton distance, the above equation reduces to

$$N_\pm = \frac{1}{\sqrt{2[1 \pm U(R)]}} \tag{3.21}$$

Problem 7: Show that

$$U(R) = \left(1 + R + \frac{1}{3}R^2\right)e^{-R} \tag{3.22}$$

Hint: Starting from

$$U(R) = \frac{1}{\pi}\int d\mathbf{r}\; e^{-|\mathbf{r}-\mathbf{R_1}|-|\mathbf{r}-\mathbf{R_2}|}$$

make a transformation to relative $\mathbf{R} = \mathbf{R_1} - \mathbf{R_2}$ and center-of-mass coordinates, $\mathbf{R_{cm}} = \frac{1}{2}(\mathbf{R_1} + \mathbf{R_2})$, and making use of the obvious translational invariance to get

$$U(R) = \frac{1}{\pi}\int d\mathbf{r}\; e^{-|\mathbf{r}-\frac{1}{2}\mathbf{R}|-|\mathbf{r}+\frac{1}{2}\mathbf{R}|}$$

The azimuthal symmetry of the integrand about the axis in the direction of $\mathbf{R}$ means we can use *elliptical coordinates* to carry out the integral. The transformation equations are given by

$$x = \frac{R}{2}\left[(\xi^2-1)(1-\eta^2)\right]^{\frac{1}{2}}\cos\phi \qquad y = \frac{R}{2}\left[(\xi^2-1)(1-\eta^2)\right]^{\frac{1}{2}}\sin\phi \qquad z = \frac{R}{2}\xi\eta$$

where $\xi \in [1,\infty), \quad \eta \in [-1,1], \quad \phi \in [0,2\pi)$. We then have

$$|\mathbf{r} \pm \frac{1}{2}\mathbf{R}| = \frac{\mathbf{R}}{\mathbf{2}}(\xi \pm \eta)$$

which gives

$$U(R) = \frac{1}{\pi}\frac{R^3}{8}\int_0^{2\pi} d\phi \int_{-1}^{1} d\eta \int_1^{\infty} d\xi \quad (\xi^2-\eta^2) \quad e^{-R\xi}$$

Carrying out the integrals gives the required result.

The expectation value of the Hamiltonian with respect to $\psi_\pm$ is given by

$$\begin{aligned}\epsilon_+(R) &= \frac{\langle\psi_1|\hat{H}|\psi_1\rangle + \langle\psi_2|\hat{H}|\psi_2\rangle + 2\langle\psi_1|\hat{H}|\psi_2\rangle}{2(1+U)} \\ \epsilon_-(R) &= \frac{\langle\psi_1|\hat{H}|\psi_1\rangle + \langle\psi_2|\hat{H}|\psi_2\rangle - 2\langle\psi_1|\hat{H}|\psi_2\rangle}{2(1-U)} \end{aligned} \tag{3.23}$$

By symmetry we have $\langle\psi_1|\hat{H}|\psi_1\rangle = \langle\psi_2|\hat{H}|\psi_2\rangle$ so the above equations reduce to

$$\epsilon_\pm(R) = \frac{\langle\psi_1|\hat{H}|\psi_1\rangle \pm \langle\psi_1|\hat{H}|\psi_2\rangle}{(1 \pm U)} \tag{3.24}$$

Problem 8: Show that

$$\begin{aligned}\langle\psi_1|\hat{H}|\psi_1\rangle &= \int d\mathbf{r}\; \psi_1^*(\mathbf{r})\hat{H}\psi_1(\mathbf{r}) \\ &= \epsilon_1^H + \frac{1}{R} - \int d\mathbf{r}\; \frac{|\psi_1(\mathbf{r})|^2}{|\mathbf{r}-\mathbf{R_2}|} \\ &= \epsilon_1^H + \frac{1}{R}\left(1+R\right)e^{-2R} \end{aligned} \tag{3.25}$$

$$\langle\psi_1|\hat{H}|\psi_2\rangle = \int d\mathbf{r}\; \psi_1^*(\mathbf{r})\hat{H}\psi_2(\mathbf{r})$$

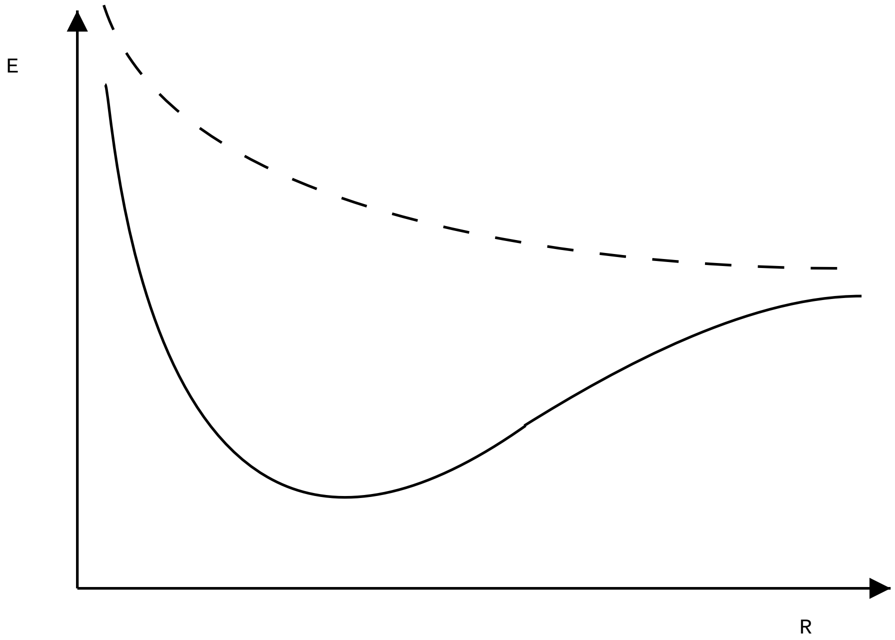
$$= \left(\epsilon_1^H + \frac{1}{R}\right)U(R) - \int d\mathbf{r}\ \frac{\psi_1^*(\mathbf{r})\psi_2(\mathbf{r})}{|\mathbf{r} - \mathbf{R_2}|}$$

$$= \left(\epsilon_1^H + \frac{1}{R}\right)U(R) + (1+R)e^{-R} \tag{3.26}$$

where ϵ_1^H is the energy of the 1s state of the hydrogen atom (-0.5 hartree).

Hint: Use elliptical coordinates as in the previous problem to carry out the integrations.

Putting all the various integrals together, we determine $\epsilon_\pm(R)$ as a function of R. This is shown in the sketch below.



Note that the state with positive parity (solid line) is bound since $\epsilon_+(R)$ has a minimum while the state with negative parity (dashed) has no minimum, i.e. that state is never bound. Experimentally the binding energy is -2.8 eV with $R = 1.06$ Åwhile our calculations gives -1.76 eV at the variational minimum value of $R = 1.3$ Å, which is not bad for such a crude trial function!

3.2.4 Generalisation: Ritz theorem

The variational theorem is generalised as follows:

Theorem: The expectation value of the Hamiltonian is **stationary** in the neighbourhood of the discrete eigenvalues.

Proof: By stationary, we mean that the change in the value of $\langle \hat{H} \rangle$ when the state vector is changed by an infinitesimal amount is zero (to first order in the change in the state vector). We need to show that each stationary expectation value is an eigenvalue of $\hat{H}$. Let

$$\langle \hat{H} \rangle = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \tag{3.27}$$

Note that this is a "functional" of the state vector $|\Psi\rangle$. Consider an infinitesimally small change to $|\Psi\rangle$:

$$|\Psi\rangle \to |\Psi\rangle + |\delta\Psi\rangle \tag{3.28}$$

We need to determine the corresponding change to $\langle \hat{H} \rangle$. From (3.27) we have

$$\langle \Psi | \Psi \rangle \langle \hat{H} \rangle = \langle \Psi | \hat{H} | \Psi \rangle \tag{3.29}$$

Inserting (3.28) into the above equation we have:

$$(\langle \Psi | + \langle \delta\Psi |)(|\Psi\rangle + |\delta\Psi\rangle)[\langle \hat{H} \rangle + \delta\langle \hat{H} \rangle] = (\langle \Psi | + \langle \delta\Psi |)\hat{H}(|\Psi\rangle + |\delta\Psi\rangle) \tag{3.30}$$

which when expanded out gives:

$$\begin{aligned} &(\langle \Psi | \delta\Psi \rangle + \langle \Psi | \Psi \rangle + \langle \delta\Psi | \Psi \rangle + \langle \delta\Psi | \delta\Psi \rangle)[\langle \hat{H} \rangle + \delta\langle \hat{H} \rangle] \\ &= \langle \Psi | \hat{H} | \delta\Psi \rangle + \langle \Psi | \hat{H} | \Psi \rangle + \langle \delta\Psi | \hat{H} | \delta\Psi \rangle + \langle \delta\Psi | \hat{H} | \Psi \rangle \end{aligned} \tag{3.31}$$

Using (3.28), dropping all terms of second order, after some algebra we finally get:

$$\delta\langle \hat{H} \rangle \langle \Psi | \Psi \rangle = \langle \Psi | (\hat{H} - \langle \hat{H} \rangle) | \delta\Psi \rangle + \langle \delta\Psi | (\hat{H} - \langle \hat{H} \rangle) | \Psi \rangle \tag{3.32}$$

Thus $\langle \hat{H} \rangle$ is stationary, i.e. $\delta\langle \hat{H} \rangle = 0$, provided the right hand side of (3.32) is zero:

$$\langle \Psi | (\hat{H} - \langle \hat{H} \rangle) | \delta\Psi \rangle + \langle \delta\Psi | (\hat{H} - \langle \hat{H} \rangle) | \Psi \rangle = 0 \tag{3.33}$$

Suppose we define $|\Phi\rangle = (\hat{H} - \langle \hat{H} \rangle)|\Psi\rangle$. Then (3.33) becomes:

$$\langle \Phi | \delta\Psi \rangle + \langle \delta\Psi | \Phi \rangle = 0 \tag{3.34}$$

This must be satisfied for **any** $|\delta\Psi\rangle$. Let us choose a $|\delta\Psi\rangle$ such that:

$$|\delta\Psi\rangle = \delta\lambda |\Phi\rangle \tag{3.35}$$

where $\delta\lambda$ is a number of first order in small quantities. This implies that (3.34) becomes:

$$2\delta\lambda \langle \Phi | \Phi \rangle = 0 \tag{3.36}$$

i.e. that the *norm* of $|\Phi\rangle$ equals zero. Therefore, $|\Phi\rangle$ must be zero, i.e.

$$(\hat{H} - \langle\hat{H}\rangle)|\Psi\rangle = 0 \tag{3.37}$$

This implies that:

$$\hat{H}|\Psi\rangle = \langle\hat{H}\rangle|\Psi\rangle \tag{3.38}$$

Therefore the right hand side of (3.27) is stationary if and only if the state vector $|\Psi\rangle$ corresponds to an eigenvector of $\hat{H}$, and the stationary values correspond to the eigenvalues of $\hat{H}$.

3.2.5 Linear variation functions

Representations of eigenstates for all except the simplest systems are complicated functions. In practice, we expand the arbitrary eigenstate $|\Psi\rangle$ as a sum of a finite number (N) of functions (whose functional form are chosen depending on the type of system under study) so that:

$$|\Psi\rangle = \sum_{i=1}^{N} c_i \; |i\rangle \tag{3.39}$$

and these functions are assumed linearly independent (but not necessarily mutually orthogonal). For example $|i\rangle$ can be plane waves, or Gaussian functions or a mixture of both, etc. Here c_i are complex numbers that are to be determined. The optimal choice for these linear coefficients, from the variational theorem, are those that make $\hat{H}$ stationary. We have

$$\langle\Psi|\hat{H}|\Psi\rangle - E\langle\Psi|\Psi\rangle = 0 \tag{3.40}$$

(We set $E = \langle\hat{H}\rangle$). Substituting (3.39) into (3.40) yields:

$$\sum_{i=1}^{N}\sum_{j=1}^{N} c_j^* \; c_i \; H_{ji} - E\sum_{i=1}^{N}\sum_{j=1}^{N} c_j^* \; c_i \; S_{ji} = 0 \tag{3.41}$$

where

$$\begin{aligned} H_{ji} &= \langle j|\hat{H}|i\rangle \\ S_{ji} &= \langle j|i\rangle \end{aligned} \tag{3.42}$$

Differentiating (3.41) with respect to c_l^* gives:

$$\sum_{i=1}^{N}(H_{li} - E\; S_{li}) \; c_i = 0 \tag{3.43}$$

Cramer's rule tells us that all the c_i's are zero unless the determinant of the coefficients vanishes:

$$|\mathbf{H} - E\,\mathbf{S}| = 0 \tag{3.44}$$

Here $\mathbf{H}$ is the $N \times N$ matrix whose coefficients are H_{ji} defined above, i.e. it is the matrix representation of the Hamiltonian in the basis $\{|i\rangle\}$. The $N \times N$ matrix $\mathbf{S}$ whose coefficients are S_{ji} is called the **overlap** matrix. Equation (3.44) is called the **secular equation** for the energy E and is an N^{th} order polynomial in E. This yields N real roots, some of which may be degenerate. Arranging these roots in order of increasing value as

$$\tilde{E}_0 \leq \tilde{E}_1 \leq \ldots \tilde{E}_{N-1}$$

we can compare them with the *exact* spectrum of the system (in order of increasing energy)

$$E_0 \leq E_1 \leq \ldots \leq E_{N-1} \leq E_N \leq \ldots$$

From the variation theorem, we know that

$$E_0 \leq \tilde{E}_0 \tag{3.45}$$

Moreover, it can be proved[1] that

$$E_1 \leq \tilde{E}_1, \quad E_2 \leq \tilde{E}_2, \ \ldots, \ E_{N-1} \leq \tilde{E}_{N-1} \tag{3.46}$$

Thus the linear variation method provides upper bounds to the energies of the lowest N eigenstates of the system. The roots of Eq. (3.44) are used as approximations to the energies of the lowest eigenstates. Increasing the value of N in (3.39) (which corresponds to increasing the number of functions to represent the eigenstates) can be shown to increase (or at worst cause no change in) the accuracy of the previously calculated energies. If the set of functions $\{|i\rangle\}$ form a complete set, then we will obtain the exact wave functions of the system. Unfortunately, to have a complete set, we usually need an infinite number of expansion functions!

To obtain an approximation to the ground state wave function, we take the lowest root $\tilde{E}_0$ of the secular equation and substitute it into the set of equations (3.43); we then solve this set of equations for the coefficients $c_1^0, \ c_2^0, \ldots, c_N^0$, where the superscript is added to indicate that these coefficients correspond to $\tilde{E}_0$. (As Eq. (3.43) constitutes a set of linear, homogeneous equations, we can only determine ratios of coefficients; we solve the $c_1^0, \ c_2^0, \ldots, c_N^0$ in terms of c_1^0, and then determine c_1^0 by normalisation).

Having found the c_i^0's, we take

$$|\Psi\rangle = \sum_{i=1}^{N} c_i^0 \ |i\rangle$$

---

[1] J.K. MacDonald, Phys. Rev. **43**, p.830 (1933); R.H. Young, Int. J. Quantum Chem. **6**, p.596 (1972).

as an approximate ground state wave function. Use of the higher roots of Eq. (3.39) in Eq. (3.43) gives approximations to the excited-state wave functions (which can be shown to be mutually orthogonal.) This approach forms the basis for most electronic structure calculations in physics and chemistry to determine the electronic structure of atoms, molecules, solids, surfaces, etc.

Problem 10†: Let $V(x) = 0$ for $-1 \leq x \leq +1$ and ∞ otherwise (the "particle in a box" problem). Use

$$\begin{aligned} f_1(x) &= (1 - x^2) \\ f_2(x) &= (1 - x^4) \end{aligned}$$

to construct the trial function

$$|\Psi\rangle = \sum_{i=1}^{2} c_i\ f_i(x)$$

Find the approximate energies and wave functions for the lowest two states and compare your results with the exact solutions to the problem.

Solution: First construct the overlap ($\mathbf{S}$) and Hamiltonian matrix ($\mathbf{H}$), in the above $f_1,\ f_2$ basis. The overlap matrix elements are:

$$S_{11} = \frac{16}{15}, \quad S_{22} = \frac{64}{45}, \quad S_{12} = S_{21} = \frac{128}{105}$$

and the hamiltonian matrix elements are:

$$H_{11} = \frac{4\hbar^2}{3m}, \quad H_{22} = \frac{16\hbar^2}{7m}, \quad H_{12} = H_{21} = \frac{8\hbar^2}{5m}$$

Now solve the generalised eigenvalue problem to get the eigenvalues $1.23\frac{\hbar^2}{m}$ and $12.77\frac{\hbar^2}{m}$.

When we compare this with the exact eigenvalues for the first three states, $\approx 1.23\frac{\hbar^2}{m}, \quad 4.93\frac{\hbar^2}{m}$ and $11.10\frac{\hbar^2}{m}$, we find that our calculation gives the upper bounds to the first and third state. (Note that the estimate for the ground state is very close but NOT equal to the exact value.) Since the basis functions are even functions it is not suprising that we do not get an estimate for the second state as this is a odd function!

Problem 10: Consider the one-dimensional infinite well of length L. The Hamiltonian for the system is $\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)$ where $V(x) = 0$ for $x \in [0, L]$ and ∞ otherwise. Find the approximate energies and wave functions for the lowest four states and compare your results with the exact solutions to the problem. The linear variation function is given by

$$|\Psi\rangle = \sum_{i=1}^{4} c_i\ f_i(x)$$

with

$$\begin{aligned} f_1(x) &= x\,(L-x) \\ f_2(x) &= x^2\,(L-x)^2 \\ f_3(x) &= x\,(L-x)\,(\frac{1}{2}L-x) \\ f_4(x) &= x^2\,(L-x)^2\,(\frac{1}{2}L-x) \end{aligned}$$

Answer: You must first note that f_1, f_2 are even functions while f_3, f_4 are odd. This simplifies evaulation of integrals when we determine the overlap and hamiltonian matrix elements. We get:

$$S_{13} = S_{31} = S_{14} = S_{41} = S_{23} = S_{32} = S_{24} = S_{42} = 0$$

and the same for the corresponding hamiltonian matrix elements. The other non-zero matrix elements are given by:

$$S_{11} = \frac{L^5}{30},\;\; S_{22} = \frac{L^9}{630},\;\; S_{12} = S_{21} = \frac{L^7}{140},\;\; S_{33} = \frac{L^7}{840},\;\; S_{44} = \frac{L^{11}}{27720},\;\; S_{34} = S_{43} = \frac{L^9}{5040}$$

and

$$H_{11} = \frac{L^3}{6}\;\; H_{22} = \frac{L^7}{105},\;\; H_{12} = H_{21} = \frac{L^5}{30},\;\; H_{33} = \frac{L^5}{40},\;\; H_{44} = \frac{L^9}{1260},\;\; H_{34} = H_{43} = \frac{L^7}{280}$$

where the Hamiltonian matrix elements are in units of $\frac{\hbar^2}{m}$.

The secular determinant reduces to a block diagonal form so that instead of having to evaluate a 4 × 4 determinant, we have two 2 × 2 determinants to work out, which is much easier! The eigenvalues (in units of $\frac{\hbar^2}{mL^2}$) are ≈ 0.125, 0.500, 1.293 and 2.539. The *exact* eigenvalues are 0.125, 0.500, 1.125, and 2.000, for the first four states. Note that now we have estimates of all four states as we have included even and odd functions in our basis set.

3.3 Perturbation methods

3.3.1 Time-independent perturbation theory

Consider a time-independent Hamiltonian, $\hat{H}$ (whose spectrum is $\{E_n\}$) where

$$\hat{H} = \hat{H}_0 + \lambda\,\hat{W} \qquad (3.47)$$

in which effects on the system due to $\hat{H}_0$ (the **unperturbed** Hamiltonian whose spectrum is $E_n^{(0)}$) are the dominant ones while those due to $\lambda\hat{W}$ (the **perturbation**) are relatively weaker. The real number $\lambda \in [0,1]$ parameterises the strength of the perturbation. The perturbation will shift energy levels and/or split degeneracies. We will distinguish between non-degenerate and degenerate cases. For the non-degenerate case we shall assume

1. that both $\hat{H}$ and $\hat{H}_0$ have discrete spectra so the corresponding eigenstates $\{|\psi_n\rangle\}$ and $\{|\psi_n^{(0)}\rangle\}$ form complete, orthonormal sets.
2. there is a one-to-one correspondence between the eigenvalues of $\hat{H}_0$ and $\hat{H}$ (i.e. quantum states do not appear or disappear when the perturbation is switched off).
3. the perturbation is "weak" (i.e. it does not alter the eigenstates very much).

The aim of perturbation theory is to estimate the corrections that need to be made to $E_n^{(0)}$ and $|\psi_n^{(0)}\rangle$ to give the required E_n and $|\psi_n\rangle$ where:

$$\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle \tag{3.48}$$

Note that this can be an iterative process (i.e. the approximate solutions obtained can be fed back into the process to obtain a better approximation).

Assume that the eigenvalues and eigenfunctions of $\hat{H}$ can be expanded in a power series in λ:

$$\begin{aligned} E_n &= E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots \\ |\psi_n\rangle &= |\psi_n^{(0)}\rangle + \lambda|\psi_n^{(1)}\rangle + \lambda^2|\psi_n^{(2)}\rangle + \cdots \end{aligned} \tag{3.49}$$

Note that:

- This series is not guaranteed to converge. However, in many cases it is an *asymptotic expansion* i.e. the first few terms nevertheless give reliable results.

An asymptotic expansion of a function $f(\lambda)$,

$$f(\lambda) = \sum_{k=0}^{m} a_k \lambda^k + R_m(\lambda) \tag{3.50}$$

is characterised by the following behaviour of the remainder:

$$\begin{aligned} \lim_{\lambda\to 0} \frac{R_m(\lambda)}{\lambda^m} &= 0 \\ \lim_{m\to\infty} R_m(\lambda) &= \infty \end{aligned} \tag{3.51}$$

- There exist cases where E_n and ψ_n are NOT expandable in λ. For example, the bound states of a potential cannot be obtained from the continuum states via PT.

Non-degenerate perturbation theory

Substituting (3.49) into (3.48) we get:

$$\begin{aligned}(\hat{H}_0+\lambda\hat{W})(|\psi_n^{(0)}\rangle+\lambda|\psi_n^{(1)}\rangle+\lambda^2|\psi_n^{(2)}\rangle+\cdots) &= (E_n^{(0)}+\lambda E_n^{(1)}+\lambda^2 E_n^{(2)}+\cdots)(|\psi_n^{(0)}\rangle \\ &+ \lambda|\psi_n^{(1)}\rangle+\lambda^2|\psi_n^{(2)}\rangle+\cdots) \qquad (3.52)\end{aligned}$$

Comparing coefficients of $\lambda^0, \lambda^1, \lambda^2, \ldots$ we have:

0^{th} Order term

$$\hat{H}_0|\psi_n^{(0)}\rangle = E_n^{(0)}|\psi_n^{(0)}\rangle \qquad (3.53)$$

1^{st} Order term

$$\hat{H}_0|\psi_n^{(1)}\rangle+\hat{W}|\psi_n^{(0)}\rangle = E_n^{(0)}|\psi_n^{(1)}\rangle+E_n^{(1)}|\psi_n^{(0)}\rangle \qquad (3.54)$$

2^{nd} Order term

$$\hat{H}_0|\psi_n^{(2)}\rangle+\hat{W}|\psi_n^{(1)}\rangle = E_n^{(0)}|\psi_n^{(2)}\rangle+E_n^{(1)}|\psi_n^{(1)}\rangle+E_n^{(2)}|\psi_n^{(0)}\rangle \qquad (3.55)$$

Problem 11: Verify Eqs. (3.53) - (3.55).

Without loss of generality, we fix the normalisation of $|\psi_n\rangle$ by

$$\langle\psi_n^{(0)}|\psi_n\rangle = 1 \qquad (3.56)$$

This implies (for any λ)

$$\lambda\langle\psi_n^{(0)}|\psi_n^{(1)}\rangle+\lambda^2\langle\psi_n^{(0)}|\psi_n^{(2)}\rangle+\ldots=0 \qquad (3.57)$$

which means that

$$\langle\psi_n^{(0)}|\psi_n^{(1)}\rangle=\langle\psi_n^{(0)}|\psi_n^{(2)}\rangle=\ldots=0 \qquad (3.58)$$

If we multiply (3.54) by $\langle\psi_n^{(0)}|$ and make use of (3.53) we end up with

$$E_n^{(1)}=\langle\psi_n^{(0)}|\hat{W}|\psi_n^{(0)}\rangle \qquad (3.59)$$

which is the 1^{st} order correction to the energy of the n^{th} eigenstate. Using the fact that $\{|\psi_m^{(0)}\rangle\}$ form a complete orthonormal set, and making use of the result (3.58) we find:

$$|\psi_n^{(1)}\rangle = \sum_{m\neq n} c_m |\psi_m^{(0)}\rangle \tag{3.60}$$

with

$$c_m = \langle\psi_m^{(0)}|\psi_n^{(1)}\rangle \tag{3.61}$$

Multiplying (3.55) by $\langle\psi_m^{(0)}|$ (which is different from $\langle\psi_n^{(0)}|$) we find

$$c_m(E_n^{(0)} - E_m^{(0)}) = \langle\psi_m^{(0)}|\hat{W}|\psi_n^{(0)}\rangle \tag{3.62}$$

giving the 1^{st} order correction to the state $|\psi_n^{(0)}\rangle$ as:

$$|\psi_n^{(1)}\rangle = \sum_{m\neq n} \frac{\langle\psi_m^{(0)}|\hat{W}|\psi_n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}} |\psi_m^{(0)}\rangle \tag{3.63}$$

Problem 12: Prove that the 2^{nd} order correction to the energy is given by:

$$\begin{aligned} E_n^{(2)} &= \langle\psi_m^{(0)}|\hat{W}|\psi_n^{(1)}\rangle \\ &= \sum_{m\neq n} \frac{|\langle\psi_m^{(0)}|\hat{W}|\psi_n^{(0)}\rangle|^2}{E_n^{(0)} - E_m^{(0)}} \end{aligned} \tag{3.64}$$

Important points to note are:

- For the ground state, the 2^{nd} order shift in energy, $E_0^{(2)}$ is negative.
- If the matrix elements of $\hat{W}$ are of comparable magnitude, neighbouring levels make a larger contribution than distant levels, in 2^{nd} order PT.
- When the perturbation is switched on, the energy levels spread apart as if they repel one another.

Problem 13: The following Hamiltonian matrix has been constructed using an orthonormal basis

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix} + \begin{pmatrix} 0 & \mathrm{c} & 0 \\ \mathrm{c} & 0 & 0 \\ 0 & 0 & \mathrm{c} \end{pmatrix}$$

Here $\mathbf{H} = \mathbf{H}_0 + \mathbf{W}$ and c is a constant.

(a) Find the exact eigenvalues of **H**.

Answer: $E_1 = 2 - \sqrt{1+c^2}, \quad E_2 = 2 + \sqrt{1+c^2}, \quad E_3 = c - 2.$

(b) Use perturbation theory to determine the eigenvalues correct to second order in c.

Answer: The energy eigenvalues, up to second order are $1 - \frac{1}{2}c^2, \quad 3 + \frac{1}{2}c^2, \quad c - 2.$

(c) Compare the results of steps (a) and (b).

Problem 14: A one-dimensional particle-in-a-box has $V(x) = 0$ for $0 \leq x \leq a$ and ∞ otherwise. It is perturbed by the potential

$$W(x) = -\lambda \sin(\pi x/a)$$

for $0 \leq x \leq a$ and 0 otherwise. Calculate the approximate ground state energy using first-order perturbation theory.

Answer: $\frac{h^2}{8ma^2} - \frac{8\lambda}{3\pi}$.

Problem 15†: For the same particle-in-a-box, consider a perturbation given by

$$W(x) = \begin{cases} -\frac{2bx}{a} + b, & 0 \leq x \leq a/2 \\ \frac{2bx}{a} - b, & a/2 \leq x \leq a \end{cases}$$

Determine the eigenvalues of the perturbed system to first order assuming that $b \ll E_1^{(0)}$ where $E_1^{(0)}$ is the ground-state energy of the unperturbed system.

Answer: For n odd, we have $E_n^{(1)} = \frac{1}{2}b - \frac{2b}{(n\pi)^2}$ while for n even, we have $E_n^{(1)} = \frac{1}{2}b$.

Problem 16: Evaluate the first and second order corrections to the energy levels of the "slightly" anharmonic oscillator with potential energy

$$V(x) = \frac{1}{2}m\omega^2\hat{x}^2 + \alpha\hat{x}^3 + \beta\hat{x}^4$$

where the perturbation is

$$W(x) = \alpha\hat{x}^3 + \beta\hat{x}^4$$

(Hint: Use the results of Chapter 1 Problem 17.)

Answer:

$$\begin{aligned} E_n &= \left(n + \frac{1}{2}\right)\hbar\omega + \frac{3}{4}\beta(2n^2 + 2n + 1)\left(\frac{\hbar}{m\omega}\right)^2 - \frac{1}{8}\frac{\alpha^2}{\hbar\omega}(30n^2 + 30n + 11)\left(\frac{\hbar}{m\omega}\right)^3 \\ &- \frac{1}{8}\frac{\beta^2}{\hbar\omega}(34n^3 + 51n^2 + 59n + 21)\left(\frac{\hbar}{m\omega}\right)^4 \end{aligned}$$

Perturbation theory for degenerate states

The formulae derived above break down when there is degeneracy in the energy levels of the unperturbed Hamiltonian, $\hat{H}_0$, i.e. when there are linearly independent eigenfunctions corresponding to the same eigenvalue. When this happens, some of the denominators in the 1^{st} order correction to the eigenfunction or the 2^{nd} order correction to the energy eigenvalues vanish, leading to divergences. Below we work out a scheme to fix this.

Suppose the eigenvalue $E_n^{(0)}$ is $g-$fold degenerate, i.e. there are g linearly independent eigenfunctions, $\{|\phi_{n\alpha}\rangle\}$, $\alpha = 1, \ldots, g$ belonging to this eigenvalue. The $\{|\phi_{n\alpha}\rangle\}$ can be chosen to be orthonormal.

$$\begin{aligned} \hat{H}_0|\phi_{n\alpha}\rangle &= E_n^{(0)}|\phi_{n\alpha}\rangle \\ \langle\phi_{n\alpha}|\phi_{n\beta}\rangle &= \delta_{\alpha\beta} \qquad (\alpha, \beta = 1, \ldots, g) \end{aligned} \tag{3.65}$$

Let

$$|\psi_{ni}^{(0)}\rangle = \sum_{\alpha=1}^{g} c_{i\alpha}|\phi_{n\alpha}\rangle \qquad (i = 1, \ldots, g) \tag{3.66}$$

where the $c_{i\alpha}$ are constants to be determined.

Then we want to find

$$\begin{aligned} E_{ni} &= E_n^{(0)} + \lambda E_{ni}^{(1)} + \lambda^2 E_{ni}^{(2)} + \cdots \\ |\psi_{ni}\rangle &= |\psi_{ni}^{(0)}\rangle + \lambda|\psi_{ni}^{(1)}\rangle + \lambda^2|\psi_{ni}^{(2)}\rangle + \cdots \qquad (i = 1, \ldots, g) \end{aligned} \tag{3.67}$$

such that

$$\hat{H}|\psi_{ni}\rangle = (\hat{H}_0 + \lambda\hat{W})|\psi_{ni}\rangle = E_{ni}|\psi_{ni}\rangle \tag{3.68}$$

In general, the states $|\psi_{ni}\rangle$, $i = 1, \ldots, g$ will have different energies E_{ni}, i.e. the effect of the perturbation $\hat{W}$ is to split the $g-$fold degenerate level $E_{ni}^{(0)}$ into several groups of levels (at most g of them). Note however that it is possible that $\hat{W}$ only shifts the levels without splitting them. As stated earlier, the perturbation expansions for the states, in the presence of degeneracies in the spectrum of $\hat{H}_0$, will in general contain terms with zero denominators. Hence these expansions will only make sense if these terms also have vanishing numerators at the same time. So we demand that

$$\langle\psi_{ni}^{(0)}|\hat{W}|\psi_{nj}^{(0)}\rangle = \langle\psi_{ni}^{(0)}|\hat{W}|\psi_{ni}^{(0)}\rangle\delta_{ij} \qquad (i, j = 1, \ldots, g) \tag{3.69}$$

i.e. the $g \times g$ matrix whose elements are $\langle\psi_{ni}^{(0)}|\hat{W}|\psi_{nj}^{(0)}\rangle$ must be a diagonal matrix. Our earlier formula (3.54) becomes:

$$(\hat{H}_0 - E_n^{(0)})|\psi_{ni}^{(1)}\rangle = (E_{ni}^{(1)} - \hat{W})|\psi_{ni}^{(0)}\rangle \tag{3.70}$$

Multiplying the above equation by $\langle\psi_{ni}^{(0)}|$ gives

$$E_{ni}^{(1)}\langle\psi_{ni}^{(0)}|\psi_{ni}\rangle = \langle\psi_{ni}^{(0)}|\hat{W}|\psi_{ni}^{(0)}\rangle \tag{3.71}$$

Again fixing the normalisation such that

$$\langle\psi_{ni}^{(0)}|\psi_{ni}\rangle = 1 \tag{3.72}$$

we have the first order correction as:

$$E_{ni}^{(1)} = \langle\psi_{ni}^{(0)}|\hat{W}|\psi_{ni}^{(0)}\rangle \tag{3.73}$$

To determine the coefficients c_{ni}, starting from (3.70) and multiplying by $\langle\phi_{n\beta}|$ we get:

$$\begin{aligned}
\langle\phi_{n\beta}|\Big[\hat{H}_0 - E_n^{(0)}\Big]|\psi_{ni}^{(1)}\rangle &= \langle\phi_{n\beta}|\Big[(E_{ni}^{(1)} - \hat{W})\Big]|\psi_{ni}^{(0)}\rangle \\
\Longrightarrow \langle\phi_{n\beta}|\hat{W}|\psi_{ni}^{(0)}\rangle &= E_{ni}^{(1)}\langle\phi_{n\beta}|\psi_{ni}^{(0)}\rangle \\
\Longrightarrow \sum_{\alpha=1}^{g} c_{i\alpha}\langle\phi_{n\beta}|\hat{W}|\phi_{n\alpha}\rangle &= E_{ni}^{(1)}\sum_{\alpha=1}^{g} c_{i\alpha}\langle\phi_{n\beta}|\phi_{n\alpha}\rangle
\end{aligned} \tag{3.74}$$

which reduces to:

$$\sum_{\alpha=1}^{g} c_{i\alpha}\Big[\langle\phi_{n\beta}|\hat{W}|\phi_{n\alpha}\rangle - \langle\phi_{n\beta}|\phi_{n\alpha}\rangle E_{ni}^{(1)}\Big] = 0 \tag{3.75}$$

Using the orthonormality of the $\{|\phi_{n\alpha}\rangle\}$, the above equation reduces to:

$$\sum_{\alpha=1}^{g} c_{i\alpha}\Big[\langle\phi_{n\beta}|\hat{W}|\phi_{n\alpha}\rangle - \delta_{\alpha\beta}E_{ni}^{(1)}\Big] = 0 \tag{3.76}$$

This is just a set of homogeneous equations for the g unknown constants $c_{i\alpha}$. For non-trivial solutions we must have:

$$\Big|\langle\phi_{n\beta}|\hat{W}|\phi_{n\alpha}\rangle - \delta_{\alpha\beta}E_{ni}^{(1)}\Big| = 0 \tag{3.77}$$

This determinant yields a $g-$degree polynomial (with g roots, not necessarily all different). These roots correspond to the 1^{st} order correction to the energy $E_n^{(0)}$ and the values of the $c_{i\alpha}$ can be determined by substituting the roots back into (3.76).

Problem 17: Given the Hamiltonian matrix $\mathbf{H} = \mathbf{H}_0 + \mathbf{W}$ constructed using orthogonal basis functions where

$$\mathbf{H} = \begin{pmatrix} 20 & 0 & 0 \\ 0 & 20 & 0 \\ 0 & 0 & 30 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 0 \end{pmatrix}$$

(a) Determine the exact eigenvalues.

Answer: The eigenvalues are 20.806, 18.805, and 30.389.

(b) Determine the eigenvalues correct to second order in the perturbation.

Answer: The eigenvalues, correct to second order, are 20.777, 18.818, and 30.404.

Problem 18: A particle of mass m moves in a two dimensional potential box where $V(x, y) = 0$ when $0 \leq x \leq a$ and $0 \leq y \leq a$, otherwise $V(x, y) = \infty$.

(a) Determine the four lowest eigenvalues and display them on a diagram. Indicate the degeneracies.

Answer: $\frac{h^2}{4ma^2}$, $\frac{5h^2}{8ma^2}$ $(\times 2)$, $\frac{h^2}{ma^2}$, $\frac{5h^2}{4ma^2}$ $(\times 2)$.

(b) When the perturbation $W(x) = bx^2$ is applied, the degeneracies are lifted. Assume that $\hat{W} \ll \hat{H}_0$ and determine the eigenvalues of the perturbed system to first order. Illustrate the shifts on the diagram in part (a).

Answer: $\frac{h^2}{4ma^2}+a^2b\left(\frac{1}{3}-\frac{1}{2\pi^2}\right)$, $\frac{5h^2}{8ma^2}+a^2b\left(\frac{1}{3}-\frac{1}{2\pi^2}\right)$, $\frac{5h^2}{8ma^2}+a^2b\left(\frac{1}{3}-\frac{1}{8\pi^2}\right)$, $\frac{h^2}{ma^2}+a^2b\left(\frac{1}{3}-\frac{1}{8\pi^2}\right)$, $\frac{5h^2}{4ma^2}+a^2b\left(\frac{1}{3}-\frac{1}{2\pi^2}\right)$, $\frac{5h^2}{4ma^2}+a^2b\left(\frac{1}{3}-\frac{1}{18\pi^2}\right)$

Problem 19: A particle of mass m is constrained to move on the $x-y$ plane so that the Hamiltonian is:

$$\hat{H} = \frac{1}{2m}\left(\hat{p}_x^2 + \hat{p}_y^2\right) + \frac{1}{2}k(\hat{x}^2 + \hat{y}^2) + a\hat{x}\hat{y}$$

(a) Solve the problem for the case when $a = 0$. (This is just the two-dimensional harmonic oscillator.)

(b) List the first few energy levels for the two-dimensional oscillator and give the degeneracies.

(c) Use degenerate perturbation theory to determine the energy splitting for the lowest degenerate states and the first order corrections to the wave functions for these states when $a \neq 0$.

Answer: This problem will be done in class.

3.3.2 Time-dependent perturbation theory

In general, if the potential energy function is independent of time $(V(\mathbf{r}, t) = V(\mathbf{r}))$, then the time-dependent Schrödinger equation:

$$\hat{H}\Psi = i\hbar\frac{\partial \Psi}{\partial t} \tag{3.78}$$

is solved by separation of variables:

$$\Psi(\mathbf{r},t) = \psi(\mathbf{r})e^{-iEt/\hbar} \tag{3.79}$$

with $\psi(\mathbf{r})$ the solution of the time-independent Schrödinger equation:

$$\hat{H}\psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{3.80}$$

These are called *stationary states* since all probabilities and expectation values calculated from them are constant in time. To allow for **transitions** between energy levels, we must have a time-dependent potential. Furthermore, if the time dependent part of the Hamiltonian is "small" compared to the time-independent part, it can be treated as a perturbation.

Starting form the Hamiltonian,

$$\hat{H} = \hat{H}_0 + \hat{W}(\mathbf{r},t) \tag{3.81}$$

we seek solutions of the time-dependent Schrödinger equation:

$$\begin{aligned} i\hbar\frac{\partial|\Psi\rangle}{\partial t} &= \hat{H}|\Psi\rangle \\ &= \left(\hat{H}_0 + \hat{W}\right)|\Psi\rangle \end{aligned} \tag{3.82}$$

Expanding $|\Psi\rangle$ in the orthonormal basis of $\hat{H}_0 : \{|\psi_n(\mathbf{r})\rangle\}$, we have:

$$|\Psi(\mathbf{r},t)\rangle = \sum_k e^{-iE_kt/\hbar}\; c_k(t)\; |\psi_k\rangle \tag{3.83}$$

where

$$\hat{H}_0|\psi_k\rangle = E_k|\psi_k\rangle \tag{3.84}$$

The quantity $|c_k(t)|^2$ is the probability of finding the system in state $|\psi_k\rangle$ at time t. Substituting (3.83) into (3.82) we get:

$$\begin{aligned} i\hbar\frac{\partial|\Psi\rangle}{\partial t} &= \sum_k E_k\; e^{-iE_kt/\hbar}\; c_k(t)\; |\psi_k\rangle + i\hbar\; \sum_k e^{-iE_kt/\hbar}\; \frac{dc_k(t)}{dt}\; |\psi_k\rangle \\ \hat{H}|\Psi\rangle &= \left(\hat{H}_0 + \hat{W}\right)\; |\Psi\rangle \\ &= \sum_k E_k e^{-iE_kt/\hbar} c_k(t)|\psi_k\rangle + \sum_k e^{-iE_kt/\hbar}\; c_k(t)\; \hat{W}\; |\psi_k\rangle \end{aligned} \tag{3.85}$$

This gives:

$$i\hbar\sum_k\; e^{-iE_kt/\hbar}\; \frac{dc_k(t)}{dt}\; |\psi_k\rangle = \sum_k e^{-iE_kt/\hbar}\; c_k(t)\; \hat{W}\; |\psi_k\rangle \tag{3.86}$$

Multiplying (3.86) by $\langle\psi_i|$ and using the orthonormality property we finally get:

$$i\hbar\frac{dc_i(t)}{dt} = \sum_k W_{ik}e^{i\omega_{ik}t}c_k(t) \tag{3.87}$$

where

$$\begin{aligned} W_{ik}(t) &= \langle\psi_i(\mathbf{r})|\hat{W}|\psi_k(\mathbf{r})\rangle \\ \omega_{ik} &= \frac{E_i - E_k}{\hbar} \end{aligned} \tag{3.88}$$

Note that (3.87) can be written as:

$$\frac{dc_i(t)}{dt} = -\frac{i}{\hbar}\, W_{ii}\; c_i(t) + \sum_{k\neq i} W_{ik}\; e^{i\omega_{ik}t}\; c_k(t) \tag{3.89}$$

Then, if $W_{ik} = 0 \;\; \forall i \neq k$ the above equation is easily solved:

$$c_i(t) = c_i(0)\; e^{-iW_{ii}t/\hbar} \tag{3.90}$$

which indicates that $|c_i(t)|^2 = |c_i(0)|^2$, i.e. no transitions from the initial state i to other states take place. **Therefore for transitions to take place the perturbation must couple the different eigenstates of $\hat{H}_0$.**

We now develop the perturbation expansion of $c_i(t)$. Assume that the system is in state i at time $t = 0$. Again introducing the parameter λ to isolate the various orders in the perturbation expansion we have:

$$\begin{aligned} c_i(t) &= c_i^{(0)}(t) + \lambda c_i^{(1)}(t) + \lambda^2 c_i^{(2)}(t) + \ldots \\ \hat{W}(t) &\rightarrow \lambda\hat{W}(t) \end{aligned} \tag{3.91}$$

Assume that the system is in state m at time $t = 0$:

$$c_i(0) = \delta_{im} = c_i^{(0)}(0) \tag{3.92}$$

which also implies

$$c_i^{(1)}(0) = c_i^{(2)}(0) = \ldots = 0 \tag{3.93}$$

Substituting (3.91) into (3.87) and isolating the coefficients of the various powers in λ, we have:

0^{th} Order Term

$$\begin{aligned} \frac{dc_i^{(0)}(t)}{dt} &= 0 \\ \Longrightarrow c_i^{(0)}(t) &= \delta_{im} \end{aligned} \tag{3.94}$$

In 0^{th} order, there are no transitions.

1^{st} Order Term

$$\begin{aligned} \frac{dc_i^{(1)}(t)}{dt} &= -\frac{i}{\hbar}\sum_k W_{ik}\; e^{i\omega_{ik}t}\; c_k^{(0)}(t) \\ &= -\frac{i}{\hbar}\; W_{im}\; e^{i\omega_{im}t} \\ \Longrightarrow c_i^{(1)}(t) &= -\frac{i}{\hbar}\int_0^t dt'\; W_{im}(t')\; e^{i\omega_{im}t'} \end{aligned} \tag{3.95}$$

This procedure can be carried on for higher order terms. What is the probability that the system, initially in state m, has made a transition to a state n by time t after the perturbation was switched on? This transition probability is denoted by P_{mn} where:

$$P_{mn}(t) = \Big|\langle\psi_n|\Psi(\mathbf{r},t)\rangle\Big|^2 = |c_n(t)|^2 \tag{3.96}$$

To 1^{st} order we therefore have

$$P_{mn}(t) = \frac{1}{\hbar^2}\Big|\int_0^t dt'\; W_{nm}(t')\; e^{i\omega_{nm}t'}\Big|^2 \tag{3.97}$$

Fermi's golden rule

We apply the formula for P_{mn} to study transitions into a continuous spectrum of final states. Examples include $\alpha-$decay (where the final states, characterised by the momenta of the emitted $\alpha-$particles, lie in a continuum) and optical transitions (a system in an excited state makes a transition to a lower state by emitting a photon).

Consider a perturbation that is switched on at time $t = 0$ and subsequently remains unchanged.

$$\hat{W}(\mathbf{r},t) = V(\mathbf{r})\;\Theta(t) \tag{3.98}$$

where $\Theta(t)$ is the unit step function defined by:

$$\begin{aligned} \Theta(t) &= 1 \quad t > 0 \\ &= 0 \quad t < 0 \end{aligned} \tag{3.99}$$

The transition probability P_{mn} now becomes:

$$\begin{aligned} P_{mn}(t) &= \frac{1}{\hbar^2}\left|\int_0^t dt'\, V_{nm}\, e^{i\omega_{nm}t'}\right|^2 \\ &= \frac{|V_{nm}|^2}{\hbar^2}\left|\frac{e^{i\omega_{nm}t}-1}{\omega_{nm}}\right|^2 \\ &= \frac{|V_{nm}|^2}{\hbar^2}\left[\frac{\sin(\omega_{nm}t/2)}{\omega_{nm}/2}\right]^2 \end{aligned} \tag{3.100}$$

Using the following definition of the Dirac delta function:

$$\lim_{t\to\infty}\frac{\sin^2(\alpha t)}{\pi\alpha^2 t} = \delta(\alpha) \tag{3.101}$$

For long times, (3.100) becomes:

$$\begin{aligned} P_{mn}(t) &= t\frac{2\pi}{\hbar^2}\,\delta(\omega_{nm})\,|V_{nm}|^2 \\ &= t\frac{2\pi}{\hbar}\,\delta(E_n - E_m)\,|V_{nm}|^2 \end{aligned} \tag{3.102}$$

The **transition rate** Γ_{nm} is defined as the transition probability per unit time and is given by:

$$\begin{aligned} \Gamma_{mn} &= \frac{P_{mn}}{t} \\ &= \frac{2\pi}{\hbar}\,\delta(E_n - E_m)\,|V_{nm}|^2 \end{aligned} \tag{3.103}$$

Since we are treating transitions into the continuum of final states, the transition rate to a group of final states (whose energies are clustered around E_n) is of interest. Assume that the matrix elements V_{nm} for all of these nearby final states are equal. Define $\rho_f(E_n)$ as the density of final states such that $\rho_f(E_n)dE_n$ gives the number of final states in the interval dE_n about E_n. The transition rate to this set of states is then:

$$\Gamma = \sum_n \Gamma_{mn} = \int \rho_f(E_n)dE_n\Gamma_{mn} = \frac{2\pi}{\hbar}\,\rho_f(E_m)\,|V_{nm}|^2 \tag{3.104}$$

Now consider the case of **periodic perturbations** where the perturbation has been switched on at $t = 0$ and varies periodically with time. (An example is a periodic external field such as a EM wave). In general:

$$\hat{W}(\mathbf{r},t) = \left[\hat{F}\ e^{-i\omega t} + \hat{F}^{\dagger}\ e^{+i\omega t}\right]\Theta(t) \tag{3.105}$$

where $\hat{F}$ is an operator (independent of time). Substituting (3.105) into (3.97) gives:

$$\begin{aligned} P_{mn}(t) &= \frac{1}{\hbar^2}\left|\int_0^t dt'\ \{e^{i(\omega_{nm}-\omega)t'}\ F_{nm} + e^{i(\omega_{nm}+\omega)t'}\ F^{\dagger}_{nm}\}\right|^2 \\ F_{nm} &= \langle\psi_n|\hat{F}|\psi_m\rangle \\ F^{\dagger}_{nm} &= \langle\psi_n|\hat{F}^{\dagger}|\psi_m\rangle \end{aligned} \tag{3.106}$$

For $t \to \infty$ we have

$$P_{mn}(t) = t\ \frac{2\pi}{\hbar^2}\left\{|F_{nm}|^2\delta(\omega_{nm}-\omega) + |F^{\dagger}_{nm}|^2\delta(\omega_{nm}+\omega)\right\} \tag{3.107}$$

Therefore the transition rate becomes

$$\Gamma_{mn} = \frac{2\pi}{\hbar}\left\{|F_{nm}|^2\delta(E_n - E_m - \hbar\ \omega) + |F^{\dagger}_{nm}|^2\delta(E_n - E_m + \hbar\ \omega)\right\} \tag{3.108}$$

Problem 20: Verify (3.107). Note that the two delta functions do not overlap so the cross terms in the square of the modulus do not contribute.

Problem 21: A hydrogen atom is placed in an electric field $\mathbf{E}(t)$ that is uniform and has the time dependence

$$\mathbf{E}(t) = \begin{cases} 0 & t < 0 \\ \mathbf{E}_0 e^{-\gamma t} & t > 0 \end{cases}$$

What is the probability that as $t \to \infty$, the hydrogen atom, if initially in the ground state, makes a transition to the $2p$ state?

Answer: The transition probability for $1s \to 2p$ is

$$e^2E_0^2\frac{|\langle\psi_{210}|z|\psi_{100}\rangle|^2}{(E_{210}-E_{100})^2 + \hbar^2\gamma^2}$$

Problem 22: Consider a harmonic oscillator described by

$$\hat{H} = \frac{1}{2m}\hat{p}_x^2 + \frac{1}{2}m\omega^2(t)\hat{x}^2$$

where

$$\omega(t) = \omega_0 + \delta\omega\cos(ft)$$

and $\delta\omega \ll \omega_0$. Calculate the probability that a transition occurs from the ground state, as a function of time, given the system is in the ground state at $t = 0$.

Answer: To be done in class.

Problem 23: Consider a two-level system with $E_1 < E_2$. There is a time-dependent potential that connects the two levels as follows:

$$V_{11} = V_{22} = 0, \quad V_{12} = \gamma e^{i\omega t}, \quad V_{21} = \gamma e^{-i\omega t} \tag{3.109}$$

where γ is real. At $t = 0$, it is known that only the lower level is populated — that is, $c_1(0) = 1,\ c_2(0) = 0$.

(a) Find $|c_1(t)|^2$ and $|c_2(t)|^2$ for $t > 0$ by *exactly* solving the coupled differential equation

$$i\hbar\dot{c}_k = \sum_{n=1}^{2} V_{kn}(t) e^{i\omega_{kn}t} c_n$$

for $k = 1, 2$.

(b) Do the same problem using time-dependent perturbation theory to lowest non vanishing order. Compare the two approaches for small values of γ. Treat the following two cases separately: (i) ω very different from ω_{12} and (ii) ω close to ω_{12}.

3.4 JWKB method

The JWKB (Jeffreys, Wentzel, Kramers, Brillouin) method is a semi-classical technique for obtaining approximate solutions to the one-dimensional Schrödinger equation. It is mainly used in calculating bound-state energies and tunnelling rates through potential barriers, and is valid in the limit $\lambda = \frac{h}{p} = \frac{h}{mv} \to 0$ or $\hbar \to 0$ or $m \to \infty$ where m is the mass of the particle, p its momentum etc.

The key idea is as follows. Imagine a particle of energy E moving through a region where the potential $V(x)$ is *constant.* If $E > V$, the wave function is of the form

$$\begin{aligned} \psi(x) &= A\, e^{\pm ikx} \\ k &= \frac{\sqrt{2m(E-V)}}{\hbar} \end{aligned}$$

The plus sign indicates particles travelling to the right etc. The wave function is oscillatory, with constant wavelength $\lambda = 2\pi/k$, and has constant amplitude, A. Consider now the case where

$V(x)$ is not a constant but varies rather slowly in comparison to λ (so that in a region containing many wavelengths the potentially is essentially constant). Then it is reasonable to suppose that ψ remains practically sinusoidal except that the wavelength and the amplitude change slowly with x. This is the central theme of the JWKB method: rapid oscillations are modulated by gradual variation in amplitude and wavelength.

Similarly, if $E < V$ (with V a constant), then ψ is exponential:

$$\begin{aligned} \psi(x) &= A\, e^{\pm Kx} \\ K &= \frac{\sqrt{2m(V-E)}}{\hbar} \end{aligned}$$

Now, if $V(x)$ is not constant but again varies slowly in comparison to $1/K$, the solution remains practically exponential except that A and K are now slowly varying functions of x.

There are of course places where this idea breaks down, e.g. in the vicinity of a classical turning point where $E \approx V$. Here, λ (or $1/K$) goes to infinity and $V(x)$ can hardly be said to vary "slowly"! Proper handling of this is the most difficult aspect of the JWKB approximation but the final results are simple and easy to implement.

3.4.1 Derivation

We seek to solve

$$\frac{d^2\psi}{dx^2} + k^2(x)\psi(x) = 0$$

$$k^2(x) = \frac{2m}{\hbar^2}(E - V(x)) \tag{3.110}$$

The semi-classical limit corresponds to k large. If k were constant, then of course the solutions would just be $e^{\pm ikx}$. This suggests that we try $\psi(x) = e^{iS(x)}$, where in general $S(x)$ is a complex function. Then,

$$\begin{aligned} \frac{d\psi}{dx} &= iS' e^{iS} \\ \frac{d^2\psi}{dx^2} &= (iS'' - S'^2)e^{iS} \end{aligned} \tag{3.111}$$

and the Schrödinger equation reduces to $(iS'' - S'^2 + k^2)e^{iS} = 0$, or

$$\begin{aligned} S' &= \pm\sqrt{k^2(x) + iS''(x)} \\ &= \pm k(x)\sqrt{1 + iS''(x)/k^2} \end{aligned} \tag{3.112}$$

(Note that if k were a constant, $S'' = 0$ and $S' = \pm k$.)

We now attempt to solve the above equation by iteration, using $S' = \pm k$ as the first guess, and as a second guess we use:

$$\begin{aligned} S' &= \pm k\sqrt{1 \pm ik'(x)/k^2} \\ &\approx \pm k\left(1 \pm \frac{i}{2}\frac{k'(x)}{k^2}\right) \\ &\approx \pm k + \frac{i}{2}\frac{k'(x)}{k} \end{aligned} \tag{3.113}$$

where we have assumed that the corrections are small. Then, we have

$$\begin{aligned} \frac{dS}{dx} &= \pm k + \frac{i}{2}\frac{k'}{k} \\ S(x) &\sim \pm\int^x k(x)dx + \frac{i}{2}\int^x \frac{k'}{k}dx + c \end{aligned} \tag{3.114}$$

The second integral is a perfect differential ($d\ln k$), so

$$\begin{aligned} S(x) &= \pm\int^x k(x)dx + \frac{i}{2}\ln k + c \\ \psi &= e^{iS} \\ &= C\, e^{\pm i\int^x k(x)dx} e^{-\frac{1}{2}\ln k} \\ &= \frac{C}{\sqrt{k(x)}} e^{\pm i\int^x k(x)dx} \end{aligned} \tag{3.115}$$

Note that in making the expansion, we have assumed that $\frac{k'}{k^2} \ll 1$ or $\frac{\lambda}{2\pi}\frac{dk}{dx} \ll k$, i.e. that the change in k in one wavelength is much smaller than k. Alternatively, one has $\lambda\frac{dV}{dx} \ll \frac{\hbar^2 k^2}{m}$ so that the change in V in one wavelength is much smaller than the local kinetic energy.

Note that in the classically forbidden regions, $k^2 < 0$, one puts $k = iK(x)$ and carries through the above derivation to get

$$\begin{aligned} \psi(x) &= \frac{C}{\sqrt{K(x)}} e^{\pm\int^x K(x)dx} \\ K^2 &= \frac{2m}{\hbar^2}(V - E) > 0 \end{aligned} \tag{3.116}$$

3.4.2 Connection formulae

In our discussion above, it was emphasised that the JWKB method works when the "short wavelength approximation" holds. This of course breaks down when we hit the classical turning points where $k^2(x) = 0$ (which happens when $E = V$). To overcome this problem, we will derive below equations relating the forms of the solution to both sides of the turning point.

If the potential can be approximated by an increasing linear potential near the turning point $x = a$ (the region $x > a$ being classically forbidden), we can write in the vicinity of the turning point

$$k^2(x) = \frac{2m}{\hbar^2}\left(-\frac{\partial V}{\partial x}\right)_{x=a}(x-a) \tag{3.117}$$

(If we have a potential which cannot be approximated linearly, we must resort to approximations with a quadratic term and find a solution in terms of parabolic cylinder functions. We will not go into the details of this case but you can look it up if you are interested.) The Schrödinger equation near the turning point becomes

$$\psi'' - \left(\frac{\partial V}{\partial x}\right)_{x=a}\frac{2m}{\hbar^2}\,(x-a)\psi = 0 \tag{3.118}$$

This is a linear potential problem which is solved in terms of **Airy** functions. If we let

$$\begin{aligned} y &= \alpha(a-x) \geq 0 \\ \alpha^3 &= \frac{2m}{\hbar^2}\,\frac{\partial V}{\partial x} \geq 0 \end{aligned} \tag{3.119}$$

then the above differential equation becomes

$$\psi''(y) + y\,\psi(y) = 0 \tag{3.120}$$

whose solutions are:

$$\begin{aligned} \psi &= A\sqrt{y}J_{-\frac{1}{3}}(z) + B\sqrt{y}J_{+\frac{1}{3}}(z) \\ z &= \frac{2}{3}\,y^{\frac{3}{2}} \\ &= \int_x^a k(x)\,dx \end{aligned} \tag{3.121}$$

The procedure is now to make asymptotic expansions for the Bessel functions, match them onto the JWKB solutions in the classically allowed and forbidden regions and thus obtain formulae relating the solutions in the two regions.

For the case $y \to \infty$, i.e. $x \ll a$, deep inside the allowed region, we have

$$\begin{aligned} J_\nu &\to \sqrt{\frac{2}{\pi z}} \cos\left(z - \frac{\nu\pi}{2} - \frac{\pi}{4}\right) \\ \psi &\to \sqrt{\frac{2}{\pi z}} \left[A\sqrt{y}\ \cos\left(z + \frac{\pi}{6} - \frac{\pi}{4}\right) + B\sqrt{y}\ \cos\left(z - \frac{\pi}{6} - \frac{\pi}{4}\right)\right] \end{aligned} \tag{3.122}$$

i.e. it oscillates as does the JWKB solution in the left region.

If $y \to 0$, i.e. when $x \to a$, near the turning point, we have

$$\begin{aligned} J_\nu &\sim \frac{\left(\frac{z}{2}\right)^\nu}{\Gamma(\nu+1)} \\ \psi &\to \frac{A\sqrt{y}(\frac{1}{2}z)^{-1/3}}{\Gamma(\frac{2}{3})} + \frac{B\sqrt{y}(\frac{1}{2}z)^{1/3}}{\Gamma(\frac{4}{3})} \end{aligned} \tag{3.123}$$

Since we know that $y = \frac{k^2}{\alpha}$ and $z = \int_x^a k(x)\, dx = \frac{2}{3}\frac{k^3}{\alpha^{3/2}}$, these can be written as

$$\begin{aligned} \psi_{left}(y \to \infty; x \ll a) &\to \sqrt{\frac{3}{\pi}} \frac{\alpha^{1/4}}{\sqrt{k(x)}} \Bigg[A \cos\left(\int_x^a k(x)\, dx - \frac{\pi}{12}\right) \\ &+ \quad B \cos\left(\int_x^a k(x)\, dx + \frac{5\pi}{12}\right)\Bigg] \\ \psi_{left}(\text{near } x = a) &\to \frac{A\ 3^{1/3}}{\Gamma(\frac{2}{3})} + \frac{By}{\Gamma(\frac{4}{3})\ 3^{1/3}} \end{aligned} \tag{3.124}$$

We now follow a similar procedure in the classically forbidden region, $x > a$. Let $y = \alpha(x-a) > 0$. The Schrödinger equation now becomes $\psi'' - y\psi = 0$, which has solutions:

$$\begin{aligned} \psi &= C\ \sqrt{y}\ I_{\frac{1}{3}}(z) + D\ \sqrt{y}\ I_{-\frac{1}{3}}(z) \\ z &= \int_a^x K(x)\, dx \\ K^2(x) &= \frac{2m}{\hbar^2}\Big(V(x) - E\Big) \end{aligned} \tag{3.125}$$

The $I's$ are Bessel functions of imaginary argument.

When the same steps described in detail above are followed, we find

$$\psi_{right}(x \gg a) \quad \to \quad \sqrt{\frac{3}{4\pi}}\ \frac{C\ \alpha^{1/4}}{\sqrt{K(x)}} \left[e^{\int_a^x K(x)\, dx} + e^{-\int_a^x K(x)\, dx}\ e^{-i\frac{5\pi}{6}}\right]$$

$$+ \quad \sqrt{\frac{3}{4\pi}}\,\frac{D\,\alpha^{1/4}}{\sqrt{K(x)}}\left[e^{\int_a^x K(x)\,dx} + e^{-\int_a^x K(x)\,dx}\,e^{i\frac{5\pi}{6}}\right]$$

$$\psi_{right}(\text{near } x = a) \quad \rightarrow \quad \frac{D\,3^{1/3}}{\Gamma(\frac{2}{3})} + \frac{Cy}{\Gamma(\frac{4}{3})\,3^{1/3}} \tag{3.126}$$

We now match these two sets of solutions near $x = a$. Matching the functions at $y = 0$ gives us $D = A$; matching derivatives gives us $B = -C$. If we let $A = 1,\ B = 0$, we find that the cosine solutions to the left correspond to dying exponentials to the right. If we manipulate the constant factors in the above asymptotic solutions, we finally find the connection formulae:

$$\begin{aligned} \frac{2}{\sqrt{k(x)}}\cos\left[\int_x^a k(x)\,dx - \frac{\pi}{4}\right] \quad &\Longleftrightarrow \quad \frac{1}{\sqrt{K(x)}}\,e^{-\int_a^x K(x)\,dx} \\ \frac{1}{\sqrt{k(x)}}\sin\left[\int_x^a k(x)\,dx - \frac{\pi}{4}\right] \quad &\Longleftrightarrow \quad -\frac{1}{\sqrt{K(x)}}\,e^{\int_a^x K(x)\,dx} \end{aligned} \tag{3.127}$$

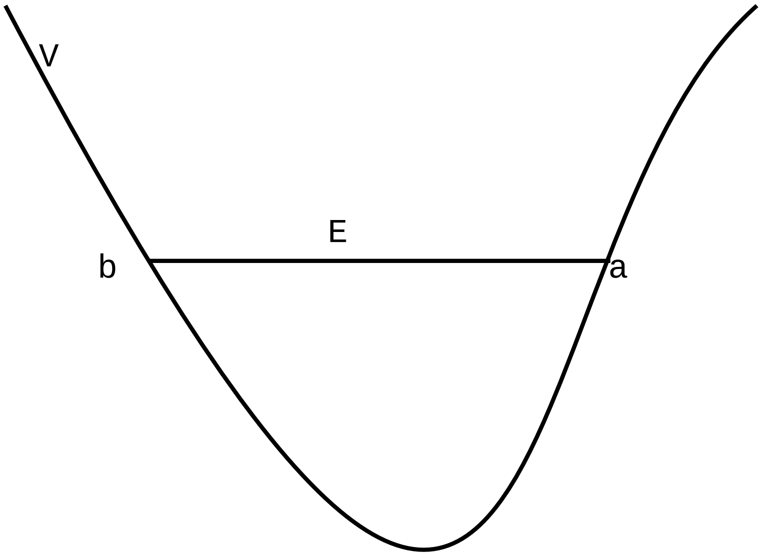
This allow a continuation of solutions. A sine solution on the left matches into a $-e^{+\int_a^x K(x)\,dx}$ solution on the right. Similar formulae for the reverse situation with the turning point on the left ($x < b$ classically forbidden) give

$$\begin{aligned} \frac{1}{\sqrt{K(x)}}\,e^{-\int_x^b K(x)\,dx} \quad &\Longleftrightarrow \quad \frac{2}{\sqrt{k(x)}}\cos\left[\int_b^x k(x)\,dx - \frac{\pi}{4}\right] \\ \frac{1}{\sqrt{K(x)}}\,e^{\int_x^b K(x)\,dx} \quad &\Longleftrightarrow \quad -\frac{1}{\sqrt{k(x)}}\sin\left[\int_b^x k(x)\,dx - \frac{\pi}{4}\right] \end{aligned} \tag{3.128}$$

Problem 24: Verify Eq. (3.127).

3.4.3 *JWKB treatment of the bound state problem

Given an arbitrary potential $V(x)$, we wish to find the approximate eigenstates. Effectively, this means that we must find energies such that the JWKB solutions in the potential well match onto dying exponentials in the classically forbidden regions (i.e. $\psi \rightarrow 0$ as $x \rightarrow \pm\,\infty$). Consider a potential as shown in the figure below:



Far to the left, we will have (as $x \to -\infty$):

$$\psi_{left} = \frac{\text{const}}{\sqrt{K(x)}}\; e^{\int_b^x K(x)\,dx} \to 0 \tag{3.129}$$

Inside the potential well, this matches to:

$$\psi_{inside} = \frac{2\cdot\text{const}}{\sqrt{k(x)}}\;\cos\left(\int_b^x k(x)\,dx - \frac{\pi}{4}\right) \tag{3.130}$$

We now move along to the next turning point $x = a$. Rewriting ψ inside (for easy matching) as:

$$\psi_{inside} = \frac{2\cdot\text{const}}{\sqrt{k(x)}}\;\cos\left(\int_b^a k(x)\,dx - \int_x^a k(x)\,dx - \frac{\pi}{4}\right) \tag{3.131}$$

We define

$$\phi = \int_b^a k(x)\,dx \tag{3.132}$$

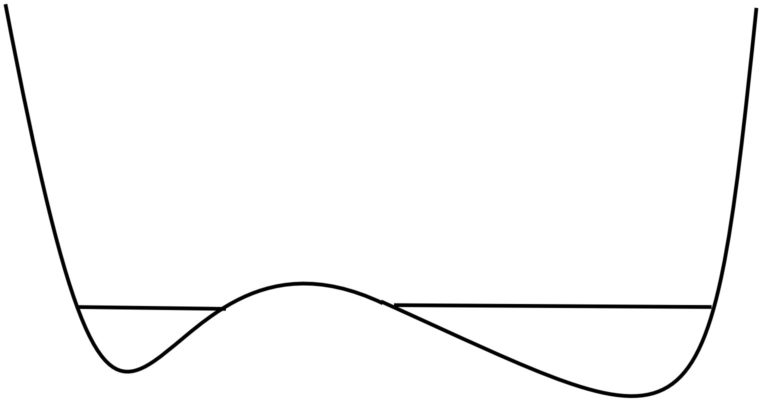
and applying the connection formulae developed in the previous section, we get

$$\begin{aligned}\psi_{inside} &= \frac{2\cdot\text{const}}{\sqrt{k(x)}}\;\cos\left[-\int_x^a k(x)\,dx + \phi - \frac{\pi}{2} + \frac{\pi}{4}\right] \\ &= \frac{2\cdot\text{const}}{\sqrt{k(x)}}\;\left[\cos\left(\phi - \frac{\pi}{2}\right)\;\cos\left(\int_x^a k(x)\,dx - \frac{\pi}{4}\right)\right. \\ &- \left.\sin\left(\frac{\pi}{2} - \phi\right)\;\sin\left(\int_x^a k(x)\,dx - \frac{\pi}{4}\right)\right]\end{aligned} \tag{3.133}$$

The cos solution matches onto the dying exponential, but the sine matches onto a growing exponential. Thus, its coefficient must be zero; i.e. $\sin(\frac{\pi}{2} - \phi) = 0$ or $\frac{\pi}{2} - \phi = -n\pi$ or

$$\int_b^a k(x)\,dx = (n + \frac{1}{2})\pi \tag{3.134}$$

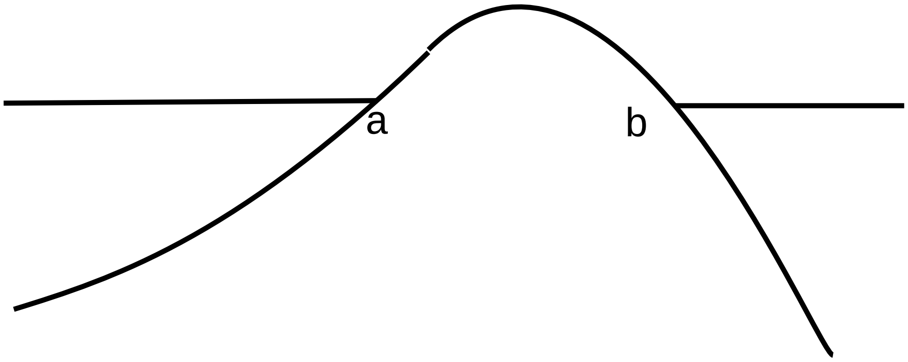
This is similar to the **Bohr-Sommerfeld Quantization Rule**, but with n shifted by $\frac{1}{2}$. This matching may need to be done many times if we have a complicated potential such as:



Problem 25†: Show that Eq. (3.134) gives the exact energy levels of a simple harmonic oscillator.

3.4.4 Barrier penetration

Given a potential like that shown below:



we know that classically there is no probability of getting through the barrier. However, from quantum mechanics we find that an incoming wave function of magnitude A will give rise to a reflected wave function of magnitude B, and a transmitted wave function of magnitude F such that

$$\psi_{left} = \frac{A}{\sqrt{k}}\, e^{i\int_a^x k(x)\,dx} + \frac{B}{\sqrt{k}}\, e^{-i\int_a^x k(x)\,dx}$$

$$\psi_{middle} = \frac{C}{\sqrt{K}}\, e^{-\int_a^x K(x)\,dx} + \frac{D}{\sqrt{K}}\, e^{\int_a^x K(x)\,dx}$$

$$\psi_{right} = \frac{F}{\sqrt{k}}\, e^{i\int_b^x k(x)\,dx} + \frac{G}{\sqrt{k}}\, e^{-i\int_b^x k(x)\,dx} \tag{3.135}$$

Obviously we will set $G = 0$ at the end (there is no wave coming in from the right). First match at the turning point $x = a$. We do this by rewriting ψ_{left} in such a form as to use the connection formulae:

$$\begin{aligned}\psi_{left} &= (A\, e^{-i\frac{\pi}{4}} + B\, e^{i\frac{\pi}{4}})\, \frac{1}{\sqrt{k}}\, \cos\left(\int_x^a k(x)\,dx - \frac{\pi}{4}\right) \\ &+ i\,(-A\, e^{-i\frac{\pi}{4}} + B\, e^{i\frac{\pi}{4}})\, \frac{1}{\sqrt{k}}\, \sin\left(\int_x^a k(x)\,dx - \frac{\pi}{4}\right)\end{aligned} \tag{3.136}$$

On applying the connection formula, we find:

$$\begin{aligned}A\, e^{-i\frac{\pi}{4}} + B\, e^{i\frac{\pi}{4}} &= 2\,C \\ -i\,B\, e^{i\frac{\pi}{4}} + i\,A\, e^{-i\frac{\pi}{4}} &= D\end{aligned} \tag{3.137}$$

We do the same at the next turning point $(x = b)$. Rewrite ψ_{middle} as

$$\begin{aligned}\psi_{middle} &= \frac{C}{\sqrt{K}}\, \Theta^{-1}\, e^{\int_x^b K(x)\,dx} + \frac{D}{\sqrt{K}}\, \Theta\, e^{-\int_x^b K(x)\,dx} \\ \Theta &= e^{\int_a^b K(x)\,dx}\end{aligned} \tag{3.138}$$

which on matching yields:

$$\begin{aligned}F\, e^{i\frac{\pi}{4}} + G\, e^{-i\frac{\pi}{4}} &= 2\,D\,\Theta \\ i\,(F\, e^{i\frac{\pi}{4}} - G\, e^{-i\frac{\pi}{4}}) &= -\frac{C}{\Theta}\end{aligned} \tag{3.139}$$

These relations can be written in a simple matrix notation as:

$$\begin{aligned}\begin{pmatrix} C \\ D \end{pmatrix} &= \frac{1}{2}\begin{pmatrix} e^{-i\frac{\pi}{4}} & e^{i\frac{\pi}{4}} \\ i\,2\,e^{-i\frac{\pi}{4}} & -i\,2\,e^{i\frac{\pi}{4}} \end{pmatrix}\begin{pmatrix} A \\ B \end{pmatrix} \\ \begin{pmatrix} C \\ D \end{pmatrix} &= \begin{pmatrix} -i\,\Theta\, e^{i\frac{\pi}{4}} & i\,\Theta\, e^{-i\frac{\pi}{4}} \\ \frac{1}{2\Theta}\, e^{i\frac{\pi}{4}} & \frac{1}{2\Theta}\, e^{-i\frac{\pi}{4}} \end{pmatrix}\begin{pmatrix} F \\ G \end{pmatrix}\end{aligned} \tag{3.140}$$

After some matrix algebra we get

$$\begin{aligned}\begin{pmatrix} A \\ B \end{pmatrix} &= \mathbf{T}\begin{pmatrix} F \\ G \end{pmatrix} \\ \mathbf{T} &= \frac{1}{2}\begin{pmatrix} 2\,\Theta + \frac{1}{2\Theta} & i\,(2\,\Theta - \frac{1}{2\Theta}) \\ -i\,(2\,\Theta - \frac{1}{2\Theta}) & 2\,\Theta + \frac{1}{2\Theta} \end{pmatrix}\end{aligned} \tag{3.141}$$

where the matrix **T** is called the **Transfer matrix**. If we let $G = 0$, then we find $A = \frac{1}{2}\,(2\,\Theta + \frac{1}{2\Theta})\,F$ or

$$\left|\frac{F}{A}\right|^2 = \left|\frac{1}{\Theta + \frac{1}{4\Theta}}\right|^2 \qquad (3.142)$$

$$\sim \Theta^{-2} \text{ for large } \Theta.$$

Therefore for large Θ the transmission coefficient will be

$$T \equiv \left|\frac{F}{A}\right|^2 \sim \Theta^{-2} = e^{-2\int_a^b K(x)\,dx} \qquad (3.143)$$

Problem 26: Verify Eq. (3.142).

Problem 27: Consider a particle with total energy $E = 0$ moving in the *anharmonic oscillator* potential

$$V(x) = \frac{1}{2} m\omega^2 (x^2 - \epsilon x^4)$$

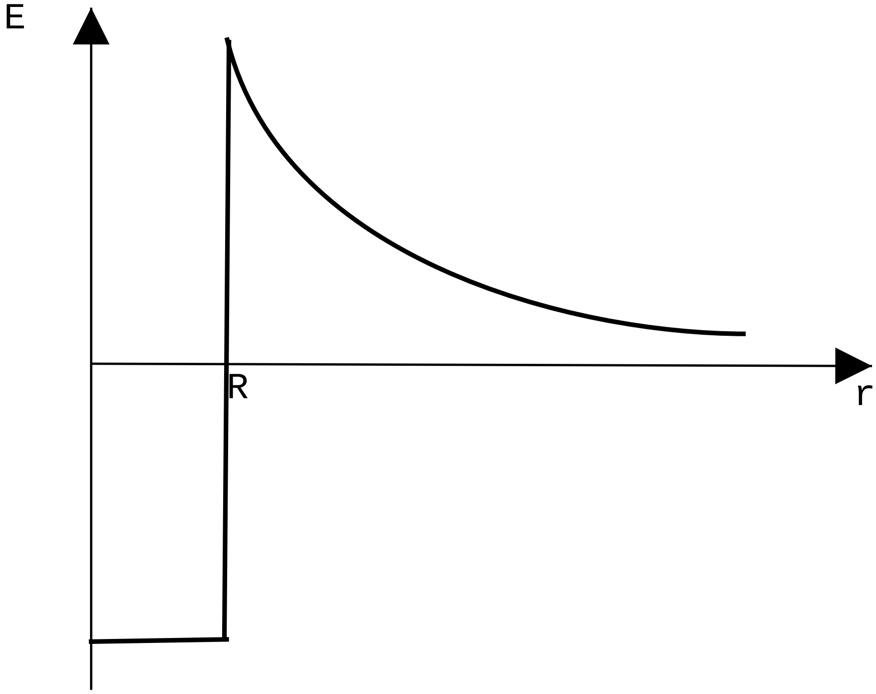
Show that, for small ϵ, the transmission coefficient of the potential barrier between $x = 0$ and $x = 1/\sqrt{\epsilon}$ is

$$T \sim e^{-2m\omega/3\hbar\epsilon}\ .$$

(Notice that this result cannot be obtained using perturbation theory, since the function on the r.h.s. does not have a power series expansion in ϵ.)

3.4.5 Alpha decay of nuclei

The potential that the alpha particle sees is shown below:



We assume that Z is the atomic number of the final nucleus, R is the nuclear radius and the potential is given by:

$$V(r) = \frac{2Ze^2}{r} \quad r > R \tag{3.144}$$

i.e. the potential is Coulombic outside the nucleus. Let the energy of the alpha particle be E and define $r_0 = \frac{2Ze^2}{E}$. We need to evaluate

$$\begin{aligned} I &= \frac{1}{\sqrt{E}} \int_R^{r_0} \left(\frac{2Ze^2}{r} - E \right)^{\frac{1}{2}} dr \\ &= \int_R^{\gamma R} \left(\frac{\gamma R}{r} - 1 \right)^{\frac{1}{2}} dr \\ \gamma &= \frac{2Ze^2}{ER} \end{aligned} \tag{3.145}$$

Substitute

$$\begin{aligned} \frac{r}{\gamma R} &= \sin^2\theta \\ dr &= 2\gamma R \sin\theta \cos\theta \, d\theta \\ \theta_0 &= \sin^{-1} \frac{1}{\sqrt{\gamma}} \end{aligned} \tag{3.146}$$

into the above integral to get

$$\begin{aligned} I &= 2\gamma R \int_{\theta_0}^{\frac{\pi}{2}} \cos^2\theta \, d\theta \\ &= \gamma R \Big[\frac{\pi}{2} - \sin^{-1}\frac{1}{\sqrt{\gamma}} - \frac{(\gamma-1)^2}{\gamma} \Big] \end{aligned} \tag{3.147}$$

Since the transmission coefficient is given by $T \sim e^{-2\int_a^b K(x)\,dx}$ where $K^2(x) = \frac{2m}{\hbar^2}\,(V(r) - E)$ we get

$$\ln T = -2\frac{\sqrt{2m}}{\hbar}\; I\,\sqrt{E} \tag{3.148}$$

For the case that $E \ll \frac{2Ze^2}{R}$, $\gamma \gg 1$ and $I \sim \gamma R[\frac{\pi}{2} - \frac{1}{\sqrt{\gamma}}]$, so we get

$$\begin{aligned} T &= A\,e^{-\beta} \\ \beta &= \pi\frac{\sqrt{2m}}{\hbar}\Big(\frac{2Ze^2}{\sqrt{E}}\Big) \\ A &= e^{-\frac{4}{\hbar}\sqrt{Ze^2 mR}} \end{aligned} \tag{3.149}$$

Every time the alpha particle is incident on the barrier, it has a probability T to escape. To find the probability per unit time (Γ) to escape, we multiply by ω, the frequency of oscillation of an alpha particle in the well of radius R, i.e.

$$\Gamma = \omega\;A\,e^{-\beta} \tag{3.150}$$

We estimate ω as follows:

$$\begin{aligned} \omega &\simeq \frac{v}{R} \\ &\cong \frac{\Delta p}{mR} \\ &\cong \frac{\hbar}{2mR^2} \end{aligned} \tag{3.151}$$

For $R = 10$ fm, one finds $\omega = 10^{20}$ sec, taking $Z = 90$, then $A = 6 \times 10^{-19}$ and $\beta = \frac{182}{\sqrt{E}}$ for E in MeV. So we have

$$\Gamma \sim 0.6\; e^{-182/\sqrt{E}} \;\times 10^{40}\;\mathrm{sec}^{-1} \quad (E \text{ in MeV}) \tag{3.152}$$

Note that although our approximation is crude (since it is not really true that $\gamma \gg 1$), from the above equation, one finds a very strong variation in Γ as a function of E. For example, doubling E from 4 to 8 MeV increases the value of Γ by eleven orders of magnitude!

Chapter 4

Scattering Theory

4.1 Introduction

The path of a moving (incident) particle is altered when it interacts with another (target) particle, such as an atom or a molecule. Phenomena of this sort are generally called **scattering**. Scattering is called **elastic** when the identities and internal properties of the incident particle and the target remain unchanged after the collision, and **inelastic** when the identities or internal properties change or when other particles are emitted or the two particles form a bound state. Analyses of scattering give information on the structure and interactions of atoms, molecules, and elementary particles. We first review the solution to the time independent Schrödinger equation for the spherically symmetric square well to develop the concepts for the general case of scattering from an arbitrary potential.

4.2 Spherically symmetric square well

The potential is defined by:

$$V(r) = \begin{cases} -V_0 & r \le a; \\ 0, & r > a. \end{cases} \tag{4.1}$$

where $V_0 > 0$. An incident particle will feel an attractive potential when it is within the spherical region of radius $r \le a$. Note that this potential is spherically symmetric i.e. it depends only on the distance from the origin. The time independent Schrödinger equation is:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(r) \right] \Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \tag{4.2}$$

Since we have spherical symmetry, we use the method of separation of variables to write $\Psi(\mathbf{r})$ as:

$$\Psi(\mathbf{r}) = R_l(r)\ Y_{lm}(\theta, \varphi) \tag{4.3}$$

where the $Y_{lm}(\theta, \varphi)$ are the Spherical Harmonics and the $R_l(r)$ are the solutions to:

$$\left[-\frac{\hbar^2}{2m}\left(\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2}\right) + V(r)\right]R_l(r) = E\ R_l(r) \tag{4.4}$$

Suppose the energy of the incident particles E are such that $(E > V\ \ \forall r)$. Define

$$k = \sqrt{2m(E-V)}/\hbar$$

Then (4.4) becomes:

$$\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2} + k^2\right]R_l(r) = 0 \tag{4.5}$$

Substituting $\rho = kr$ into (4.5) gives

$$\left[\frac{d^2}{d\rho^2} + \frac{2}{\rho}\frac{d}{d\rho} - \frac{l(l+1)}{\rho^2} + 1\right]R_l(\rho) = 0 \tag{4.6}$$

The general solution to (4.6) are the **spherical Bessel functions**, $j_l(\rho)$ and the **spherical Neumann functions** $n_l(\rho)$:

$$R_l(\rho) = Aj_l(\rho) + Bn_l(\rho) \tag{4.7}$$

(The mathematical properties of these functions are discussed in the next section).

When there is no potential (free particle case), the solution to (4.6) which is regular at the origin, involves only the spherical Bessel function:

$$R_l(r) = Aj_l(kr) \tag{4.8}$$

where now, $k = \sqrt{2mE}/\hbar$. The asymptotic behaviour of this solution far from the origin is:

$$R_l(r) \underset{r\to\infty}{\longrightarrow} A\frac{\sin(kr - l\pi/2)}{kr} \tag{4.9}$$

If we now have the potential (4.1) acting, with $E > 0$ then the solution to (4.6) take the form:

$$R_l(r) = \begin{cases} Aj_l(qr) & r < a; \\ Bj_l(kr) + Cn_l(kr), & r > a. \end{cases}$$

where A, B and C are constants to be determined from the boundary conditions and

$$\begin{aligned} k &= \sqrt{2mE}/\hbar \\ q &= \sqrt{2m(E+V_0)}/\hbar \end{aligned} \tag{4.10}$$

From the continuity conditions (i.e. both $R_l(r)$ and its first derivative must be continuous at $r = a$), we have

$$q\frac{j_l'(qr)}{j_l(qr)}\bigg|_{r=a} = k\left[\frac{Bj_l'(kr) + Cn_l'(kr)}{Bj_l(kr) + Cn_l(kr)}\right]_{r=a} \tag{4.11}$$

from which we can get the ratio C/B. The asymptotic form (see next section) of the solution for the symmetric spherical well (for $r \gg a$) is

$$R_l(r) \underset{r\to\infty}{\longrightarrow} \frac{B}{kr}\left[\sin\left(kr - \frac{l\pi}{2}\right) - \frac{C}{B}\cos\left(kr - \frac{l\pi}{2}\right)\right] \tag{4.12}$$

Introducing

$$\frac{C}{B} = -\tan\delta_l(k) \tag{4.13}$$

the above asymptotic form reduces to:

$$R_l(r) \underset{r\to\infty}{\longrightarrow} \frac{B}{\cos\delta_l(k)}\frac{1}{kr}\left[\sin\left(kr - \frac{l\pi}{2} + \delta_l(k)\right)\right] \tag{4.14}$$

Comparing this with the solution to the free particle case (4.9), we find that the presence of the potential induces a **phase shift** $\delta_l(k)$, which depends of the value of l, the energy of the incident particle (through k) and the strength of the potential (since C/B and therefore $\delta_l(k)$ depends on q).

Problem 1: Show that for $l = 0$ i.e. *s-wave scattering*, the value of the phase-shift, $\delta_0(k)$ is determined by:

$$q\cot qa = k\cot(ka + \delta_0(k)) \tag{4.15}$$

and hence

$$\delta_0(k) = \arctan\left(\frac{k}{q}\tan qa\right) - ka \tag{4.16}$$

4.3 Mathematical interlude

In this section we will review some mathematical tools needed in our study of the quantum mechanical scattering problem. For a more detailed account of the topics discussed here, refer to Arfken's book, "Mathematical Methods for Physicists", Third Edition, which is in the Rayleigh Library.

4.3.1 Brief review of complex analysis

- **Analytic functions:** A function f of the complex variable $z = x + iy$, defined in some region $\mathcal{D}$ of the complex plane, is said to be *analytic* if it satisfies the **Cauchy-Riemann**

conditions. Specifically,

$$\begin{aligned} f(z) &= u(x,y) + i\, v(x,y) \\ \frac{\partial u}{\partial x} &= \frac{\partial v}{\partial y} \\ \frac{\partial u}{\partial y} &= -\frac{\partial v}{\partial x} \end{aligned} \tag{4.17}$$

Analytic functions have some very important properties among which is that their derivatives to all orders exist and are also analytic.

- **Cauchy theorem:** If $f(z)$ is analytic in a simply connected domain $\mathcal{D}$, and if C is a simply closed curve in $\mathcal{D}$, then
$$\oint_C f(z)\, dz = 0$$
which can be proved by applying the Cauchy-Riemann conditions. Note that this result is *independent* of the contour C.

- **Cauchy integral formula:** If $f(z)$ is analytic in a simply connected domain $\mathcal{D}$, and if C is a simply closed curve in $\mathcal{D}$, then
$$\oint_C \frac{f(z)}{z - z_0} dz = \begin{cases} 2\pi i f(z_0) & \text{if } z_0 \in \mathcal{D} \\ 0 & \text{otherwise} \end{cases}$$

- **Residue theorem:** Let $f(z)$ be analytic in some neighbourhood of $z = z_0$, and let C be a simple closed contour lying in this neighbourhood and surrounding $z = z_0$. The quantity
$$\frac{1}{2\pi i} \oint_C f(z) dz = \mathrm{Res} f(z_0)$$
is independent of the choice of C and is called the *residue* of $f(z)$ at the point $z = z_0$. Evaluating many line integrals can therefore be reduced to simply finding the residue.

- **Isolated singularity:** A function $f(z)$, analytic in the neighbourhood of $z = z_0$ with the exception of the point $z = z_0$ itself, is said to have an isolated singularity at $z = z_0$. If there are a finite number of isolated singularities at $z = z_i$ within C, then the integral is given by the sum of residues,
$$\oint_C f(z) dz = 2\pi i \sum_i \mathrm{Res} f(z_i)$$

- **Laurent series:** If $f(z)$ is analytic over a region $r_1 < |z - z_0| < r_2$, it may be represented there by a generalisation of Taylor's series called a *Laurent series*:

  $$f(z) = \sum_{n=-\infty}^{\infty} A_n (z - z_0)^n$$

  where the expansion coefficients are given by

  $$A_n = \frac{1}{2\pi i} \oint_C \frac{f(z)}{(z - z_0)^{n+1}} dz$$

  Then $\text{Res}f(z_0)$ is equal to the coefficient A_{-1}.

Example: The function $f(z) = \exp(z) + \exp(1/z) - 1$ is analytic for $|z| > 0$ and has Laurent expansion

$$f(z) = \sum_{n=-\infty}^{\infty} \frac{1}{|n|!} z^n$$

Therefore in this case $\text{Res}f(0) = 1$.

4.3.2 Properties of spherical Bessel/Neumann functions

The **spherical Bessel/Neumann functions** are the solutions to the $2^{nd}-$order linear differential equation (4.6):

$$\begin{aligned} j_l(\rho) &= (-\rho)^l \left(\frac{1}{\rho}\frac{d}{d\rho}\right)^l \left(\frac{\sin\rho}{\rho}\right) \\ n_l(\rho) &= -(-\rho)^l \left(\frac{1}{\rho}\frac{d}{d\rho}\right)^l \left(\frac{\cos\rho}{\rho}\right) \qquad (4.18) \end{aligned}$$

where $l = 0, 1, 2, \ldots$. They can be expressed by the ascending power series

$$\begin{aligned} j_l(\rho) &= \frac{\rho^l}{(2l+1)!!}\left[1 - \frac{\frac{1}{2}\rho^2}{1!(2l+3)} + \frac{(\frac{1}{2}\rho^2)^2}{2!(2l+3)(2l+5)} - \cdots\right] \\ n_l(\rho) &= -\frac{(2l-1)!!}{\rho^{l+1}}\left[1 - \frac{\frac{1}{2}\rho^2}{1!(1-2l)} + \frac{(\frac{1}{2}\rho^2)^2}{2!(1-2l)(3-2l)} - \cdots\right] \qquad (4.19) \end{aligned}$$

where $(2l+1)!! = 1 \cdot 3 \cdot 5 \cdots (2l+1)$.

The **spherical Hankel functions** $h_l^{(1)}(\rho)$ and $h_l^{(2)}(\rho)$ are defined in terms of j_l and n_l:

$$\begin{aligned} h_l^{(1)}(\rho) &= j_l(\rho) + i n_l(\rho) \\ h_l^{(2)}(\rho) &= [h_l^{(1)}(\rho)]^* \qquad (4.20) \end{aligned}$$

The explicit forms for the first few values of l are:

$$\begin{aligned}
j_0(\rho) &= \frac{\sin\rho}{\rho} \\
n_0(\rho) &= -\frac{\cos\rho}{\rho} \\
h_0^{(1)}(\rho) &= \frac{e^{i\rho}}{i\rho} \\
j_1(\rho) &= \frac{\sin\rho}{\rho^2} - \frac{\cos\rho}{\rho} \\
n_1(\rho) &= -\frac{\cos\rho}{\rho^2} - \frac{\sin\rho}{\rho} \\
h_1^{(1)}(\rho) &= -\frac{e^{i\rho}}{\rho}\left(1 + \frac{i}{\rho}\right) \qquad (4.21)
\end{aligned}$$

Note that in general, $n_l(\rho)$ is **singular** (i.e. diverges) at the origin (while $j_l(\rho)$ is **regular**). This means that if the interval over which the solution to (4.6) is sought includes the origin, then we have to drop the $n_l(\rho)$ part. All the functions and their first derivatives defined above satisfy the following **recursion formulae**:

$$\begin{aligned}
\frac{2l+1}{\rho} z_l(\rho) &= z_{l-1}(\rho) + z_{l+1}(\rho) \\
z_l'(\rho) &= \frac{1}{(2l+1)}\Big[l z_{l-1}(\rho) - (l+1) z_{l+1}(\rho)\Big] \qquad (4.22)
\end{aligned}$$

where the $z_l(\rho)$ are any of the functions $j_l(\rho), n_l(\rho), h_l^{(1)}(\rho), h_l^{(2)}(\rho)$. The spherical Bessel functions have the **integral representation**:

$$j_l(\rho) = \frac{(-i)^l}{2}\int_{-1}^{1} dz P_l(z) e^{i\rho z} \qquad (4.23)$$

where the $P_l(z)$ are the Legendre polynomials. (You can easily verify this for $l = 0, 1$ etc. This formula will be used later when we derive the expansion of plane-waves in terms of spherical waves). The functions have the following **asymptotic behaviour**:

For *small arguments* ($\rho \ll 1, l$):

$$\begin{aligned}
j_l(\rho) &\to \frac{\rho^l}{(2l+1)!!} \\
n_l(\rho) &\to -\frac{(2l-1)!!}{\rho^{l+1}} \qquad (4.24)
\end{aligned}$$

For *large arguments* ($\rho \gg l$):

$$\begin{aligned} j_l(\rho) &\to \frac{1}{\rho}\sin\left(\rho - \frac{l\pi}{2}\right) \\ n_l(\rho) &\to -\frac{1}{\rho}\cos\left(\rho - \frac{l\pi}{2}\right) \\ h_l^{(1)}(\rho) &\to -\frac{i}{\rho}e^{i(\rho - l\pi/2)} \end{aligned} \tag{4.25}$$

4.3.3 Expansion of plane waves in spherical harmonics

In this section, we derive a formula relating a plane wave to an expansion involving the spherical Bessel functions and the spherical harmonics. This will be used later in the partial-wave analysis of spherically symmetric potentials.

The set of spherical wave solutions to the free particle Schrödinger equation, $\{j_l(kr)\}$ is complete. We can therefore expand a plane wave, given by $e^{i\mathbf{k}\cdot\mathbf{r}}$ in terms of these solutions:

$$e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} c_{lm}(\mathbf{k})\, j_l(kr)\, Y_{lm}(\theta,\varphi) \tag{4.26}$$

where the expansion coefficients $c_{lm}(\mathbf{k})$ are to be determined.

First let $\mathbf{k}$ point along the z-axis and let θ be the angle between $\mathbf{r}$ and the z-axis. Then we have

$$\mathbf{k}\cdot\mathbf{r} = kr\cos\theta$$

Note that there is no φ dependence as we have azimuthal symmetry about the z-axis. This implies that

$$\begin{aligned} c_{lm}(\mathbf{k}) &\to A_l \\ Y_{lm}(\theta,\varphi) &\to \left[\frac{2l+1}{4\pi}\right]^{\frac{1}{2}} P_l(\cos\theta) \end{aligned} \tag{4.27}$$

and

$$e^{ikr\cos\theta} = \sum_{l=0}^{\infty} A_l \left[\frac{2l+1}{4\pi}\right]^{\frac{1}{2}} j_l(kr)\, P_l(\cos\theta) \tag{4.28}$$

Using the orthonormality of the Legendre polynomials:

$$\int_{-1}^{1} d\cos\theta\, P_l(\cos\theta)P_{l'}(\cos\theta) = \frac{2}{2l+1}\delta_{ll'} \tag{4.29}$$

we have

$$A_l\; j_l(kr) = \frac{1}{2}[4\pi(2l+1)]^{\frac{1}{2}} \int_{-1}^{1} dz P_l(z) e^{ikrz} \tag{4.30}$$

where $z = \cos\theta$. Using the integral representation of the spherical Bessel function, (4.30) reduces to

$$A_l = i^l\; [4\pi(2l+1)]^{\frac{1}{2}} \tag{4.31}$$

Therefore (4.26) becomes:

$$e^{ikr\cos\theta} = \sum_{l=0}^{\infty} i^l\; (2l+1)\; j_l(kr)\; P_l(\cos\theta) \tag{4.32}$$

For an arbitrary $\mathbf{k}$, all we need to do to generalise (4.32) is use the **addition theorem of spherical harmonics**. If θ is the angle between the vectors $\mathbf{k}$ and $\mathbf{r}$ then this theorem tells us that:

$$P_l(\cos\theta) = \frac{4\pi}{2l+1} \sum_{m=-l}^{l} Y^*_{lm}(\Omega_{\mathbf{k}})\; Y_{lm}(\Omega_{\mathbf{r}}) \tag{4.33}$$

where $\Omega_{\mathbf{k}} \equiv (\theta_{\mathbf{k}}, \varphi_{\mathbf{k}})$ and $\Omega_{\mathbf{r}} \equiv (\theta_{\mathbf{r}}, \varphi_{\mathbf{r}})$. Substituting (4.33) into (4.32) gives the general expansion:

$$e^{i\mathbf{k}\cdot\mathbf{r}} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{l} i^l\; j_l(kr)\; Y^*_{lm}(\Omega_{\mathbf{k}})\; Y_{lm}(\Omega_{\mathbf{r}}) \tag{4.34}$$

4.4 The quantum mechanical scattering problem

In a typical scattering experiment, one might measure the number of particles that are scattered by an angle (θ, φ) into the element $d\Omega$ of the solid angle. The **differential cross section** $\frac{d\sigma}{d\Omega}$ is defined by:

$$\frac{d\sigma}{d\Omega} d\Omega = \frac{\text{number of particles scattered into } d\Omega \text{ per unit time}}{\text{number of incident particles crossing unit area per unit time}}$$

In terms of the incident and scattered fluxes of particles we have:

$$\frac{d\sigma}{d\Omega} d\Omega = \frac{r^2 d\Omega |\mathbf{j}_{scatt}|^2}{|\mathbf{j}_{incid}|^2} \tag{4.35}$$

where $\mathbf{j}_{incid}$ and $\mathbf{j}_{scatt}$ are the incident and scattered flux densities respectively. The **total cross section** σ_{tot} is given by

$$\begin{aligned} \sigma_{tot} &= \int_{\text{unit sphere}} \frac{d\sigma}{d\Omega} d\Omega \\ &= \int_0^{2\pi} d\varphi \int_{-1}^{1} d\cos\theta \frac{d\sigma}{d\Omega} \end{aligned} \tag{4.36}$$

The experimental setup is such that we have a beam of particles incident from $z = -\infty$ travelling in the $+z$ direction (characterised by a plane-wave e^{ikz}) scattered by the target particles (represented mathematically by a potential $V(\mathbf{r})$), so that we have a spherical wave that emanates from the target that describes the scattered particles. We therefore seek a solution to the Schrödinger equation that has the asymptotic form:

$$\psi(\mathbf{r}) \underset{r \to \infty}{\longrightarrow} e^{ikz} + f(\theta, \varphi) \frac{e^{ikr}}{r} \tag{4.37}$$

The function $f(\theta, \varphi)$ is called the **scattering amplitude** and is related to the differential cross section by:

$$\frac{d\sigma}{d\Omega} = |f(\theta, \varphi)|^2 \tag{4.38}$$

Partial wave analysis

We now apply the tools developed in the previous sections to study scattering from a spherically symmetric potential $V(r)$. We assume that $V(r)$ is **short ranged** in the sense that:

$$\left| \int_0^\infty r^2 V(r)\, dr \right| < \infty \tag{4.39}$$

(This ensures that we not only have a well defined scattering problem but also ensures the convergence of the various expressions we shall encounter below.) The time independent Schrödinger equation (4.2) for a particle in such a spherically symmetric potential is given by:

$$\begin{aligned} &[\nabla^2 + k^2 - U(r)]\psi(\mathbf{r}) = 0 \\ &U(r) = \frac{2mV(r)}{\hbar^2} \\ &k^2 = \frac{2mE}{\hbar^2} \end{aligned} \tag{4.40}$$

We now decompose the $\psi(\mathbf{r})$ into spherical waves:

$$\psi(\mathbf{r}) = \sum_{l=0}^{\infty} i^l \; (2l+1) \; R_l(r) \; P_l(\cos\theta) \tag{4.41}$$

where we have taken advantage of the spherical symmetry of $V(r)$. Substituting (4.41) into (4.40) gives the radial equation for $R_l(r)$

$$\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2} + k^2 - U(r) \right] R_l(r) = 0 \tag{4.42}$$

The boundary condition for $R_l(r)$ is that for $r \to 0$, we want $R_l(r)$ to be *finite*. Equation (4.42) reduces for $r \to 0$ to

$$\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2} \right] R_l(r) = 0 \tag{4.43}$$

and the solution that is finite at the origin is

$$R_l(r) \underset{r\to 0}{\longrightarrow} Ar^l \tag{4.44}$$

The remaining boundary condition is determined by the physics of the situation. For an incoming beam $e^{i\mathbf{k}\cdot\mathbf{r}}$ (where it is understood that $\mathbf{k}$ points along the $+z$-axis), we have

$$\begin{aligned} e^{i\mathbf{k}\cdot\mathbf{r}} &= \sum_{l=0}^{\infty} i^l\,(2l+1) j_l(kr)\,P_l(\cos\theta) \\ &= \frac{1}{2}\sum_{l=0}^{\infty} i^l(2l+1)\left[h_l^{(2)}(kr)+h_l^{(1)}(kr)\right] P_l(\cos\theta) \end{aligned} \tag{4.45}$$

The effect of the potential is to cause scattering and hence to modify the amplitude of each of the **outgoing** spherical waves. Therefore, asymptotically, the full solution must be of the form

$$\psi(\mathbf{r}) \underset{r\to\infty}{\longrightarrow} \sum_{l=0}^{\infty} i^l\,(2l+1)\,\frac{1}{2}\left[h_l^{(2)}(kr)+S_l(k)h_l^{(1)}(kr)\right] P_l(\cos\theta) \tag{4.46}$$

where S_l contains the total effect of the scattering. (Note that we have incorporated into (4.46) the fact that, for $r\to\infty$, $\psi(\mathbf{r})$ must satisfy the free particle Schrödinger equation due to the short-range nature of the scattering potential as well as the fact that the scattering effects only the outgoing spherical waves.) Equation (4.46) can now be rewritten as:

$$\begin{aligned} \psi(\mathbf{r}) \quad &\underset{r\to\infty}{\longrightarrow} \quad \frac{1}{2}\sum_{l=0}^{\infty} i^l\,(2l+1)\left[h_l^{(1)}(kr)+h_l^{(2)}(kr)+[S_l(k)-1]h_l^{(1)}(kr)\right] P_l(\cos\theta) \\ &= \sum_{l=0}^{\infty} i^l(2l+1)\left\{j_l(kr)+\frac{1}{2}[S_l(k)-1]h_l^{(1)}(kr)\right\}P_l(\cos\theta) \\ &= e^{i\mathbf{k}\cdot\mathbf{r}}+\sum_{l=0}^{\infty} i^l\,(2l+1)\,\frac{1}{2}\,[S_l(k)-1]\,h_l^{(1)}(kr)\,P_l(\cos\theta) \end{aligned} \tag{4.47}$$

Since this is the solution for large r, we replace $h_l^{(1)}(kr)$ by its asymptotic form to get the final asymptotic form for $\psi(\mathbf{r})$

$$\psi(\mathbf{r}) \underset{r\to\infty}{\longrightarrow} e^{i\mathbf{k}\cdot\mathbf{r}}+f(\theta)\frac{e^{ikr}}{r} \tag{4.48}$$

where

$$f(\theta)=\frac{1}{2ik}\sum_{l=0}^{\infty}(2l+1)\,[S_l(k)-1]\,P_l(\cos\theta) \tag{4.49}$$

(Note that $f(\theta)$ also depends on the energy of the incident particle). Since from (4.38), the differential cross section $\frac{d\sigma}{d\Omega}$ depends only on the scattering amplitude $f(\theta)$, our scattering problem will be solved if we find the $f(\theta)$ or alternatively the $S_l(k)$. The complex numbers $S_l(k)$ can be expressed in terms of real numbers $\delta_l(k)$ called **phase shifts** which are defined by

$$S_l(k) = e^{2i\delta_l(k)} \tag{4.50}$$

(The factor of 2 is conventional). Using this definition of the phase-shifts, the scattering amplitude (4.49) reduces to,

$$f(\theta) = \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)\, e^{i\delta_l(k)}\ \sin\delta_l(k)\ P_l(\cos\theta) \tag{4.51}$$

and the differential cross section is

$$\begin{aligned} \frac{d\sigma}{d\Omega} &= |f(\theta)|^2 \\ &= \frac{1}{k^2}\sum_{l=0}^{\infty}\sum_{l'=0}^{\infty}(2l+1)\ (2l'+1)\ \sin\delta_l\ \sin\delta_{l'}\ \cos(\delta_l-\delta_{l'})\ P_l(\cos\theta)\ P_{l'}(\cos\theta) \end{aligned} \tag{4.52}$$

The total cross section, σ_{tot} is given by

$$\begin{aligned} \sigma_{tot} &= \int_{\text{unit sphere}} \frac{d\sigma}{d\Omega}\, d\Omega \\ &= \frac{4\pi}{k^2}\ \sum_{l=0}^{\infty}(2l+1)\ \sin^2\delta_l \end{aligned} \tag{4.53}$$

Problem 2: Verify equation (4.53).

The total cross section can also be written as

$$\sigma_{tot} = \sum_{l=0}^{\infty}\sigma_l \tag{4.54}$$

where the l^{th} partial cross section σ_l is given by

$$\sigma_l = \frac{4\pi}{k^2}(2l+1)\sin^2\delta_l \tag{4.55}$$

For elastic scattering, the phase shifts are real so it follows from (4.55) that for elastic scattering

$$\sigma_l \le \frac{4\pi}{k^2}(2l+1) \tag{4.56}$$

The value $4\pi(2l+1)/k^2$ is known as the **unitarity bound** and is reached only for

$$\delta_l = \left(n+\frac{1}{2}\right)\pi\ , \quad (n = 0, 1, \ldots) \tag{4.57}$$

This is the condition for **resonance** which shows up as a local maximum in the cross section for the corresponding partial wave.

On the other hand,

$$\begin{aligned}\mathrm{Im} f(\theta) &= \frac{1}{2k}\sum_{l=0}^{\infty}(2l+1)\,\mathrm{Re}[1-S_l]\;P_l(\cos\theta) \\ &= \frac{1}{2k}\sum_{l=0}^{\infty}(2l+1)\;[2\sin^2\delta_l]\;P_l(\cos\theta) \qquad (4.58)\end{aligned}$$

Setting $\theta = 0$ and making use of $P_l(1) = 1$, we get

$$\mathrm{Im} f(\theta = 0) = \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)\;\sin^2\delta_l \qquad (4.59)$$

Comparing this with (4.53), we obtain the **optical theorem**:

$$\sigma_{tot} = \frac{4\pi}{k}\,\mathrm{Im} f(\theta = 0) \qquad (4.60)$$

Problem 3: Consider the elastic scattering of a low energy particle from a "hard sphere" potential ($V(r) = \infty$ for $r \leq a$, 0 otherwise.) Derive an expression for $\tan\delta_l$ and show that for $l = 0$, $\tan\delta_0 = -\tan(ka)$ where $k^2 = 2mE/\hbar^2$. Show that as $k \to 0$ the total cross section approaches $4\pi a^2$. Hence obtain an expression for the s-wave contribution to the forward scattering amplitude $f(\theta = 0)$ and verify the optical theorem for $ka \ll 1$.

4.4.1 Born approximation

Another technique for determining the scattering amplitude, valid for scattering potentials that are weak (and so may be regarded as a perturbation), is by means of the Born approximation. To derive the expression for the scattering amplitude in this case, we first express the time independent Schrödinger equation in integral form.

Integral form of the Schrödinger equation

The time independent Schrödinger equation

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r}) = E\psi(\mathbf{r}) \qquad (4.61)$$

can be written as

$$[\nabla^2 + k^2]\psi(\mathbf{r}) = Q(\mathbf{r}) \tag{4.62}$$

where

$$\begin{aligned} k &= \sqrt{2mE}/\hbar \\ Q(\mathbf{r}) &= \frac{2m}{\hbar^2}\, V(\mathbf{r})\, \psi(\mathbf{r}) \end{aligned} \tag{4.63}$$

Note that

- $V(\mathbf{r})$ in (4.61) need not be spherically symmetric,
- $Q(\mathbf{r})$, the "inhomogeneous source term" itself depends on $\psi(\mathbf{r})$
- (4.62) is the **inhomogeneous Helmholtz equation**.

Suppose there exists a function, $G(\mathbf{r}-\mathbf{r}')$, that is the solution to the following differential equation

$$[\nabla^2 + k^2]G(\mathbf{r}-\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}') \tag{4.64}$$

where the Dirac delta function $\delta(\mathbf{r}-\mathbf{r}')$ is defined by

$$\delta(\mathbf{r}-\mathbf{r}') = \frac{1}{(2\pi)^3}\int_{all\ space} e^{i\mathbf{s}\cdot(\mathbf{r}-\mathbf{r}')} d\mathbf{s} \tag{4.65}$$

In what follows, it is understood that the Laplacian ∇^2 acts on the argument $\mathbf{r}$. We can then express $\psi(\mathbf{r})$ as an integral

$$\psi(\mathbf{r}) = \int_{all\ space} G(\mathbf{r}-\mathbf{r}')Q(\mathbf{r}')d\mathbf{r}' \tag{4.66}$$

Note that (4.66) is a **integral equation** since the unknown function $\psi(\mathbf{r})$ appears under the integral sign. To show that (4.66) satisfies (4.62) consider

$$\begin{aligned} [\nabla^2 + k^2]\psi(\mathbf{r}) &= [\nabla^2 + k^2]\int_{all\ space} G(\mathbf{r}-\mathbf{r}')Q(\mathbf{r}')d\mathbf{r}' \\ &= \int_{all\ space} [(\nabla^2 + k^2)G(\mathbf{r}-\mathbf{r}')]Q(\mathbf{r}')d\mathbf{r}' \\ &= \int_{all\ space} \delta(\mathbf{r}-\mathbf{r}')Q(\mathbf{r}')d\mathbf{r}' \\ &= Q(\mathbf{r}) \end{aligned} \tag{4.67}$$

The function $G(\mathbf{r}-\mathbf{r}')$ is called the **Green's function** for the equation (4.62). In general, the Green's function for a given differential equation represents the "response" to a delta-function source. The easiest way to determine the $G(\mathbf{r}-\mathbf{r}')$ here is by taking the Fourier transform (which turns the *differential* equation into a *algebraic* equation). Let

$$G(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \int_{all\ space} e^{i\mathbf{s}\cdot\mathbf{r}} g(\mathbf{s}) d\mathbf{s} \tag{4.68}$$

Then

$$\begin{aligned}(\nabla^2+k^2)G(\mathbf{r}) &= \frac{1}{(2\pi)^{3/2}} \int_{all\ space} [(\nabla^2+k^2)e^{i\mathbf{s}\cdot\mathbf{r}}] g(\mathbf{s}) d\mathbf{s} \\ &= \frac{1}{(2\pi)^{3/2}} \int_{all\ space} (k^2-s^2) e^{i\mathbf{s}\cdot\mathbf{r}} g(\mathbf{s}) d\mathbf{s} \end{aligned} \tag{4.69}$$

Problem 4: Verify Eq (4.69).

This means that

$$\frac{1}{(2\pi)^{3/2}} \int_{all\ space} (k^2-s^2) e^{i\mathbf{s}\cdot\mathbf{r}} g(\mathbf{s}) d\mathbf{s} = \frac{1}{(2\pi)^3} \int_{all\ space} e^{i\mathbf{s}\cdot\mathbf{r}} d\mathbf{s} \tag{4.70}$$

which implies

$$g(\mathbf{s}) = \frac{1}{(2\pi)^{3/2}} \frac{1}{(k^2-s^2)} \tag{4.71}$$

so that

$$G(\mathbf{r}) = \frac{1}{(2\pi)^3} \int_{all\ space} e^{i\mathbf{s}\cdot\mathbf{r}} \frac{1}{(k^2-s^2)} d\mathbf{s} \tag{4.72}$$

To evaluate the integral (4.72) we first note that $\mathbf{r}$ is fixed as far as the $\mathbf{s}$ integration is concerned. It is advantageous to do the integration in spherical coordinates such that $\mathbf{s} = (s, \theta, \varphi)$ and with the polar axis pointing along $\mathbf{r}$. Then since $\mathbf{s}\cdot\mathbf{r} = sr\cos\theta$ and $d\mathbf{s} = s^2 \sin\theta\, d\theta\, d\varphi\, ds$, (4.72) becomes

$$G(\mathbf{r}) = \frac{1}{(2\pi)^3} \int_0^{2\pi} d\varphi \int_0^{\infty} \frac{s^2 ds}{(k^2-s^2)} \int_{-1}^{1} d\cos\theta\, e^{isr\cos\theta} \tag{4.73}$$

The φ integration is trivial ($= 2\pi$) and the θ integration yields

$$\int_{-1}^{1} d\cos\theta\, e^{isr\cos\theta} = \frac{2\sin(sr)}{sr} \tag{4.74}$$

Thus

$$\begin{aligned} G(\mathbf{r}) &= \frac{1}{(2\pi)^2} \frac{2}{r} \int_0^{\infty} \frac{s\sin(sr)}{(k^2-s^2)} ds \\ &= \frac{1}{4\pi^2 r} \int_{-\infty}^{\infty} \frac{s\sin(sr)}{(k^2-s^2)} ds \end{aligned} \tag{4.75}$$

Writing $\sin(sr) = \frac{1}{2i}(e^{isr} - e^{-isr})$ (4.75) becomes

$$\begin{aligned} G(\mathbf{r}) &= \frac{i}{8\pi^2 r}[I_1 - I_2] \\ I_1 &= \int_{-\infty}^{\infty} \frac{se^{isr}}{(s-k)(s+k)} ds \\ I_2 &= \int_{-\infty}^{\infty} \frac{se^{-isr}}{(s-k)(s+k)} ds \end{aligned} \tag{4.76}$$

The integrals I_1 and I_2 can be evaluated using **Cauchy's integral formula**.

Problem 5: In (4.76), the integration is along the real axis and passes right over the pole singularities at $\pm k$. How to deal with integration in the vicinity of these singularities is fixed by the boundary conditions; the result is that the contour of integration should go *over* the singularity at $-k$ and *under* the singularity at $+k$. We must close the contour in such a way that the semicircle at infinity contributes nothing.

- Show that for I_1, we must close the contour *above* the real axis. Hence, show that

  $$I_1 = i\pi e^{ikr}$$

- For I_2 show that the contour must be closed below the real axis so that

  $$I_2 = -i\pi e^{ikr}$$

  (Remember that when you go round the contour in the clockwise direction, you pick up a minus sign).

Using the results of the above problem, we have

$$G(\mathbf{r}) = -\frac{e^{ikr}}{4\pi r} \tag{4.77}$$

so that the general solution to (4.66) takes the form

$$\psi(\mathbf{r}) = \psi_0(\mathbf{r}) - \frac{m}{2\pi\hbar^2}\int_{all\ space} \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} V(\mathbf{r}')\psi(\mathbf{r}')d\mathbf{r}' \tag{4.78}$$

where the $\psi_0(\mathbf{r})$ satisfies the free particle Schrödinger equation

$$[\nabla^2 + k^2]\psi_0(\mathbf{r}) = 0 \tag{4.79}$$

Equation (4.78) is the integral form of the Schrödinger equation and is equivalent to the differential form plus boundary conditions.

Problem 6: Check that (4.78) satisfies (4.62) by direct substitution.

Hint: The following identity will be useful:

$$\nabla^2\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|}\right) = -4\pi\delta(\mathbf{r}-\mathbf{r}')$$

The ∇ acts on the $\mathbf{r}$ argument.

First Born approximation

Suppose $V(\mathbf{r}')$ is **localised** about $\mathbf{r}' = 0$. (This means that the potential drops to zero outside some finite region). We want to determine $\psi(\mathbf{r})$ at points *far away* (i.e. $r \gg r'$) from the scattering centre. For this case we have

$$\begin{aligned} |\mathbf{r}-\mathbf{r}'|^2 &\cong r^2\left(1-\frac{2\mathbf{r}\cdot\mathbf{r}'}{r^2}\right) \\ \implies |\mathbf{r}-\mathbf{r}'| &\cong r - \hat{\mathbf{r}}\cdot\mathbf{r}' \end{aligned} \tag{4.80}$$

where $\hat{\mathbf{r}}$ denotes a unit vector pointing along $\mathbf{r}$. Let

$$\mathbf{k}' \equiv k\hat{\mathbf{r}}$$

then

$$e^{ik|\mathbf{r}-\mathbf{r}'|} \cong e^{ikr}e^{-i\mathbf{k}'\cdot\mathbf{r}'} \tag{4.81}$$

and therefore

$$\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \cong \frac{e^{ikr}}{r}e^{-i\mathbf{k}'\cdot\mathbf{r}'} \tag{4.82}$$

In the case of scattering where the incident beam is along the $z-$axis, we require

$$\psi_0(\mathbf{r}) = e^{ikz} \tag{4.83}$$

In the asymptotic limit (large r), (4.78) reduces to

$$\psi(\mathbf{r}) \cong e^{ikz} - \frac{m}{2\pi\hbar^2}\frac{e^{ikr}}{r}\int_{all\ space} e^{-i\mathbf{k}'\cdot\mathbf{r}'}V(\mathbf{r}')\psi(\mathbf{r}')d\mathbf{r}' \tag{4.84}$$

from which we can read off the scattering amplitude

$$f(\theta,\varphi) = -\frac{m}{2\pi\hbar^2}\int_{all\ space} e^{-i\mathbf{k}'\cdot\mathbf{r}'}V(\mathbf{r}')\psi(\mathbf{r}')d\mathbf{r}' \tag{4.85}$$

This expression for $f(\theta,\varphi)$ is *exact.* We now invoke the **Born approximation**: Suppose the incoming plane wave is *not substantially altered by the potential* (i.e. the scattering potential is weak), than it makes sense to substitute

$$\psi(\mathbf{r}') \approx \psi_0(\mathbf{r}') = e^{ikz'} = e^{i\mathbf{k}\cdot\mathbf{r}'} \tag{4.86}$$

where $\mathbf{k} = k\hat{\mathbf{z}}$, into the integral (4.84). The scattering amplitude in the Born approximation then reduces to

$$f(\mathbf{k}',\mathbf{k}) = f(\theta,\varphi) = -\frac{m}{2\pi\hbar^2}\int_{all\ space} e^{-i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}'}V(\mathbf{r}')d\mathbf{r}' \tag{4.87}$$

(Note: $\mathbf{k}$ and $\mathbf{k}'$ both have magnitude k but the former points in the direction of the incident beam while the latter points towards the detector). Equation (4.87) indicates that $f(\theta,\varphi)$ in the Born approximation is just the Fourier transform of the scattering potential.

For a spherically symmetric potential, $V(\mathbf{r}) = V(r)$, the Born approximation for $f(\theta)$ reduces to

$$f(\theta) \cong -\frac{2m}{q\hbar^2}\int_0^\infty rV(r)\sin(qr)dr \tag{4.88}$$

where

$$\begin{aligned}
\mathbf{q} &= \mathbf{k}' - \mathbf{k} \\
q &= |\mathbf{k}' - \mathbf{k}| \\
&= 2k\sin(\theta/2) \\
(\mathbf{k}' - \mathbf{k})\cdot\mathbf{r}' &= qr'\cos\theta'
\end{aligned} \tag{4.89}$$

For **low-energy scattering**, where the de Broglie wavelength of the scattered particle is much larger than the extent of the scattering region, the Born approximation simplifies to

$$f(\theta,\varphi) \cong -\frac{m}{2\pi\hbar^2}\int_{all\ space} V(\mathbf{r}')d\mathbf{r}' \tag{4.90}$$

Problem 7: A potential of considerable utility in both nuclear physics and chemistry is the Yukawa or "screened" Coulomb potential:

$$V(r) = -A\frac{e^{-r/r_0}}{r}$$

where $r_0 > 0$. Determine the scattering amplitude in the first Born approximation for this potential. Hence, obtain the differential scattering cross section in the limit $r_0 \to \infty$.

Answer: The scattering amplitude is given by

$$f(\theta) = \frac{2Am}{\hbar^2}\left(\frac{r_0^2}{1+q^2r_0^2}\right)$$

and the differential cross section at the limit $r_0 \to \infty$ is

$$\frac{d\sigma}{d\Omega} = \frac{4A^2m^2}{\hbar^4q^4}$$

Problem 8†: Determine the scattering amplitude in the first Born approximation for the spherically symmetric square well.

Answer: For an attractive square well, we have

$$f(\theta) = \frac{2mV_0}{\hbar^2q^3}(\sin qa - qa\cos qa)$$

Problem 9†: Determine the scattering amplitude in the first Born approximation for the potential

$$V(r) = -Ae^{-br^2}$$

where $A,\ b > 0$.

Answer: The scattering amplitude is given by

$$f(\theta) = \frac{Am\sqrt{\pi}}{2\hbar^2b^{3/2}}e^{-q^2/4b}$$

4.5 *Formal time-independent scattering theory

We now develop a time-independent formulation of scattering processes. Assume that the Hamiltonian for the scattering problem is

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{4.91}$$

where $\hat{H}_0$ stands for the kinetic energy operator

$$\hat{H}_0 = \frac{\hat{p}^2}{2m} \tag{4.92}$$

In the absence of a scatterer, $\hat{V}$ would be zero and the energy eigenstate would be just a free particle state $|\mathbf{p}\rangle$ where

$$\hat{H}_0|\mathbf{p}\rangle = \frac{p^2}{2m}|\mathbf{p}\rangle \tag{4.93}$$

The presence of $\hat{V}$ therefore causes the energy eigenstate to be different from the free-particle state. For *elastic scattering*, where no change in energy is involved, we need to find solutions to Eq. (4.91) with the *same* energy eigenvalue as Eq. (4.93).

In general, let $|\phi\rangle$ be an energy eigenstate of $\hat{H}_0$

$$\hat{H}_0|\phi\rangle = E|\phi\rangle \tag{4.94}$$

where in general, $|\phi\rangle$ stands for either plane waves or free spherical wave states. We want to solve

$$\hat{H}|\psi\rangle = \left(\hat{H}_0 + \hat{V}\right)|\psi\rangle = E|\psi\rangle \tag{4.95}$$

with both $\hat{H}_0$ and $\hat{H}$ exhibiting **continuous** energy spectra. We look for solutions to Eq. (4.95) such that

$$|\psi\rangle \rightarrow |\phi\rangle \quad \text{as} \quad \hat{V} \rightarrow 0 \tag{4.96}$$

where $|\phi\rangle$ is the solution to the free particle Schrödinger equation Eq. (4.94) with the *same* energy eigenvalue.

The desired solution (at least in a formal sense) is

$$|\psi\rangle = |\phi\rangle + \left(E - \hat{H}_0\right)^{-1}\hat{V}|\psi\rangle \tag{4.97}$$

although there are complications arising from the singular nature of the operator $(E - \hat{H}_0)^{-1}$.

Problem 10†: Verify that Eq. (4.97) is indeed a solution to Eq. (4.95)

Moreover, the prescription for dealing with the singular nature of $(E - \hat{H}_0)^{-1}$ is to treat E as a **complex variable** and to add an infinitesimal imaginary number to it, i.e.

$$|\psi\rangle = |\phi\rangle + \frac{1}{(E - \hat{H}_0 + i\epsilon)}\hat{V}|\psi\rangle \tag{4.98}$$

Eq. (4.98) is called the **Lippmann-Schwinger equation** and is a vector equation independent of any particular representations.

4.5.1 *Lippmann-Schwinger equation in the position representation

We now focus on expressing the operator Lippmann-Schwinger equation Eq. (4.98) in the position representation, which enables us to obtain the scattering amplitude in the (first) Born approximation. In the position basis (4.98) becomes

$$\langle\mathbf{r}|\psi\rangle = \langle\mathbf{r}|\phi\rangle + \int d\mathbf{r}'\langle\mathbf{r}|\frac{1}{(E - \hat{H}_0 + i\epsilon)}|\mathbf{r}'\rangle\langle\mathbf{r}'|\hat{V}|\psi\rangle \tag{4.99}$$

which is an **integral equation** for scattering because the unknown ket $|\psi\rangle$ appears under the integral sign. The quantity

$$G(\mathbf{r},\mathbf{r}') = \frac{\hbar^2}{2m}\langle\mathbf{r}|\frac{1}{(E-\hat{H}_0+i\epsilon)}|\mathbf{r}'\rangle \tag{4.100}$$

is just the Green's function (equation (4.77)) derived in the previous section. It was worked out to be

$$G(\mathbf{r},\mathbf{r}') = -\frac{1}{4\pi}\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}$$

where $E = \hbar^2k^2/2m$. Equation (4.98) then just reduces to the integral equation (4.66). By looking at the large distance behaviour of r, we again get the expression for the scattering amplitude describing an incident beam of particles in the direction $\mathbf{k}$ being scattered in the direction $\mathbf{k}'$:

$$f(\mathbf{k}',\mathbf{k}) = -\frac{m}{2\pi\hbar^2}\langle\mathbf{k}'|\hat{V}|\psi\rangle \tag{4.101}$$

which is *exact*.

4.5.2 *Born again!

We now define the **transition operator** $\hat{T}$, such that

$$\hat{V}|\psi\rangle = \hat{T}|\phi\rangle \tag{4.102}$$

Multiplying the Lippmann-Schwinger equation Eq.(4.98) by $\hat{V}$, we obtain

$$\hat{T}|\phi\rangle = \hat{V}|\phi\rangle + \hat{V}\frac{1}{(E-\hat{H}_0+i\epsilon)}\hat{T}|\phi\rangle \tag{4.103}$$

which is supposed to hold for any $|\phi\rangle$ (taken to be any plane-wave state for example, which we know are complete). Therefore the following operator equation is satisfied:

$$\hat{T} = \hat{V} + \hat{V}\frac{1}{(E-\hat{H}_0+i\epsilon)}\hat{T} \tag{4.104}$$

The scattering amplitude can now be written as

$$f(\mathbf{k}',\mathbf{k}) = -\frac{m}{2\pi\hbar^2}\langle\mathbf{k}'|\hat{T}|\mathbf{k}\rangle \tag{4.105}$$

which shows that to determine the scattering amplitude, it is sufficient to know the transition operator $\hat{T}$.

An iterative solution for $\hat{T}$ is obtained as follows:

$$\hat{T} = \hat{V} + \hat{V}\frac{1}{(E-\hat{H}_0+i\epsilon)}\hat{V} + \hat{V}\frac{1}{(E-\hat{H}_0+i\epsilon)}\hat{V}\frac{1}{(E-\hat{H}_0+i\epsilon)}\hat{V} + \cdots \tag{4.106}$$

Correspondingly, we can expand $f(\mathbf{k}', \mathbf{k})$ as follows:

$$f(\mathbf{k}', \mathbf{k}) = \sum_{n=1}^{\infty} f^{(n)}(\mathbf{k}', \mathbf{k}) \tag{4.107}$$

where n is the number of times the $\hat{V}$ operator enters. We have

$$\begin{aligned} f^{(1)}(\mathbf{k}', \mathbf{k}) &= -\frac{m}{2\pi\hbar^2}\langle \mathbf{k}'|\hat{V}|\mathbf{k}\rangle \\ f^{(2)}(\mathbf{k}', \mathbf{k}) &= -\frac{m}{2\pi\hbar^2}\langle \mathbf{k}'|\hat{V}\frac{1}{(E - \hat{H}_0 + i\epsilon)}\hat{V}|\mathbf{k}\rangle \end{aligned} \tag{4.108}$$

etc., which are just the first, second etc. order Born approximations and is called the **Born series**. A physical interpretation is that the scattering process is viewed as a multi-step process with, for example, $f^{(2)}(\mathbf{k}', \mathbf{k})$ being viewed as the incident wave (with wave vector $\mathbf{k}$ undergoing two sequential interactions before being scattered into the direction $\mathbf{k}'$ and so on. In this context, the Green's function is called the **propagator** – it tells us how a disturbance propagates between one interaction and the next. The Born series was the inspiration for Feynman's formulation of relativistic quantum mechanics, which is expressed entirely in terms of **vertex functions** $\hat{V}$ and propagators G, connected together in **Feynman diagrams**. It is this technique that forms the basis for the treatment of the quantum theory of fields and its application to a wide range of phenomena including elementary particle and condensed matter physics.

Chapter 5

Identical Particles in Quantum Mechanics

5.1 Introduction

Consider a system of N identical particles. In classical physics it is possible (at least in principle), to keep track of the individual particles at each instant of time. In this sense, the particles retain their individuality and are distinguishable. In QM the situation is very different. When two identical particles come together, interact, and then separate, their individuality is lost. This is because we cannot 'tag' the particles nor can we follow their trajectories (since this would entail a position measurement at each instant in time which necessarily disturbs the system). Therefore, whereas in classical physics there is no necessity to treat distinguishable or non-distinguishable particles differently, in QM the fact that identical particles are indistinguishable must be "built in" to the theory.

5.2 Multi-particle systems

Consider the Hamiltonian for the $N-$particle system

$$\hat{H}(1,2,\ldots,N) = \sum_{i=1}^{N}\left[-\frac{\hbar^2}{2m}\nabla_i^2 + v_{ext}(i)\right] + V_{int}(1,2,\ldots,N) \tag{5.1}$$

where the labels $1,2,\ldots$ represent **all** of particle i's internal (for example spin) and external (for example position) degrees of freedom, collectively. $v_{ext}(i)$ is the external potential (for example due to an electric/magnetic field) acting on the i^{th} particle, and $V_{int}(1,2,\ldots,N)$ represents the mutual interaction between the particles. If, as is usually the case, the interaction between the particles is due to two-body interactions (as in the case of the Coulomb interaction between charged particles) then

$$V_{int}(1,2,\ldots,N) = \frac{1}{2}\sum_{i\neq j}^{N} V_{int}(i,j) \tag{5.2}$$

where due to Newton's 3^{rd} law

$$V_{int}(i,j) = V_{int}(j,i) \tag{5.3}$$

Therefore the $N-$particle Hamiltonian is given by

$$\hat{H}(1,2,\ldots,N) = \sum_{i=1}^{N}\Big[-\frac{\hbar^2}{2m}\nabla_i^2 + v_{ext}(i)\Big] + \frac{1}{2}\sum_{i\neq j}^{N} V_{int}(i,j) \tag{5.4}$$

Note that this Hamiltonian in **invariant under the interchange of particles**. Let $\hat{P}_{ij}$ be a particle interchange (or permutation) operator which interchanges the labels $i \leftrightarrow j$ such that its action on $f(1,2,\ldots,N)$ is

$$\hat{P}_{ij} f(1,2,\ldots,i,\ldots,j,\ldots,N) = f(1,2,\ldots,j,\ldots,i,\ldots,N) \tag{5.5}$$

Note that from (5.5) it is apparent that $\hat{P}_{ij}^2 = \hat{I}$ which implies that $\hat{P}_{ij}$ has ±1 as eigenvalues. (In fact, $\hat{P}_{ij}$ is both Hermitian and unitary). Let $\hat{P}$ denote some arbitrary particle interchange, not necessarily pair-wise. (Note that $\hat{P}$ can be represented as the product of pair-wise interchanges, $\hat{P}_{ij}$. $\hat{P}$ is called even(odd) if it is made up of an even(odd) number of the $\hat{P}_{ij}$). The property $\hat{P}^2 = \hat{I}$ still holds. Then by the symmetry of the Hamiltonian under any number of interchanges of identical particles, we have

$$[\hat{H}, \hat{P}] = 0 \tag{5.6}$$

Theorem: Let $\Psi(1,2,\ldots,N)$ denote the $N-$body wave function for this system. If $\Psi(1,2,\ldots,N)$ is an eigenfunction of $\hat{H}(1,2,\ldots,N)$ with eigenvalue E, then so is $\hat{P}\Psi$.

Proof: Given $\hat{H}\Psi = E\Psi$, from (5.6) we have

$$\hat{H}(\hat{P}\Psi) = \hat{P}(\hat{H}\Psi) = E(\hat{P}\Psi) \tag{5.7}$$

i.e. $\hat{P}\Psi$ is also an eigenstate corresponding to the same eigenvalue E.

This means that we can use the eigenvalues of $\hat{P}$ (which are ±1) together with the energy E to label the eigenstates of the Hamiltonian. In general for any symmetrical operator (i.e. one which is even under interchange of particle labels), $\hat{S}(1,2,\ldots,N)$ we have

$$[\hat{P}, \hat{S}] = 0 \tag{5.8}$$

If $|\Phi\rangle = \hat{P}|\Psi\rangle$ for some arbitrary $|\Psi\rangle$ then

$$\begin{aligned}\langle\Phi|\hat{S}|\Phi\rangle &= \langle\Psi|\hat{P}^\dagger\hat{S}\hat{P}|\Psi\rangle \\ &= \langle\Psi|\hat{P}^\dagger\hat{P}\hat{S}|\Psi\rangle \\ &= \langle\Psi|\hat{S}|\Psi\rangle \qquad (5.9)\end{aligned}$$

using the unitary property of $\hat{P}$. Therefore the expectation value (and more generally the matrix elements) of the states $|\Psi\rangle$ and $\hat{P}|\Psi\rangle$ of a symmetrical operator $\hat{S}$ are equal. Since identical particles are influenced equivalently by any physical process, **all physical operators are symmetric** and thus the states $\Psi(1,2,\ldots,N)$ and $\hat{P}\Psi(1,2,\ldots,N)$ **cannot** be distinguished from one another. As we have seen, we must in fact have

$$\hat{P}\Psi(1,2,\ldots,N) = \pm\Psi(1,2,\ldots,N) \qquad (5.10)$$

All experimental observations to date imply that the sign factor is either always +1 (i.e. Ψ is completely symmetric and represents **bosons**) or else it is ±1 for even/odd $\hat{P}$ (i.e. Ψ is antisymmetric with respect to any pair-wise interchange and represents **fermions**). Furthermore, the nature of a given type of particle is intimately connected with its intrinsic spin. This connection between spin and symmetry follows from the **spin-statistics theorem** of quantum field theory which states that all elementary particles with integer spin (including spin 0) are bosons (obeying Bose-Einstein statistics) and all particles with half-odd integer spin are fermions (obeying Fermi-Dirac statistics).

5.2.1 Pauli exclusion principle

Let $\Psi_A(1,2,\ldots,N)$ denote a totally antisymmetric wave function. Then

$$\begin{aligned}\hat{P}_{ij}\Psi_A(1,\ldots,i,\ldots,j,\ldots N) &= \Psi_A(1,\ldots,j,\ldots,i,\ldots N) \\ &= -\Psi_A(1,\ldots,i,\ldots,j,\ldots N) \qquad (5.11)\end{aligned}$$

If we set $i = j$ in (5.11) we get

$$\Psi_A(1,\ldots,i,\ldots,i,\ldots N) = -\Psi_A(1,\ldots,i,\ldots,i,\ldots N) \qquad (5.12)$$

i.e.

$$\Psi_A(1,\ldots,i,\ldots,i,\ldots N) = 0$$

Thus two fermions in the same spin state cannot occupy the same position in space. (Remember that i includes both the spatial and spin degrees of freedom). This is the **Pauli Exclusion Principle**. Note that this is true independent of whether the particles interact with one another or not. The key point is that they are identical fermions.

5.2.2 Representation of $\Psi(1,2,\ldots,N)$

For N identical **non-interacting** particles, the Hamiltonian is

$$\hat{H} = \sum_{i=1}^{N} \hat{H}(i) \tag{5.13}$$

i.e the sum of $N-$identical one-particle Hamiltonians $\hat{H}(i)$. Assume that the eigenfunctions of $\hat{H}(i)$ are given by the solution to the following one-particle Schrödinger equation

$$\hat{H}(i)\psi_{\alpha_i}(i) = E_{\alpha_i}\psi_{\alpha_i}(i) \tag{5.14}$$

where i labels the particles (and is in the range $1 \to N$), while α_i labels the different one-particle eigenstates (and ranges from $1 \to \infty$). We can use these one-particle eigenfunctions to construct the $N-$particle wave function in various ways.

5.2.3 Neglecting the symmetry of the many-body wave function

If we set

$$\Psi(1,2,\ldots,N) = \psi_{\alpha_1}(1)\psi_{\alpha_2}(2)\cdots\psi_{\alpha_N}(N) \tag{5.15}$$

then

$$\begin{aligned}
\hat{H}\Psi &= \sum_{i=1}^{N} \hat{H}(i)\Psi \\
&= \sum_{i=1}^{N} \hat{H}(i)\Big[\psi_{\alpha_1}(1)\psi_{\alpha_2}(2)\cdots\psi_{\alpha_N}(N)\Big] \\
&= \sum_{i=1}^{N} \psi_{\alpha_1}(1)\cdots\Big[\hat{H}(i)\psi_{\alpha_i}(i)\Big]\cdots\psi_{\alpha_N}(N) \\
&= \sum_{i=1}^{N} \psi_{\alpha_1}(1)\cdots\Big[E_{\alpha_i}\psi_{\alpha_i}(i)\Big]\cdots\psi_{\alpha_N}(N) \\
&= \sum_{i=1}^{N} E_{\alpha_i}\Psi \\
&= E\Psi
\end{aligned} \tag{5.16}$$

which implies $E = \sum_{i=1}^{N} E_{\alpha_i}$. These states are in general neither symmetric nor antisymmetric, and therefore cannot describe real systems of identical particles.

5.3 Fermions

For two fermions, the antisymmetric state is

$$\Psi_A(1,2) = \frac{1}{\sqrt{2}}\Big[\psi_{\alpha_1}(1)\psi_{\alpha_2}(2) - \psi_{\alpha_1}(2)\psi_{\alpha_2}(1)\Big] \tag{5.17}$$

Problem 1: Show that if the $\psi_{\alpha_i}(i)$'s are orthonormal, then $\Psi_A(1,2)$ defined above is normalised to one.

In general, for N identical fermions the many-body wave function is given by

$$\begin{aligned}\Psi_A(1,\cdots N) &= \frac{1}{\sqrt{N!}}\sum_P(-1)^P\hat{P}\Big[\psi_{\alpha_1}(1)\psi_{\alpha_2}(2)\cdots\psi_{\alpha_N}(N)\Big] \\ &= \frac{1}{\sqrt{N!}}\begin{vmatrix}\psi_{\alpha_1}(1) & \psi_{\alpha_1}(2) & \cdots & \psi_{\alpha_1}(N) \\ \vdots & \vdots & \vdots & \vdots \\ \psi_{\alpha_N}(1) & \psi_{\alpha_N}(2) & \cdots & \psi_{\alpha_N}(N)\end{vmatrix}\end{aligned} \tag{5.18}$$

The normalisation factor of $1/\sqrt{N!}$ is introduced because the determinant consists of $N!$ mutually orthogonal terms. (Actually orthogonality is not really necessary, all that is required is that the $\psi_{\alpha_i}(i)$'s are **linearly independent**.) Such determinants of one-particle states are called **Slater determinants**. Since the interchange of any two particle coordinates corresponds to interchanging columns in the Slater determinant (which introduces a factor of –1), the antisymmetry is already built in. Also, the $\Psi_A(1,\cdots N)$ defined above is zero if we set $\alpha_i = \alpha_j$ for any $i \neq j$, i.e. no state can be multiply occupied (this is another statement of the Pauli exclusion principle).

5.4 Bosons

For two bosons, the symmetric state for $\alpha_1 \neq \alpha_2$ is

$$\Psi_S(1,2) = \frac{1}{\sqrt{2}}\Big[\psi_{\alpha_1}(1)\psi_{\alpha_2}(2) + \psi_{\alpha_1}(2)\psi_{\alpha_2}(1)\Big] \tag{5.19}$$

whereas for $\alpha_1 = \alpha_2$ we have simply

$$\Psi_S(1,2) = \psi_{\alpha_1}(1)\psi_{\alpha_1}(2) \tag{5.20}$$

In general, for N bosons the many-body wave function is defined as

$$\Psi_S(1,\cdots N) = \sqrt{\frac{N_1!N_2!\ldots}{N!}}\sum_{P'}\hat{P}'\Big[\psi_{\alpha_1}(1)\psi_{\alpha_2}(2)\cdots\psi_{\alpha_N}(N)\Big] \tag{5.21}$$

where N_1 is the multiplicity of state α_1 etc. The summation $\sum_{P'}$ is only over those permutations leading to *distinct* terms and includes $N!/N_1!N_2!\dots$ different terms.

5.5 Exchange forces

We now examine the physical impact of the symmetrization requirement for identical particles. Consider a simple one-dimensional case with two non-interacting identical particles, one in state ψ_a and the other in state ψ_b with these two states being orthonormal.

If the two particles are **distinguishable**, with particle 1 in state $\psi_a(x_1)$ and particle 2 in state $\psi_b(x_2)$, then the total wave function for the two particle system is

$$\Psi(x_1, x_2) = \psi_a(x_1)\psi_b(x_2) \tag{5.22}$$

If the two particles are **identical bosons**, then the total normalised wave function is

$$\Psi_S(x_1, x_2) = \frac{1}{\sqrt{2}}\Big[\psi_a(x_1)\psi_b(x_2) + \psi_a(x_2)\psi_b(x_1)\Big] \tag{5.23}$$

and if they are **identical fermions** it is

$$\Psi_A(x_1, x_2) = \frac{1}{\sqrt{2}}\Big[\psi_a(x_1)\psi_b(x_2) - \psi_a(x_2)\psi_b(x_1)\Big] \tag{5.24}$$

What is the expectation value of the square of the separation distance between the two particles X_{12} ? Here X_{12} is defined as

$$\begin{aligned} X_{12} &= \langle (x_1 - x_2)^2 \rangle \\ &= \int dx_1 dx_2 \Psi^*(x_1, x_2)(x_1 - x_2)^2 \Psi(x_1, x_2) \\ &= \langle x_1^2 \rangle + \langle x_2^2 \rangle - 2\langle x_1 x_2 \rangle \end{aligned} \tag{5.25}$$

where the expectation value is with respect to the total two particle wave function. There are three cases.

Case 1: Distinguishable particles

Using (5.22), we have

$$\begin{aligned} \langle x_1^2 \rangle &= \int dx_1 dx_2 \Psi^*(x_1, x_2) x_1^2 \Psi(x_1, x_2) \\ &= \int x_1^2 |\psi_a(x_1)|^2 dx_1 \int |\psi_b(x_2)|^2 dx_2 \\ &= \langle x^2 \rangle_a \end{aligned} \tag{5.26}$$

i.e. this is the expectation value of x^2 in the one-particle state ψ_a. (Also remember that ψ_a and ψ_b are normalised). Similarly,

$$\begin{aligned}\langle x_2^2\rangle &= \int x_2^2|\psi_b(x_2)|^2dx_2\int|\psi_a(x_1)|^2dx_1 \\ &= \langle x^2\rangle_b \qquad (5.27)\end{aligned}$$

and

$$\begin{aligned}\langle x_1x_2\rangle &= \int x_1|\psi_a(x_1)|^2dx_1\int x_2|\psi_b(x_2)|^2dx_2 \\ &= \langle x\rangle_a\langle x\rangle_b \qquad (5.28)\end{aligned}$$

which implies that

$$X_{12} = \langle x^2\rangle_a + \langle x^2\rangle_b - 2\langle x\rangle_a\langle x\rangle_b \qquad (5.29)$$

Note that the above result would have been the same if particle 1 had been in state ψ_b and particle 2 had been in ψ_a.

Case 2: Identical particles Using (5.23) and (5.24) we have

$$\begin{aligned}\langle x_1^2\rangle &= \frac{1}{2}\Big[\int x_1^2|\psi_a(x_1)|^2dx_1\int|\psi_b(x_2)|^2dx_2\Big] \\ &+ \frac{1}{2}\Big[\int x_1^2|\psi_b(x_1)|^2dx_1\int|\psi_a(x_2)|^2dx_2\Big] \\ &\pm \frac{1}{2}\Big[\int x_1^2\psi_a^*(x_1)\psi_b(x_1)dx_1\int\psi_b^*(x_2)\psi_a(x_2)dx_2\Big] \\ &\pm \frac{1}{2}\Big[\int x_1^2\psi_b^*(x_1)\psi_a(x_1)dx_1\int\psi_a^*(x_2)\psi_b(x_2)dx_2\Big] \\ &= \frac{1}{2}\Big[\langle x^2\rangle_a + \langle x^2\rangle_b \pm 0 \pm 0\Big] \\ &= \frac{1}{2}\Big(\langle x^2\rangle_a + \langle x^2\rangle_b\Big) \qquad (5.30)\end{aligned}$$

where the + sign corresponds to bosons and the − sign to fermions. Similarly,

$$\langle x_2^2\rangle = \frac{1}{2}\Big(\langle x^2\rangle_b + \langle x^2\rangle_a\Big) \qquad (5.31)$$

Problem 2: Verify (5.31) above. (Note that you would expect $\langle x_1\rangle = \langle x_2\rangle$ since one can't tell the particles apart.) Then show that

$$\langle x_1x_2\rangle = \langle x\rangle_a\langle x\rangle_b \pm |\langle x\rangle_{ab}|^2 \qquad (5.32)$$

where

$$\langle x \rangle_{ab} = \int x \psi_a^*(x) \psi_b(x) dx \tag{5.33}$$

Therefore, we have

$$X_{12} = \begin{cases} \langle x^2 \rangle_a + \langle x^2 \rangle_b - 2\langle x \rangle_a \langle x \rangle_b & \text{distinguishable particles} \\ \langle x^2 \rangle_a + \langle x^2 \rangle_b - 2\langle x \rangle_a \langle x \rangle_b - 2|\langle x \rangle_{ab}|^2 & \text{bosons} \\ \langle x^2 \rangle_a + \langle x^2 \rangle_b - 2\langle x \rangle_a \langle x \rangle_b + 2|\langle x \rangle_{ab}|^2 & \text{fermions} \end{cases}$$

This indicates that identical bosons tend to be somewhat closer together and identical fermions somewhat farther apart than distinguishable particles in the same pair of states. Note that $\langle x \rangle_{ab}$ vanishes unless the two wave functions actually overlap. As a *practical* matter, it is reasonable to assume that identical particles with non-overlapping wave functions (if for example they are very far apart) are distinguishable. If there is some overlap of the wave functions, then the system behaves as though there were a "force of attraction" between identical bosons and a "force of repulsion" between identical fermions (which tends to keep them apart). It is called the **exchange force**, although it is not really a force at all as no physical agency is acting on the particles; rather it is a purely *geometrical* consequence of the symmetrization requirement. It is a strictly quantum mechanical phenomenon with no classical counterpart.

Problem 3: Imagine two noninteracting particles each of mass m in an infinite square well. If one of the particles is in state ψ_n where

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right)$$

and the other is in state ψ_m, calculate X_{12} assuming (a) they are distinguishable, (b) they are identical bosons and (c) they are identical fermions.

Answer: To be done in class.

5.6 Helium atom

The accurate determination of the properties of atoms is a difficult task. The two main sources of the difficulty are:

- the Coulomb interaction between the electrons, and
- the spin-orbit interaction between the electron spins and the electric fields in the atom.

In this section we deal with the simplest atom after hydrogen, the helium atom. (Note that other two-electron ions such as H^-, Li^+, Be^{++} etc. can be treated in the same way.) The spin-orbit interaction can be regarded as a weak perturbation; to a first approximation it can be neglected entirely and we only consider the Coulomb interactions between the electrons and the between electrons and the positively charged nuclei. The Hamiltonian (in atomic units) is given by

$$\hat{H} = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 - \frac{Z}{r_1} - \frac{Z}{r_2} + \frac{1}{r_{12}} \tag{5.34}$$

where for generality we write the Hamiltonian in terms of Z rather then setting it to 2 for the specific case of helium. In the equation above, the first two terms correspond to the kinetic energies of the two electrons, the next two terms are the electron-nuclei coulombic attraction, and the last term is the coulombic repulsion between the electrons. We will use both perturbation theory and the variational method to "solve" the time-independent Schrödinger equation to determine the ground state of this atom.

5.6.1 Ground state

We apply perturbation theory to this problem by considering the inter-electronic repulsion term to be a perturbation. In this case, the unperturbed hamiltonian $\hat{H}_0$ is given by

$$\hat{H}_0 = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 - \frac{Z}{r_1} - \frac{Z}{r_2} \tag{5.35}$$

and the zeroth-order ground state wave function by

$$\psi^{(0)}(\mathbf{r}_1, \mathbf{r}_2) = \psi_{1s}(\mathbf{r}_1)\,\psi_{1s}(\mathbf{r}_2) \tag{5.36}$$

$$\psi_{1s}(\mathbf{r}) = \left(\frac{Z^3}{\pi}\right)^{\frac{1}{2}} e^{-Zr} \tag{5.37}$$

Here we are using the fact that in the ground state the two electrons are a state of total spin $S = 0$, which is already antisymmetric with respect to interchange of their spins (see Chapter 2), so that they can have the same *spatial* wave function.

Since the ground state energy of a hydrogen-like system is $-Z^2/2$, the energy $E^{(0)}$ in atomic units is given by

$$E^{(0)} = -\frac{Z^2}{2} - \frac{Z^2}{2} = -Z^2 \tag{5.38}$$

The first-order correction to $E^{(0)}$ (see Chapter 3) is

$$\begin{aligned} E^{(1)} &= \int\int d\mathbf{r}_1\, d\mathbf{r}_2\, \psi^{(0)*}(\mathbf{r}_1, \mathbf{r}_2)\frac{1}{r_{12}}\psi^{(0)}(\mathbf{r}_1, \mathbf{r}_2) \\ &= \int\int d\mathbf{r}_1\, d\mathbf{r}_2\, \psi_{1s}(\mathbf{r}_1)\,\psi_{1s}(\mathbf{r}_2)\,\frac{1}{r_{12}}\,\psi_{1s}(\mathbf{r}_1)\,\psi_{1s}(\mathbf{r}_2) \end{aligned} \tag{5.39}$$

The integral is evaluated in the problem below.

Problem 4: We evaluate the integral in (5.39) as follows. Let $\mathbf{r}_1$ and $\mathbf{r}_2$ be the radius vectors of electron 1 and 2 respectively and let θ be the angle between these two vectors. In general, this is *not* the polar angle from the z-axis but if we choose on of the radius vectors, say $\mathbf{r}_1$, to be the z-axis, then we can treat θ as the polar angle in spherical coordinates. From the cosine rule, we have

$$r_{12} = \left(r_1^2 + r_2^2 - 2r_1 r_2 \cos\theta\right)^{\frac{1}{2}}$$

Show that $E^{(1)}$ becomes

$$\begin{aligned} E^{(1)} &= \frac{Z^6}{\pi^2}\int_0^\infty e^{-2Zr_1}\; 4\pi r_1^2 dr_1 \int_0^\infty e^{-2Zr_2} r_2^2 dr_2 \\ &\times \int_0^{2\pi} d\phi \int_0^\pi \frac{\sin\theta\; d\theta}{(r_1^2 + r_2^2 - 2r_1 r_2\cos\theta)^{\frac{1}{2}}} \end{aligned} \tag{5.40}$$

Letting $x = \cos\theta$, show that the integral over θ is

$$\begin{aligned} \int_0^\pi \frac{\sin\theta\; d\theta}{(r_1^2 + r_2^2 - 2r_1 r_2\cos\theta)^{\frac{1}{2}}} &= \int_{-1}^1 \frac{dx}{(r_1^2 + r_2^2 - 2r_1 r_2 x)^{\frac{1}{2}}} \\ &= \frac{2}{r_1} \qquad r_1 > r_2 \\ &= \frac{2}{r_2} \qquad r_1 < r_2 \end{aligned} \tag{5.41}$$

Substituting this result into $E^{(1)}$, show that

$$\begin{aligned} E^{(1)} &= 16Z^6 \int_0^\infty e^{-2Zr_1}\; r_1^2\; dr_1 \left(\frac{1}{r_1}\int_0^{r_1} e^{-2Zr_2} r_2^2\; dr_2 + \int_{r_1}^\infty e^{-2Zr_2} r_2\; dr_2\right) \\ &= 4Z^3 \int_0^\infty e^{-2Zr_1}\; r_1^2\; dr_1 \left[\frac{1}{r_1} - e^{-2Zr_1}\left(Z + \frac{1}{r_1}\right)\right] \\ &= \frac{5}{8} Z \end{aligned} \tag{5.42}$$

The final result is

$$E^{(1)} = \frac{5}{8}\, Z \tag{5.43}$$

so that the ground state energy of the helium atom through first order is

$$E = E^{(0)} + E^{(1)} + \cdots = -Z^2 + \frac{5}{8}\, Z + \cdots \tag{5.44}$$

Setting $Z = 2$, we find that $E = -\frac{11}{4} = -2.75$ or -74.83 eV, while the experimental value is -79.00 eV or -2.9033 au. This means that first-order perturbation theory already gives a result within 5 % of the experimental (exact) value. Note that more extensive perturbation calculations to many orders by Scheer and Knight (Rev. Mod. Phys. **35**, 426 (1963)) give

$$E = -Z^2 + \frac{5}{8}\,Z - 0.157666254 + \frac{0.008698679}{Z} + \cdots \tag{5.45}$$

which yields a value of -2.9037 au, in excellent agreement with experiment.

Problem 5: In this problem, we study how to get the ground state energy of the helium atom using the variational method. We start with the trial wave function (5.36) with Z as a variational parameter. Note that this wave function is already normalised. Thus we need to evaluate

$$E = \int d\mathbf{r}_1\, d\mathbf{r}_2\; \psi^* \hat{H} \psi$$

with the Hamiltonian defined above. The evaluation of the integral is greatly simplified if one recalls that the ψ_{1s} is an eigenfunction of a hydrogen-like Hamiltonian with a nucleus having a charge Z. It is convenient to rewrite this Hamiltonian as

$$\hat{H} = -\frac{1}{2}\nabla_1^2 - \frac{Z}{r_1} - \frac{1}{2}\nabla_2^2 - \frac{Z}{r_2} + \frac{Z-2}{r_1} + \frac{Z-2}{r_2} + \frac{1}{r_{12}} \tag{5.46}$$

where

$$\left(-\frac{1}{2}\nabla_1^2 - \frac{Z}{r_1}\right)\left(\frac{Z^3}{\pi}\right)^{\frac{1}{2}} e^{-Zr} = -\frac{Z^2}{2}\left(\frac{Z^3}{\pi}\right)^{\frac{1}{2}} e^{-Zr} \tag{5.47}$$

Show that

$$\begin{aligned} E(Z) &= \frac{Z^6}{\pi^2}\int\!\!\int e^{-Z(r_1+r_2)}\Bigg[-\frac{Z^2}{2} - \frac{Z^2}{2} \\ &+ \frac{Z-2}{r_1} + \frac{Z-2}{r_2} + \frac{1}{r_{12}}\Bigg] e^{-Z(r_1+r_2)}\, d\mathbf{r}_1\, d\mathbf{r}_2 \\ &= -Z^2 + 2(Z-2)\frac{Z^3}{\pi}\int d\mathbf{r}\,\frac{e^{-2Zr}}{r} + \frac{5}{8}Z \\ &= Z^2 - \frac{27}{8}Z \end{aligned} \tag{5.48}$$

Now minimise E with respect to Z and show that

$$E = -2.8477 \text{ au}$$

with $Z_{min} = \frac{27}{16}$. This value can be interpreted as an *effective nuclear charge*. The fact that this value is less than 2 reflects the fact that each electron partially screens the nucleus from the other.

While the agreement between the above result and experiment looks impressive, it can be misleading. Consider the ionisation energy (IE) for example, which is the difference between the energy of a singly ionised helium atom and the energy of the helium atom in its ground state. From our calculation above we get this to be IE $= -2.0 - (-\frac{11}{4}) = 0.750$ au $= 20.4$ eV, while the experimental value is 24.6 eV. This is indeed a big difference and we should be able to do better. One way is to use a more general form of the trial wave function (still using the variational method). We could go beyond the one-electron product form used above and use a functional form that includes the inter-electronic distance *explicitly.* This was first done by Hylleras in 1930, who used

$$\psi(r_1, r_2, r_{12}) = e^{-Zr_1}\, e^{-Zr_2}\Big[1 + g(r_1, r_2, r_{12})\Big] \tag{5.49}$$

with g chosen to be a polynomial in r_1, r_2, r_{12} and whose expansion coefficients are variationally optimised. Using 14 terms for g, he obtained a value of -2.9037 for the ground state energy (exact to five significant figures). A more extensive calculation using this approach, by Pekeris in 1959 yielded -2.903724375 au with 1078 terms!

While this showed that one could obtain essentially exact energies using the variational method with more elaborate trial wave functions in which the inter-electronic distance r_{12} is incorporated explicitly, this approach is computationally very demanding and does not lend itself to the study of large atoms and molecules. More state-of-the-art methods will be discussed later.

5.7 Hydrogen molecule

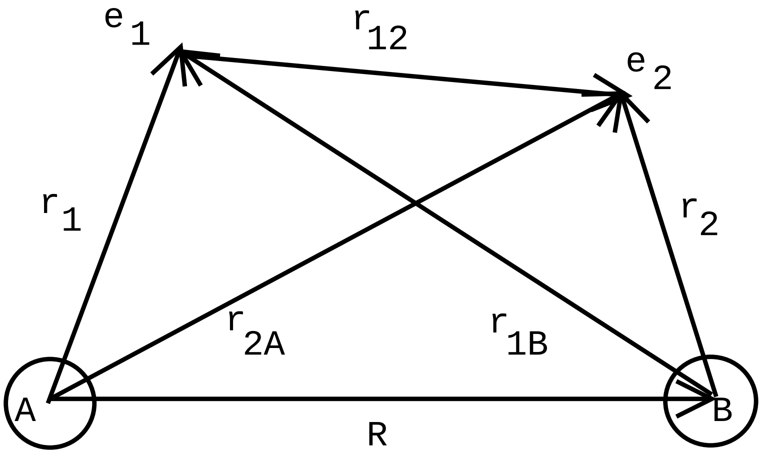
We now consider the ground state of the hydrogen molecule, the main point of this calculation being to study the effect of the exchange energy on the **chemical bonding** between atoms. This energy is entirely a consequence of the antisymmetry of the electron wave function under the interchange of the two electrons.

The hydrogen molecule consists of two hydrogen atoms bound together. Since the protons are much more massive than the electrons, we neglect their motion and treat them as fixed centres of force (i.e. the **Born-Oppenheimer** approximation.) With this approximation, the Hamiltonian for a hydrogen molecule becomes

$$\hat{H}(1,2) = -\frac{\hbar^2}{2m}\nabla_1^2 - \frac{e^2}{r_1} - \frac{\hbar^2}{2m}\nabla_2^2 - \frac{e^2}{r_2} + \frac{e^2}{r_{12}} + \frac{e^2}{R} - \frac{e^2}{r_{1B}} - \frac{e^2}{r_{2A}} \tag{5.50}$$

where the first and third term corresponds to the kinetic energies of the electrons; the second and fourth terms correspond to the potential energy of interaction between electron 1 and nucleus A

and electron 2 and nucleus B, respectively; the fifth and sixth term correspond to the potential energy of inter-electronic and inter-nuclear repulsions; the seventh and eighth term corresponds to the potential energy of interaction between electron 1 and nucleus B and electron 2 and nucleus A respectively, (the various coordinates are labelled in the figure below).



We write the Hamiltonian as

$$\hat{H} = \hat{H}_0 + \hat{W} \tag{5.51}$$

where $\hat{H}_0$ is the "hydrogen atom" part of the Hamiltonian

$$-\frac{\hbar^2}{2m}\nabla_1^2 - \frac{e^2}{r_1} - \frac{\hbar^2}{2m}\nabla_2^2 - \frac{e^2}{r_2} \tag{5.52}$$

and

$$\hat{W} = \frac{e^2}{r_{12}} + \frac{e^2}{R} - \frac{e^2}{r_{1B}} - \frac{e^2}{r_{2A}} \tag{5.53}$$

Here the two-body potential is $\frac{e^2}{r_{12}}$, whereas $-(\frac{e^2}{r_1} + \frac{e^2}{r_{1B}})$ is the "external" potential for particle 1 and $-(\frac{e^2}{r_2} + \frac{e^2}{r_{2A}})$ is the "external" potential for particle 2. The repulsive potential $\frac{e^2}{R}$ between the two protons is just a constant i.e. independent of the electron positions. To determine the ground state energy of the hydrogen molecule, we treat $\hat{W}$ as a perturbation. In this case we only need to find a good approximation for the ground state wave function. The energy is then approximated by the expectation value of $\hat{H}$ in this approximate ground state. For R sufficiently large, we have two unperturbed hydrogen atoms with the Hamiltonian $\hat{H}_0$, and we use the eigenfunctions of $\hat{H}_0$ as our approximate ground state.

Since the electrons are fermions, the total wave function (which must be a direct product of the spatial and spin parts), must be antisymmetric in the particle labels (which include the spin degrees of freedom). From the discussion of the addition of angular momenta for two spin$-\frac{1}{2}$ particles in

Chapter 2, we know that the possible values of the total spin angular momentum for the electrons are the $S = 1$ *triplet* state, $\chi^{(t)}$ where

$$\chi^{(t)} = \begin{cases} |1,1\rangle = |++\rangle \\ |1,0\rangle = \frac{1}{\sqrt{2}}(|+-\rangle + |-+\rangle) \\ |1,-1\rangle = |--\rangle \end{cases}$$

and the $S = 0$ *singlet* state, $\chi^{(s)}$ with

$$\chi^{(s)} = |0,0\rangle = \frac{1}{\sqrt{2}}(|+-\rangle - |-+\rangle) \tag{5.54}$$

Note that the triplet state is even under the interchange of the spin coordinates while the singlet state is odd. The corresponding spatial wave functions are therefore

$$\begin{aligned} \Psi_s(\mathbf{r}_1, \mathbf{r}_2) &= A_s\Big[\psi_A(1)\ \psi_B(2) + \psi_A(2)\ \psi_B(1)\Big] \\ \Psi_t(\mathbf{r}_1, \mathbf{r}_2) &= A_t\Big[\psi_A(1)\ \psi_B(2) - \psi_A(2)\ \psi_B(1)\Big] \end{aligned} \tag{5.55}$$

where A_s and A_t are normalisation constants, and the $\psi_A(i)$, $\psi_B(i)$ represent hydrogen atom ground state wave functions centred at A and B respectively. Thus

$$\begin{aligned} \psi_A(1) &= \frac{1}{\sqrt{\pi a^3}} e^{-r_1/a}\ , \quad \psi_A(2) = \frac{1}{\sqrt{\pi a^3}} e^{-r_{2A}/a} \\ \psi_B(1) &= \frac{1}{\sqrt{\pi a^3}} e^{-r_{1B}/a}\ , \quad \psi_B(2) = \frac{1}{\sqrt{\pi a^3}} e^{-r_2/a} \end{aligned} \tag{5.56}$$

where a is the Bohr radius, $a = \hbar^2/2me^2$.

Problem 6: Show that the normalisations A_s and A_t in equations (5.55) are given by

$$\begin{aligned} A_s &= \Big[2(1 + A^2)\Big]^{-\frac{1}{2}} \\ A_t &= \Big[2(1 - A^2)\Big]^{-\frac{1}{2}} \end{aligned} \tag{5.57}$$

where the **overlap integral**, A, is given by

$$A = \frac{1}{\pi a^3} \int d\mathbf{r}_1 e^{-(r_1 + r_{1B})/a} = \langle \psi_A | \psi_B \rangle \tag{5.58}$$

To evaluate integral (5.58), we transform to elliptical coordinates

$$\begin{aligned} \xi &= \frac{r_1 + r_{1B}}{R} \\ \eta &= \frac{r_1 - r_{1B}}{R} \end{aligned} \tag{5.59}$$

and ϕ (which is the angle of rotation about the line joining the two protons. The volume element in these coordinates is

$$d\mathbf{r}_1 = \frac{R^3}{8}\,(\xi^2 - \eta^2)\; d\xi\; d\eta\; d\phi \tag{5.60}$$

with the range of integration being

$$1 \le \xi < \infty, \quad -1 \le \eta \le +1, \quad 0 \le \phi \le 2\pi \tag{5.61}$$

Problem 7: Verify (5.60) and, using (5.60) and (5.61), show that

$$A = \left[1 + \frac{R}{a} + \frac{1}{3}\left(\frac{R}{a}\right)^2\right] e^{-R/a} \tag{5.62}$$

To the extent that Ψ_s and Ψ_t are good wave functions, the ground state energy of the hydrogen molecule can be approximated in the singlet and triplet states by

$$\begin{aligned} E_s(R) &= \int \Psi_s^*(\mathbf{r}_1, \mathbf{r}_2)\; \hat{H}\; \Psi_s(\mathbf{r}_1, \mathbf{r}_2)\; d\mathbf{r}_1\; d\mathbf{r}_2 \\ E_t(R) &= \int \Psi_t^*(\mathbf{r}_1, \mathbf{r}_2)\; \hat{H}\; \Psi_t(\mathbf{r}_1, \mathbf{r}_2)\; d\mathbf{r}_1\; d\mathbf{r}_2 \end{aligned} \tag{5.63}$$

These integrals can be rewritten with a little algebra in the form

$$\begin{aligned} E_s(R) &= \frac{J+K}{1+A^2} \\ E_t(R) &= \frac{J-K}{1-A^2} \end{aligned} \tag{5.64}$$

where

$$\begin{aligned} J &= \int |\psi_A(1)\psi_B(2)|^2 \left[\frac{e^2}{r_{12}} + \frac{e^2}{R} - \frac{e^2}{r_{1B}} - \frac{e^2}{r_{2A}}\right] d\mathbf{r}_1\; d\mathbf{r}_2 \\ &= \frac{e^2}{R} - \int |\psi_A(1)|^2 \frac{e^2}{r_{1B}} d\mathbf{r}_1 - \int |\psi_B(2)|^2 \frac{e^2}{r_{2A}}\; d\mathbf{r}_2 \\ &+ \int |\psi_A(1)|^2 \frac{e^2}{r_{12}} |\psi_B(2)|^2\; d\mathbf{r}_1\; d\mathbf{r}_2 \end{aligned} \tag{5.65}$$

and

$$\begin{aligned} K &= \int \psi_A(1)\; \psi_B(2) \left[\frac{e^2}{r_{12}} + \frac{e^2}{R} - \frac{e^2}{r_{1B}} - \frac{e^2}{r_{2A}}\right] \psi_A(2)\; \psi_B(1)\; d\mathbf{r}_1\; d\mathbf{r}_2 \\ &= \frac{e^2}{R} A^2 - A \int \psi_B(2) \frac{e^2}{r_{2A}} \psi_A(2) d\mathbf{r}_2 - A \int \psi_A(1) \frac{e^2}{r_{1B}} \psi_B(1) d\mathbf{r}_1 \\ &+ \int \psi_A(1)\psi_B(2) \frac{e^2}{r_{12}} \psi_A(2)\psi_B(1)\; d\mathbf{r}_1\; d\mathbf{r}_2 \end{aligned} \tag{5.66}$$

Problem 8: Work out the steps leading to (5.64) - (5.66).

The integral J (5.65) is called the **Coulomb interaction integral** and contains the following terms. The first term gives the Coulomb repulsion of the two protons. The second term gives the energy due to the interaction of the proton at B with the "charge density" $-e|\psi_A(1)|^2$ due to electron 1, while the third term gives the interaction energy between the proton at A and the "charge density" $-e|\psi_B(2)|^2$ due to electron 2. (Since protons, like electrons, are identical particles, the second and third terms must be equal.) The last term gives the interaction between the two "charge densities" $-e|\psi_A(1)|^2$ and $-e|\psi_B(2)|^2$ centred at A and B. The integral K is called the **exchange integral** and results strictly from the indistinguishability of the two electrons. It is this type of term that gives rise to covalent bonding, ferromagnetism, etc.

Problem 9: The various integrals in (5.65) and (5.66) can be evaluated by the use of elliptical coordinates.

- (a) Verify that

$$\int |\psi_A(1)|^2 \frac{e^2}{r_{1B}}\, d\mathbf{r}_1 = \frac{\pi a^3}{R}\left[1 - e^{-2R/a}\left(1 + \frac{R}{a}\right)\right] \tag{5.67}$$

- (b) The last term in (5.65) can be written as

$$\int |\psi_A(1)|^2 \frac{e^2}{r_{12}} |\psi_B(2)|^2 \; d\mathbf{r}_1\, d\mathbf{r}_2 = \int \rho_A(1)\Phi_B(1)\, d\mathbf{r}_1 \tag{5.68}$$

  where

$$\rho_A(1) = -e\,|\psi_A(1)|^2 \tag{5.69}$$

  is the "charge density" due to electron 1 centred at A and $\Phi_B(1)$ is the "potential" at r_1 due to electron 2 with charge density $-e|\psi_B(2)|^2$ centred at B. Thus $\Phi_B(1)$ satisfies the Poisson equation

$$\nabla^2\Phi_B(\mathbf{r}_1) = 4\pi e\,|\psi_B(2)|^2 \tag{5.70}$$

  Evaluate this potential and hence show that

$$J = \frac{e^2}{R}e^{-2R/a}\left[1 + \frac{5}{8} - \frac{3}{4}\left(\frac{R}{a}\right)^2 - \frac{1}{6}\left(\frac{R}{a}\right)^3\right] \tag{5.71}$$

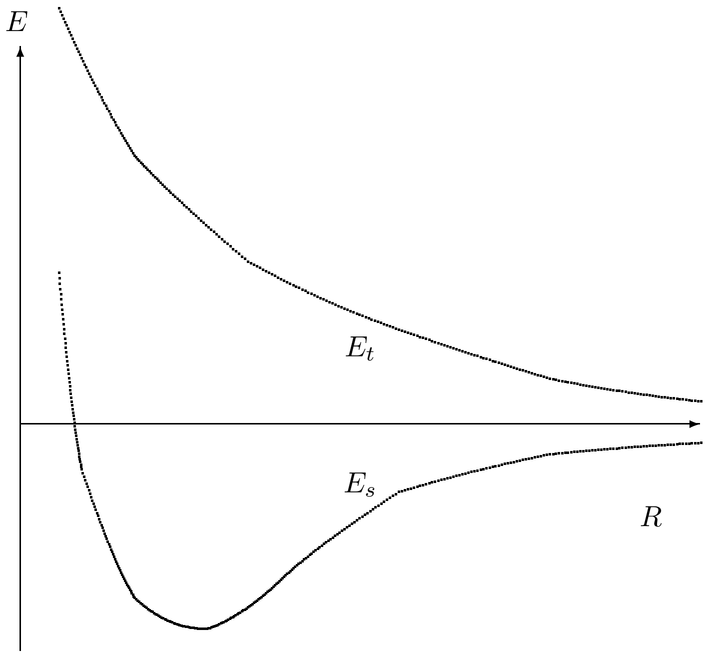
For the exchange integral K, the terms involving only one integration are again obtained by going to elliptical coordinates. The last term cannot be expressed in terms of elementary functions. It can, however, be expressed in terms of the exponential integral

$$\mathrm{Ei}(x) = -\int_{-x}^{\infty} \frac{e^{-t}}{t} dt \tag{5.72}$$

whose properties can be found on p.228 of the "Handbook of Mathematical Functions" by Abramowitz and Stegun (in the Rayleigh Library). The final result for K (which you are welcome to verify but the details can be found in the paper by Sugiura, **Z. Physik, Vol. 45**, p.484 (1927)) is:

$$\begin{aligned} K &= \frac{e^2}{R}A^2\left[1+\frac{6}{5}(C+\ln\frac{R}{a})\right] \\ &+ \frac{e^2}{a}e^{-2R/a}\left[\frac{11}{8}+\frac{103}{20}\frac{R}{a}+\frac{49}{15}\left(\frac{R}{a}\right)^2+\frac{11}{15}\left(\frac{R}{a}\right)^3\right] \\ &+ \frac{6}{5}\frac{e^2}{R}e^{2R/a}\left[1-\frac{R}{a}+\frac{1}{3}\left(\frac{R}{a}\right)^2\right]^2\mathrm{Ei}(-4R/a) \\ &- \frac{12}{5}Ae^{R/a}\left[1-\frac{R}{a}+\frac{1}{3}\left(\frac{R}{a}\right)^2\right]\mathrm{Ei}(-2R/a) \end{aligned} \tag{5.73}$$

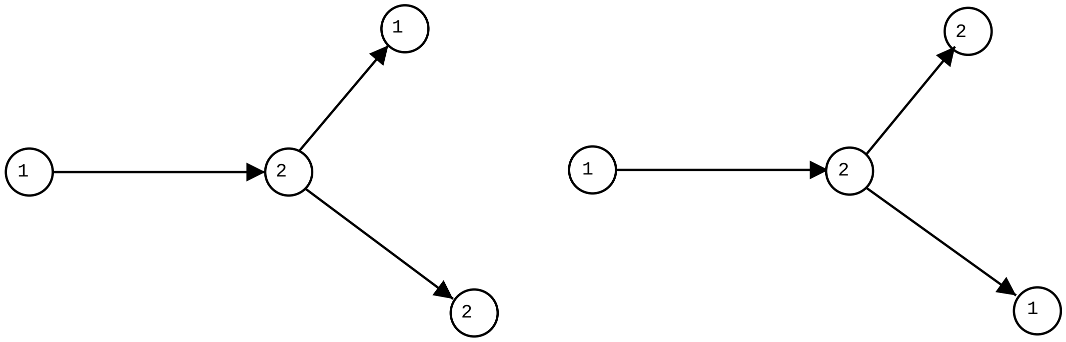
where $C = 0.577215\ldots$ is Euler's constant. A plot of E_s and E_t as functions of R is shown below.



Although quantitative agreement with experimental values for the equilibrium bond length and dissociation energy is not very good, we find that E_t is **always higher** in energy than E_s. In fact the singlet state is called the **bonding state** while the triplet state is called the **anti-bonding** state. Better agreement with experiment is obtained if we take into account the **correlation** between the electrons i.e. the fact that electrons avoid each other as much as possible. One can "build" this into the ground state wave function by constructing the two-electron wave function as a function of r_1, r_2 and r_{12} (with variational parameters) and then using the variational principle to get the best estimate for these parameters.

5.8 Scattering of identical particles

If the particles in the target and the incident beam are identical, it is impossible to tell if we are observing a scattered beam particle or a recoiling target particle, as shown in the following figure:



There is no way of distinguishing between a deflection of a particle through an angle θ and a deflection of $(\pi - \theta)$ in the centre-of-mass frame, since momentum conservation demands that if one of the particles scatters through θ then the other goes in the direction $(\pi - \theta)$. For classical scattering, there is an **increase** in the scattering cross section relative to the non-identical particle case since the number of counts at a certain detector will be the sum due to both particles. The consequence is that the classical cross section $\sigma_{cl}(\theta)$ is symmetric about $\theta = \pi/2$:

$$\sigma_{cl}(\theta) = \sigma(\theta) + \sigma(\pi - \theta) \tag{5.74}$$

In QM there is no way of distinguishing the two final states. For quantum scattering, interference occurs at the scattering amplitude level, i.e. between the scattering amplitudes $f(\theta)$ and $f(\pi - \theta)$, which can result in dramatic changes.

5.8.1 Scattering of identical spin zero bosons

The differential scattering cross section, for the scattering of two identical spin zero bosons (for example, α-particles) is given by

$$\begin{aligned}\frac{d\sigma}{d\Omega} &= |f(\theta)+f(\pi-\theta)|^2 \\ &= |f(\theta)|^2+|f(\pi-\theta)|^2+2Re\Big[f^*(\theta)f(\pi-\theta)\Big] \qquad (5.75)\end{aligned}$$

which differs from the classical result (5.74) by the interference term (the last term on the r.h.s.). It leads to an **enhancement** of the differential cross section at $\theta=\pi/2$

$$\left.\frac{d\sigma}{d\Omega}\right|_{\theta=\pi/2} = 4\times|f(\theta=\pi/2)|^2 \qquad (5.76)$$

as compared to the result that would have been obtained without interference:

$$\left.\frac{d\sigma}{d\Omega}\right|_{\theta=\pi/2} = 2\times|f(\theta=\pi/2)|^2 \qquad (5.77)$$

Problem 10: Show that in the partial wave analysis of the cross section for boson-boson scattering, only the even l values contribute to the differential cross section. (This means that it is impossible to measure or even define odd-l phase shifts for boson-boson scattering).

5.8.2 Scattering of fermions

When the scattering of two spin$-\frac{1}{2}$ particles is considered (for example electron-electron or proton-proton scattering), then the scattering amplitude should reflect the basic antisymmetry of the total wave function under the interchange of the two particles. If the two particles are in a spin singlet state, then the spatial wave function is symmetric (even) and

$$\frac{d\sigma_s}{d\Omega} = |f(\theta)+f(\pi-\theta)|^2 \qquad (5.78)$$

and we have the same situation as the boson case discussed above. If the two particles are in a spin triplet state, then the spatial wave function is antisymmetric and

$$\begin{aligned}\frac{d\sigma_t}{d\Omega} &= |f(\theta)-f(\pi-\theta)|^2 \\ &= |f(\theta)|^2+|f(\pi-\theta)|^2-2Re\Big[f^*(\theta)f(\pi-\theta)\Big] \qquad (5.79)\end{aligned}$$

In the scattering of two unpolarised protons for example (where all spin states are equally likely), the probability of finding two protons in a triplet state is three times as large as finding them in the singlet state. Therefore the differential cross section for this case is

$$\begin{aligned} \frac{d\sigma}{d\Omega} &= \frac{3}{4}\frac{d\sigma_t}{d\Omega} + \frac{1}{4}\frac{d\sigma_s}{d\Omega} \\ &= |f(\theta)|^2 + |f(\pi-\theta)|^2 - Re\Big[f^*(\theta)f(\pi-\theta)\Big] \qquad (5.80) \end{aligned}$$

Analogously to the boson case discussed in the last problem, the singlet term in (5.80) is a sum over only even l's while the triplet term is a sum over the odd l's. The resulting differential cross section therefore has mixed symmetry with respect to θ.

5.9 *Modern electronic structure theory

In this section we review the theoretical foundations of some of the computational techniques that are currently used in electronic structure calculations for *real* systems. The relative merits of the different techniques for calculating the electronic structure of atoms, molecules and solids are considered. The progress made over the past decade in developing quantum Monte Carlo (QMC) methods as a tool for tackling realistic continuum electronic structure problems is described.

5.9.1 *The many-electron problem

Within the Born-Oppenheimer approximation, the time independent Schrödinger equation for a fully interacting many-electron system is

$$\hat{H}\Psi = \sum_{i=1}^{N}\left(-\frac{\hbar^2}{2m}\nabla_i^2 - \sum_{\alpha}\frac{Z_\alpha e^2}{4\pi\epsilon_0|\vec{r}_i - \vec{d}_\alpha|}\right)\Psi + \frac{1}{2}\sum_i\sum_{j\neq i}\frac{e^2}{4\pi\epsilon_0|\vec{r}_i - \vec{r}_j|}\Psi = E\Psi \quad , \qquad (5.81)$$

where Ψ is the N-electron wave functions, $\vec{r}_i$ are the electron positions, $\vec{d}_\alpha$ are the positions of the ions and Z_α are the ionic charges. This equation is impossible to solve exactly so approximate solutions must be sought. One of the main challenges of condensed matter physics is to try to find good, workable approximations that contain the essence of the physics involved in a particular problem and to obtain the most accurate solutions possible. For the rest of this chapter all equations will be written in atomic units, $e = m_e = \hbar = 4\pi\epsilon_0 = 1$.

5.9.2 *One-electron methods

One of the most common ways of dealing with the many-electron problem is to assume that each electron can be considered separately. Each electron is treated as moving in a mean field potential, $U(\vec{r})$. This potential models the effects of all the other particles in the system, as well as any external potential acting on the system.

The one-electron equations are of the form

$$-\frac{1}{2}\nabla^2\psi_i(\vec{r}) + U(\vec{r})\psi_i(\vec{r}) = \epsilon_i\psi_i(\vec{r}) \quad , \tag{5.82}$$

where $\psi_i(\vec{r})$ is a one-electron wave function and ϵ_i are Lagrange multipliers which arise from the fact that the one-electron wave functions are normalised. Choosing an appropriate $U(\vec{r})$ for the single electron is still a very complicated problem. $U(\vec{r})$ depends upon the interactions between the electrons and therefore on the one-electron wave functions. Since initially neither of these quantities, $U(\vec{r})$ or $\psi_i(\vec{r})$, is known, it is necessary to solve Eq.(5.82) in a self-consistent manner. What this means is that we make a guess for the ψ_i, then construct the U (how one does this is described below) and proceed to solve the above equation to get the new value for ψ_i. One then repeats this procedure until the ψ_i have converged in some prescribed manner.

5.9.3 *Hartree approximation

This approximation starts from the one-electron equations (5.82). $U(\vec{r})$ is chosen to try to model the interaction terms in this equation. The ions contribute a potential

$$U_{ion}(\vec{r}) = -\sum_{\alpha}\frac{Z_\alpha}{|\vec{r}-\vec{d}_\alpha|} \quad . \tag{5.83}$$

All the other electrons in the system also contribute to the potential. The potential due to the electrons is approximated by the electrostatic interaction with all the others, which can be written in terms of the electron density, $\rho(\vec{r})$, as

$$U_H(\vec{r}) = \int d\vec{r}'[\rho(\vec{r}') - \rho_i(\vec{r}')]\frac{1}{|\vec{r}-\vec{r}'|} \quad , \tag{5.84}$$

where the self-interaction potential due to electron i has been removed.

To actually calculate the Hartree potential it is necessary to know the electronic charge distribution of the system. If the electrons are assumed to be independent of each other, then it is straightforward to construct $\rho(\vec{r})$ from the single electron eigenstates

$$\rho(\vec{r}) = \sum_i |\psi_i(\vec{r})|^2 \quad , \tag{5.85}$$

where the summation over i includes all occupied states. Using this charge density the total one-electron potential is

$$U_i(\vec{r}) = U_{ion}(\vec{r}) + \sum_{j\neq i} \int d\vec{r}' |\psi_j(\vec{r}')|^2 \frac{1}{|\vec{r}-\vec{r}'|} \quad . \tag{5.86}$$

The potential $U_i(\vec{r})$ is different for each orbital, and therefore the orbitals are not orthogonal. Note that $U_i(\vec{r})$ depends on all the other orbitals, ψ_j, and so the solution of Eq.(5.82) must be found self-consistently.

The choice of $U(\vec{r})$ in Eq.(5.86) all seems a bit like guesswork, but it can also be derived using the variational principle. If the electrons are assumed to be non-interacting, and so the N-electron wave function is just the product of the one-electron wave functions,

$$\Psi = \psi_1(\vec{r}_1)\psi_2(\vec{r}_2)\ldots\psi_N(\vec{r}_N) \quad . \tag{5.87}$$

then using the variational principle, by minimising the expectation of $\hat{H}$ with respect to the single-particle orbitals, one can derive the Hartree equations:

$$\left(-\frac{1}{2}\nabla^2 + U_{ion}(\vec{r})\right)\psi_i(\vec{r}) + \sum_{j(\neq i)} \int d\vec{r}' \frac{|\psi_j(\vec{r}')|^2}{|\vec{r}-\vec{r}'|}\psi_i(\vec{r}) = \epsilon_i \psi_i(\vec{r}) \quad , \tag{5.88}$$

5.9.4 *Hartree-Fock approximation

The Hartree-Fock approximation is an extension of the above Hartree approximation to include the permutation symmetry of the wave function, which leads to the exchange interaction. Exchange is due to the Pauli exclusion principle, which states that the total wave function for the system must be antisymmetric under particle exchange. This means that when two arguments are swapped the wave function changes sign as follows:

$$\Psi(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i, \ldots, \vec{x}_j, \ldots, \vec{x}_N) = -\Psi(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_j, \ldots, \vec{x}_i, \ldots, \vec{x}_N) \quad , \tag{5.89}$$

where $\vec{x}_i$ includes coordinates of position and spin. Therefore no two electrons can have the same set of quantum numbers, and electrons with the same spin cannot occupy the same spatial state simultaneously.

Instead of using the simple product form of the wave function shown in Eq.(5.87), a Slater determinant wave function which satisfies antisymmetry is used

$$D = \begin{vmatrix} \psi_1(\vec{x}_1) & \psi_1(\vec{x}_2) & \cdots & \psi_1(\vec{x}_N) \\ \psi_2(\vec{x}_1) & \psi_2(\vec{x}_2) & \cdots & \psi_2(\vec{x}_N) \\ \vdots & \vdots & & \\ \psi_N(\vec{x}_1) & \psi_N(\vec{x}_2) & \cdots & \psi_N(\vec{x}_N) \end{vmatrix} \quad , \tag{5.90}$$

where $\psi_i(\vec{x}_j)$ are the one-electron wave functions.

Following exactly the same method of minimising the expectation value of $\hat{H}$ with respect to the one-electron wave functions as was used in the derivation of the Hartree equations, results in the following set of one-electron equations, the Hartree-Fock equations;

$$\begin{aligned} \epsilon_i \psi_i(\vec{r}) &= \left(-\frac{1}{2}\nabla^2 + U_{ion}(\vec{r})\right)\psi_i(\vec{r}) + \sum_j \int d\vec{r}' \frac{|\psi_j(\vec{r}')|^2}{|\vec{r}-\vec{r}'|}\psi_i(\vec{r}) \\ &- \sum_j \delta_{s_i s_j} \int d\vec{r}' \frac{\psi_j^*(\vec{r}')\psi_i(\vec{r}')}{|\vec{r}-\vec{r}'|}\psi_j(\vec{r}) \quad , \end{aligned} \tag{5.91}$$

where s_i labels the spin of particle i. Note the self-interaction cancels out from the second and third terms. The extra term in these equations, when compared to Eq.(5.88), is known as the exchange term and is only non-zero when considering electrons of the same spin. The effect of exchange on the many-body system is that electrons of like spin tend to avoid each other. As a result of this, each electron has a “hole” associated with it which is known as the exchange hole (or the Fermi hole). This is a small volume around the electron which like-spin electrons avoid. The charge contained in the exchange hole is positive and exactly equivalent to the absence of one electron.

Unlike all the other terms acting on ψ_i , the exchange term is a non-local integral operator and this makes the Hartree-Fock equations hard to solve in all but a few special cases.

Hartree-Fock calculations, which include the exchange interaction between electrons, are most useful for performing calculations on relatively small systems as they are considerably more computationally expensive than Hartree and DFT-LDA calculations (see below), due to the non-local exchange term. Even for atoms, however, Hartree-Fock theory is not ideal. For example, H^- is predicted to be unstable in contradiction to reality.

Various improvements to Hartree-Fock theory have been attempted. Unrestricted Hartree-Fock theory ignores some of the simplifying restrictions which are normally applied to Hartree-Fock wave functions. The exchange interaction is allowed to make the spatial parts of spin up and spin down electron wave functions different for the same state. However, although for some systems this results in an improvement, especially for open shell systems, it also sometimes produces worse results than conventional Hartree-Fock theory. In general, Hartree-Fock theory is most useful as a tool for providing qualitative answers. It is also used as the starting point for other methods, such as some quantum Monte Carlo calculations.

5.9.5 *Density functional methods

Computations based on density functional structure theory are the most popular approach to determining the electronic properties of real materials in physics, chemistry and material science. Density functional theory is a formally exact theory based on the charge density of a system. Kohn-Sham density functional theory is a formally exact one-electron theory. Working within the Born-Oppenheimer approximation, the many-body Schrödinger equation,

$$\hat{H}\Psi = \sum_{i=1}^{N}\left(-\frac{1}{2}\nabla_i^2 - \sum_{\alpha}\frac{Z_\alpha}{|\vec{r}_i - \vec{d}_\alpha|}\right)\Psi + \frac{1}{2}\sum_i\sum_{j\neq i}\frac{1}{|\vec{r}_i - \vec{r}_j|}\Psi = E\Psi \quad , \tag{5.92}$$

where Ψ is the many-body wave function, is replaced by a set of N one-electron equations of the form

$$\left(-\frac{1}{2}\nabla^2 + V(\vec{r})\right)\psi_i(\vec{r}) = \epsilon\psi_i(\vec{r}) \quad , \tag{5.93}$$

where $\psi_i(\vec{r})$ is a single-electron wave function. These one-electron equations contain a potential $V(\vec{r})$ produced by all the ions and the electrons. Density functional theory properly includes all parts of the electron-electron interaction, i.e. the Hartree potential

$$V_H(\vec{r}) = \int d\vec{r}'\frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} \quad , \tag{5.94}$$

where ρ is the charge density of all the electrons, a potential due to exchange and correlation effects, $V_{XC}(\vec{r})$, and the external potential due to the ions, $V_{\text{ext}}(\vec{r})$,

$$V(\vec{r}) = V_{\text{ext}}(\vec{r}) + V_H(\vec{r}) + V_{XC}(\vec{r}) \quad . \tag{5.95}$$

Hohenberg and Kohn (*Phys. Rev.* **136**, B864 (1964)) originally developed density functional theory for application to the ground state of a system of 'spinless fermions'. In such a system the particle density is given by

$$\rho(\vec{r}) = N\int |\Psi_o(\vec{r}, \vec{r}_2, \ldots, \vec{r}_N)|^2 d\vec{r}_2 \ldots d\vec{r}_N \quad , \tag{5.96}$$

with Ψ_0 being the many-body ground state wave function of the system. It can be shown that the total ground state energy of the system is a functional of the density, $E[\rho(\vec{r})]$, and that if the energy due to the electron-ion interactions is excluded the remainder of the energy is a universal functional of the density, $F[\rho(\vec{r})]$ (i.e. $F[\rho(\vec{r})]$ does not depend on the potential from the ions).

Kohn and Sham (*Phys. Rev.* **140**, A1133 (1965)) introduced a method based on the Hohenberg-Kohn theorem that enables one to minimise the functional $E[\rho(\vec{r})]$ by varying $\rho(\vec{r})$ over all densities containing N electrons. This constraint is introduced by the Lagrange multiplier, μ, chosen so that $\int \rho(\vec{r})d\vec{r} = N$,

$$\frac{\delta}{\delta\rho(\vec{r})}\left[E[\rho(\vec{r})] - \mu\int \rho(\vec{r})d\vec{r}\right] = 0$$

$$\Rightarrow \quad \frac{\delta E[\rho(\vec{r})]}{\delta \rho(\vec{r})} = \mu \quad . \tag{5.97}$$

Kohn and Sham chose to separate $F[\rho(\vec{r})]$ into three parts, so that $E[\rho(\vec{r})]$ becomes

$$E\left[\rho(\vec{r})\right] = T_s\left[\rho(\vec{r})\right] + \frac{1}{2}\int\int \frac{\rho(\vec{r})\rho(\vec{r}')}{|\vec{r}-\vec{r}'|}d\vec{r}d\vec{r}' + E_{XC}[\rho(\vec{r})] + \int \rho(\vec{r})V_{\text{ext}}(\vec{r})d\vec{r} \quad , \tag{5.98}$$

where $T_s\left[\rho(\vec{r})\right]$ is defined as the kinetic energy of a *non-interacting* electron gas with density $\rho(\vec{r})$,

$$T_s\left[\rho(\vec{r})\right] = -\frac{1}{2}\sum_{i=1}^{N}\int \psi_i^*(\vec{r})\nabla^2\psi_i(\vec{r})d\vec{r} \quad . \tag{5.99}$$

Eq.(5.98) also acts as a definition for the *exchange-correlation energy functional*, $E_{XC}[\rho(\vec{r})]$. We can now rewrite Eq.(5.97) in terms of an effective potential, $V_{\text{eff}}(\vec{r})$, as follows

$$\frac{\delta T_s\left[\rho(\vec{r})\right]}{\delta\rho(\vec{r})} + V_{\text{eff}}(\vec{r}) = \mu \quad , \tag{5.100}$$

where

$$V_{\text{eff}}(\vec{r}) = V_{\text{ext}}(\vec{r}) + \int \frac{\rho(\vec{r}')}{|\vec{r}-\vec{r}'|}d\vec{r}' + V_{XC}(\vec{r}) \quad , \tag{5.101}$$

and

$$V_{XC}(\vec{r}) = \frac{\delta E_{XC}[\rho(\vec{r})]}{\delta\rho(\vec{r})} \quad . \tag{5.102}$$

Now, if one considers a system that really contained *non-interacting* electrons moving in an external potential equal to $V_{\text{eff}}(\vec{r})$, as defined in Eq.(5.101), then the same analysis would lead to exactly the same Eq.(5.100). Therefore, to find the ground state energy and density, E_0 and $\rho_0(\vec{r})$ all one has to do is solve the one-electron equations

$$\left(-\frac{1}{2}\nabla_i^2 + V_{\text{eff}}(\vec{r}) - \epsilon_i\right)\psi_i(\vec{r}) = 0 \quad . \tag{5.103}$$

As the density is constructed according to

$$\rho(\vec{r}) = \sum_{i=1}^{N}|\psi_i(\vec{r})|^2 \quad , \tag{5.104}$$

these equations (5.101-5.103) must be solved self-consistently with Eq.(5.104).

The above derivation assumes that the exchange-correlation functional is known. At present numerical exchange-correlation potentials have only been determined for a few simple model systems, and so most current density functional calculations use the local density approximation (LDA). The LDA approximates the XC functional by a simple function of the density at any position, $\vec{r}$. The value of this function is the XC energy per electron in a uniform homogeneous electron gas of density $n(\vec{r})$. The LDA expression for $E_{XC}[n(\vec{r})]$ is

$$E_{XC}[n(\vec{r})] \approx \int \epsilon_{XC}(n(\vec{r}))n(\vec{r})d\vec{r} \quad . \tag{5.105}$$

The LDA is only expected to be accurate for systems with slowly varying electronic charge densities, which is not the case in most real systems, but despite this it has been surprisingly successful. The success is currently understood to be due to two main features:

- The sum rule on the exchange-correlation hole is conserved, i.e. within the LDA, the exchange-correlation hole contains exactly an equal and opposite amount of charge to the electron it surrounds, and
- The exchange-correlation energy only depends on the spherical average of the exchange-correlation hole, i.e.
$$E_{XC}[\rho] = \frac{1}{2}\int\int \frac{\rho(\vec{r})\,\overline{\rho}_{XC}(\vec{r},|\vec{r}-\vec{r}'|)}{|\vec{r}-\vec{r}'|}d\vec{r}\,d\vec{r}' \quad , \tag{5.106}$$
where $\overline{\rho}_{XC}(\vec{r},|\vec{r}-\vec{r}'|)$ is the spherical average of $\rho_{XC}(\vec{r},\vec{r}')$ about $\vec{r}' = \vec{r}$ for each $\vec{r}$.

Therefore, the fact that in the LDA ρ_{XC} is constrained to be spherically symmetric about $\vec{r}' = \vec{r}$ is not a handicap.

The LDA is remarkably accurate, but often fails when the electrons are strongly correlated, as in systems containing *d* and *f* orbital electrons. In strongly correlated systems, the correlations may change the whole nature of the ground state and the local density approximation, derived from homogeneous electron gas results, is not successful. For example, the high T_c superconductor La_2CuO_4 is an anti-ferromagnetic insulator but the LDA finds it to be metallic. Also FeO, MnO and NiO all have Mott metal-insulator transitions but the LDA predicts that they are either semiconductors or metals. Other failings of the LDA are that it tends to overbind atoms in solids, that it finds stable negative ions to be unstable and that it predicts iron to be *fcc* paramagnetic, when it is actually *bcc* ferromagnetic.

5.9.6 *Shortcomings of the mean-field approach

The main problem with Hartree, Hartree-Fock and LDA methods is the approximations they introduce in the process of reducing the many-body problem to a one-electron problem. Hartree and Hartree-Fock calculations do not, in general, provide satisfactory results and are best used as a qualitative guide to the expected ground state properties. The configuration interaction method (popular in quantum chemistry), while in principle exact, is in practice only useful for small systems; for condensed matter systems it is not of practical value.

Density functional theory within the LDA provides the current staple method of performing electronic structure calculations and for many purposes gives good results. However, it fails for highly correlated systems and tends to underestimate band gaps and overestimates cohesive energies and hence is not ideal.

A straightforward, accurate approach to many-body systems that is now gaining popularity in the electronic structure theory community is the quantum Monte Carlo method, described in the next section.

5.9.7 *Quantum Monte Carlo methods

Applying QMC techniques to study the electronic structure of real systems is relatively new. The variational principle provides the starting point with the aim of finding an approximate solution to Schrödinger's equation. As we saw in Chapter 3, the expectation value of a Hamiltonian, $\hat{H}$, calculated using a trial wave function, Ψ_T, is never lower in value than the true ground state energy, ϵ_0, which is the expectation value of $\hat{H}$ calculated using the true ground state wave function, Ψ_0. The success of the method depends on making a physically plausible guess at the form of the ground state wave function, Ψ_T, of the Hamiltonian, $\hat{H}$. This guess will be referred to as the trial/guiding wave function. The "trial" part of the name refers to the use of the wave function as a guess of the true ground state wave function to be used as the input wave function in a variational quantum Monte Carlo (VMC) calculation. The "guiding" part refers to the use of the same wave function as an input wave function in the diffusion quantum Monte Carlo (DMC) algorithm as part of the mechanism to introduce importance sampling. The trial/guiding wave function depends on a number of variable parameters which can be adjusted to minimise the energy expectation value. If the guessed values of these parameters are good and the chosen functional form builds in enough variational freedom to adequately describe the physics of the system being studied, then very accurate estimates of the ground state energy can be obtained. Variational quantum Monte Carlo (VMC) calculations are direct applications of the above variational principle.

The variational quantum Monte Carlo (VMC) method is the simpler of the two quantum Monte Carlo methods. It is based on a combination of the variational principle and Monte Carlo evaluation of integrals using importance sampling based on the Metropolis algorithm.

Within the Born-Oppenheimer approximation, the Hamiltonian for a many-body system can be written as

$$\hat{H} = \sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 - \sum_i \sum_\alpha \frac{Z_\alpha}{|\vec{r}_i - \vec{d}_\alpha|} + \frac{1}{2}\sum_i \sum_{j\neq i} \frac{1}{|\vec{r}_i - \vec{r}_j|} + \frac{1}{2}\sum_\alpha \sum_{\beta\neq\alpha} \frac{Z_\alpha Z_\beta}{|\vec{d}_\alpha - \vec{d}_\beta|} \quad . \tag{5.107}$$

The VMC method relies on one being able to construct a trial wave function, Ψ_T, that is a reasonably good approximation to the true ground state wave function, Ψ_0. The energy associated with the trial wave function is given by

$$\begin{aligned} E_T &= \frac{\langle \Psi_T | \hat{H} | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} \\ &= \int \frac{\Psi_T^*(\vec{R}) \hat{H} \Psi_T(\vec{R}) d\vec{R}}{\Psi_T^*(\vec{R}) \Psi_T(\vec{R}) d\vec{R}} \quad . \end{aligned} \tag{5.108}$$

where $\vec{R}$ denotes the $3N$-dimensional vector of electronic positions. The variational principle ensures that E_T is a rigorous upper bound to the true ground state energy, E_0.

The VMC method is a Monte Carlo method for evaluating the multi-dimensional integral in Eq.(5.108). This is achieved by rewriting Eq.(5.108) in the following form:

$$E_T = \int \frac{\left| \Psi_T(\vec{R}) \right|^2 \frac{\hat{H}\Psi_T(\vec{R})}{\Psi_T(\vec{R})} d\vec{R}}{\int \left| \Psi_T(\vec{R}) \right|^2} \quad . \tag{5.109}$$

The Metropolis algorithm is used to sample a series of points, $\vec{R}$, in configuration space. At each of these points the “local energy”, $\frac{\hat{H}\Psi_T(\vec{R})}{\Psi_T(\vec{R})}$, is evaluated. After a sufficient number of evaluations of the local energy have been made, the average is taken:

$$E_{VMC} = \frac{1}{N} \sum_{i=1}^{N} \frac{\hat{H}\Psi_T(\vec{R}_i)}{\Psi_T(\vec{R}_i)} \underset{N \to \infty}{\longrightarrow} E_T \; . \tag{5.110}$$

(The Metropolis algorithm is a method of ensuring that, in the limit of large N, the $\vec{R}_i$ are sampled from $|\Psi_T(\vec{R})|^2$.)

Chapter 6

Density Operators

6.1 Introduction

A quantum-mechanical wave function (or state vector), when it exists, conveys the maximum amount of information permitted by quantum mechanics concerning the properties of a physical system in the state described by the wave function. Situations in which we have accurate wave functions for a physical system are actually quite rare. More often, the complexity of the system owing to its many degrees of freedom precludes the possibility of constructing a wave function. It is then necessary to resort to statistical methods. When the state of an incompletely prepared system is only partially known, we resort to assigning a probability to all possible state vectors that the system could be in. The synthesis of this statistical nature with the probabilities arising from the quantum mechanics of state vectors can be made using a mathematical entity, the **density operator**. The density operator formalism was introduced independently by Landau and von Neumann, and enables us to reformulate the laws of quantum mechanics more generally than with the formalism using state vectors or wave functions alone. All predictions – of a statistical nature – that one can make at a given time about a physical system can be found once we know its density operator. Furthermore, the density operator formalism enables us to extend quantum mechanics to the description of statistical mixtures representing systems which are not well known or to describe the state of a part of a quantum system (i.e. a subsystem). Here we will examine the properties of this operator and its application to quantum statistical mechanics.

6.2 Pure and mixed states

In classical mechanics, the dynamical state of a system is completely determined once the values of the positions and momenta of all the particles are known. The state of the system at any subsequent time can be predicted with certainty (since all one needs to do is integrate the equations of motion). In quantum mechanics however this is not the case. A precise simultaneous measurement of two physical variables is only possible if the two operator corresponding to the two variables *commute*. The largest set of mutually commuting independent observables, $\{\hat{A}, \hat{B}, \ldots\}$ that can be found will give the most complete characterisation possible. (This is just the **complete set of commuting observables (CSCO)** discussed in Chapter 1.) The measurement of another variable whose operator is not contained in the above set of operators necessarily introduces *uncertainty* into at least one of those already measured. This means that it is not possible to give a more complete specification of the system. In general, the maximum information which can be obtained on a system consists of the eigenvalues of the CSCO. The system is then completely specified by assigning the state vector $|a, b, \ldots\rangle$ in a Hilbert space H to it. If the measurement of the observables $\{\hat{A}, \hat{B}, \ldots\}$ on the state $|a, b, \ldots\rangle$ is immediately repeated, we get the same values $a, b, \ldots$ again. The existence of such a set of experiments (for which the results can be predicted with certainty) gives a necessary and sufficient characterisation for the state of "maximum knowledge". The states of maximum knowledge are called **pure states**. Pure states represent the ultimate limit of precise observation as permitted by the uncertainty principle and are the quantum mechanical analog of classical states where all positions and momenta of all particles are known.

In practice, the state of a system is not pure and cannot be represented by a single state vector. However, it can be described by stating that the system has certain probabilities $p_1, p_2, \ldots$ of being in the pure states $|\Psi_1\rangle, |\Psi_2\rangle, \ldots$, respectively. Therefore in the case of incomplete knowledge about the state of the system, it is necessary to use a *statistical description* in the same sense as classical statistical mechanics. Systems which cannot be characterised by a single-state vector are called **mixed states**.

Consider an ensemble of particles in the pure state $|\Psi\rangle$. If this state is not one of the eigenstates of the observable $\hat{A}$ then measurements of the corresponding physical quantity will produce a variety of results, each of which is an eigenvalue of $\hat{A}$. If similar measurements are made on a very large number of particles, all of which were in the same state $|\Psi\rangle$, then, in general, all the possible eigenvalues of $|\Psi\rangle$ would be obtained. The *average* of these values is given by the expectation value $\langle\hat{A}\rangle$ of the observable corresponding to $\hat{A}$ which is defined by the matrix element

$$\langle\hat{A}\rangle = \langle\Psi|\hat{A}|\Psi\rangle \tag{6.1}$$

where we have assumed that $|\Psi\rangle$ is normalised.

In order to obtain $\langle\hat{A}\rangle$ for a mixture of states, $|\Psi_1\rangle, |\Psi_2\rangle, \ldots$, the expectation values $\langle\Psi_i|\hat{A}|\Psi_i\rangle$ of each of the pure state components must be calculated and then averaged by summing over all pure states multiplied by its corresponding statistical weight p_i:

$$\langle\hat{A}\rangle = \sum_i p_i \,\langle\Psi_i|\hat{A}|\Psi_i\rangle \tag{6.2}$$

where we have again assumed that the $|\Psi_n\rangle$ are normalised. Note that statistics enter into Eq. (6.2) in two ways: First of all in the quantum mechanical expectation value $\langle\Psi_i|\hat{A}|\Psi_i\rangle$ and secondly in the ensemble average over these values with the weights p_i. While the first type of averaging is connected with the perturbation of the system during the measurement (and is therefore inherent in the nature of quantisation), the second averaging is introduced because of the lack of information as to which of the several pure states the system may be in. This latter averaging closely resembles that of classical statistical mechanics and it can be conveniently performed by using **density operator** techniques.

6.3 Properties of the Density Operator

The density operator is defined by

$$\hat{\rho} = \sum_i p_i |\Psi_i\rangle\langle\Psi_i| \tag{6.3}$$

where p_i is the probability of the system being in the normalised state $|\Psi_i\rangle$ and the sum is over all states that are accessible to the system. The probabilities p_i satisfy

$$0 \leq p_i \leq 1, \quad \sum_i p_i = 1, \quad \sum_i p_i^2 \leq 1 \tag{6.4}$$

For a *pure state* there is just one p_i (which is equal to unity) and all the rest are zero. In that case

$$\hat{\rho} = |\Psi\rangle\langle\Psi| \quad \text{(pure state)} \tag{6.5}$$

Let $\{|\psi_i\rangle\}$ be a complete orthonormal set which serves as a basis for the expansion of $|\Psi_i\rangle$ (and from which we can construct the matrix representation of state vectors and operators). We have

$$|\Psi_i\rangle = \sum_n c_{ni}\,|\psi_n\rangle \tag{6.6}$$

and from the orthonormality of the $\{|\psi_i\rangle\}$,

$$c_{ni} = \langle\psi_n|\Psi_i\rangle \tag{6.7}$$

We now construct the **density matrix** which consists of the matrix elements of the density operator in the $\{|\psi_i\rangle\}$ basis:

$$\begin{aligned}\langle\psi_n|\hat{\rho}|\psi_m\rangle &= \sum_i p_i\langle\psi_n|\Psi_i\rangle\langle\Psi_i|\psi_m\rangle \\ &= \sum_i p_i\, c_{ni}\, c^*_{mi} \qquad (6.8)\end{aligned}$$

which characterises $\hat{\rho}$ as a Hermitian operator since

$$\langle\psi_n|\hat{\rho}|\psi_m\rangle = \langle\psi_m|\hat{\rho}|\psi_n\rangle^* \qquad (6.9)$$

(given that the p_i are real), i.e we have

$$\hat{\rho} = \hat{\rho}^\dagger \qquad (6.10)$$

From Eq. (6.8), the probability of finding the system in the state $|\psi_n\rangle$ is given by the diagonal element

$$\langle\psi_n|\hat{\rho}|\psi_n\rangle = \sum_i p_i\,|c_{ni}|^2 \qquad (6.11)$$

which gives a physical interpretation of the diagonal elements of the density operator. Because probabilities are positive numbers, we have

$$\langle\psi_n|\hat{\rho}|\psi_n\rangle \geq 0 \qquad (6.12)$$

The trace of $\hat{\rho}$ (i.e. the sum of the diagonal matrix elements) is

$$\begin{aligned}\mathrm{Tr}\,\hat{\rho} &= \sum_n \langle\psi_n|\hat{\rho}|\psi_n\rangle \\ &= \sum_i\sum_n p_i\,\langle\psi_n|\Psi_i\rangle\,\langle\Psi_i|\psi_n\rangle \\ &= \sum_i p_i\langle\Psi_i|\Psi_i\rangle \\ &= \sum_i p_i \\ &= 1 \qquad (6.13)\end{aligned}$$

(Since the trace of an operator is an invariant quantity, the above result is independent of the basis.) As $\hat{\rho}$ is Hermitian, the diagonal elements $\langle\psi_n|\hat{\rho}|\psi_n\rangle$ must be real and from Eq. (6.8) it follows that

$$0 \leq \langle\psi_n|\hat{\rho}|\psi_n\rangle \leq 1 \qquad (6.14)$$

Note that for a pure state, $\langle\psi_n|\hat{\rho}|\psi_n\rangle = |c_n|^2$, which is the probability of finding the system in the state ψ_n.

Consider the matrix elements of $\hat{\rho}^2$:

$$\begin{aligned}\langle\psi_n|\hat{\rho^2}|\psi_m\rangle &= \sum_k \langle\psi_n|\hat{\rho}|\psi_k\rangle\,\langle\psi_k|\hat{\rho}|\psi_m\rangle \\ &= \sum_i\sum_j\sum_k p_i\,p_j\,\langle\psi_n|\Psi_i\rangle\langle\Psi_i|\psi_k\rangle\,\langle\psi_k|\Psi_j\rangle\langle\Psi_j|\psi_m\rangle \qquad (6.15)\end{aligned}$$

where we have used Eq. (6.3).

Problem 1: Using (6.15), show that

$$\mathrm{Tr}\,\hat{\rho}^2 \le 1 \qquad (6.16)$$

For a pure state, there is only one p_i and it is equal to unity. Therefore

$$\mathrm{Tr}\,\hat{\rho}^2 = 1 \quad \text{(pure state)} \qquad (6.17)$$

and

$$\begin{aligned}\hat{\rho}^2 &= |\Psi\rangle\langle\Psi|\Psi\rangle\langle\Psi| \\ &= |\Psi\rangle\langle\Psi| \\ &= \hat{\rho} \quad \text{(pure state)} \qquad (6.18)\end{aligned}$$

i.e. $\hat{\rho}$ is *idempotent* for a pure state. Thus whether a state is pure or not can be established by testing whether (6.17) or (6.18) is satisfied or not.

We now derive the expectation value of an operator $\hat{A}$ for pure as well as mixed states. Let

$$\langle\hat{A}\rangle_i = \langle\Psi_i|\hat{A}|\Psi_i\rangle \qquad (6.19)$$

and

$$\langle\hat{A}\rangle = \sum_i p_i\,\langle\hat{A}\rangle_i \qquad (6.20)$$

The distinction between $\langle\hat{A}\rangle_i$ and $\langle\hat{A}\rangle$ is that the former is a quantum-mechanical average or the expectation value of an operator $\hat{A}$ when the system is definitely in the state $|\Psi_i\rangle$. On the other hand, $\langle\hat{A}\rangle$ is a statistical or *ensemble average* which from (6.20), is seen to be the weighted average of $\langle\hat{A}\rangle_i$ taken over all states that the system may occupy. For pure states, we have

$$\langle\hat{A}\rangle = \langle\hat{A}\rangle_i \quad \text{(pure state)} \qquad (6.21)$$

Now consider the operator $\hat{\rho}\hat{A}$. From (6.3) we have

$$\hat{\rho}\hat{A} = \sum_i p_i \, |\Psi_i\rangle\langle\Psi_i| \, \hat{A} \tag{6.22}$$

In the $\{|\psi_i\rangle\}$ basis,

$$\langle\psi_n|\hat{\rho}\hat{A}|\psi_m\rangle = \sum_i p_i \, \langle\psi_n|\Psi_i\rangle\langle\Psi_i|\hat{A}|\psi_m\rangle \tag{6.23}$$

Taking the trace of $\hat{\rho}\hat{A}$,

$$\begin{aligned}
\mathrm{Tr}\,\hat{\rho}\hat{A} &= \sum_n \langle\psi_n|\hat{\rho}\hat{A}|\psi_n\rangle \\
&= \sum_i \sum_n p_i \, \langle\psi_n|\Psi_i\rangle\langle\Psi_i|\hat{A}|\psi_n\rangle \\
&= \sum_i p_i \, \langle\Psi_i|\hat{A}|\Psi_i\rangle \\
&= \langle\hat{A}\rangle
\end{aligned} \tag{6.24}$$

Thus the average value of an operator for a system in either a pure or mixed state, is known as soon as the density operator is known. **Therefore the density operator contains all physically significant information on the system**.

To summarise, the density operator $\hat{\rho}$ has the following properties:

- $\hat{\rho}$ is Hermitean: $\hat{\rho} = \hat{\rho}^\dagger$. This follows from the fact that the p_i are real. This property means that the *expectation value of any observable is real.*
- $\hat{\rho}$ has unit trace: $\mathrm{Tr}\,\hat{\rho} = 1$.
- $\hat{\rho}$ is non-negative: $\langle\Phi|\hat{\rho}|\Phi\rangle \geq 0 \;\; \forall \;\; |\Phi\rangle \in H$
- The expectation value of an operator $\hat{A}$ is given by $\langle\hat{A}\rangle = \mathrm{Tr}\,\hat{\rho}\hat{A}$.

6.3.1 Density operator for spin states

Suppose the spin state of an electron is given by

$$|\Psi\rangle = |\uparrow\rangle \tag{6.25}$$

so that the density operator is

$$\hat{\rho} = |\uparrow\rangle\langle\uparrow| \tag{6.26}$$

In the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$ (i.e. the eigenstates of $\hat{S}_z$, the z-component of the spin angular momentum of the electron), the density matrix is

$$\hat{\rho} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{6.27}$$

Problem 2: Verify (6.27) and hence show that the expectation values of the operators $\hat{S}_x, \hat{S}_y, \hat{S}_z$ are 0,0 and $\frac{1}{2}\hbar$ respectively.

More generally, if the electron is in a state described by

$$|\Psi\rangle = a_1|\uparrow\rangle + a_2|\downarrow\rangle \tag{6.28}$$

with

$$|a_1|^2 + |a_2|^2 = 1$$

the density operator is

$$\hat{\rho} = \begin{pmatrix} |a_1|^2 & a_1 a_2^* \\ a_2 a_1^* & |a_2|^2 \end{pmatrix} \tag{6.29}$$

which indicates that the diagonal elements $|a_1|^2$ and $|a_2|^2$ are just the probabilities that the electron is the state $|\uparrow\rangle$ and $|\downarrow\rangle$ respectively.

Another useful form for the density matrix for spin-$\frac{1}{2}$ particles is obtained by writing

$$\hat{\rho} = c_0\, I + c_1\, \hat{S}_x + c_2\, \hat{S}_y + c_3\, \hat{S}_z \tag{6.30}$$

where I is the unit 2×2 matrix and the c_i's are real numbers. The density matrix becomes

$$\hat{\rho} = \begin{pmatrix} c_0 + \frac{1}{2}c_3 & \frac{1}{2}(c_1 - ic_2) \\ \frac{1}{2}(c_1 + ic_2) & c_0 - \frac{1}{2}c_3 \end{pmatrix} \tag{6.31}$$

(where we have set $\hbar = 1$).

Problem 3: Verify (6.31) using the definition of the spin operators in terms of the Pauli matrices. Show that $c_0 = \frac{1}{2}$ and the expectation values of $\hat{S}_x, \hat{S}_y, \hat{S}_z$ are given by $\frac{1}{2}c_1, \frac{1}{2}c_2, \frac{1}{2}c_3$ respectively. Hence show that the density operator can be written compactly as

$$\hat{\rho} = \frac{1}{2}\Big[I + \langle\hat{\sigma}\rangle \cdot \sigma\Big]$$

where $\sigma = (\sigma_x, \sigma_y, \sigma_z)$ is the vector whose components are the Pauli matrices.

Problem 4: By analogy with the polarisation of the spin-$\frac{1}{2}$ case discussed in this section, the polarisation of a light quantum can be described by a two-component wave function $\begin{pmatrix} a \\ b \end{pmatrix}$, where

$|a|^2$ and $|b|^2$ are the probabilities that the photon is polarised in one or the other of two mutually perpendicular directions (or that the photon is right- or left-hand circularly polarised). If we want to determine the polarisation of a photon, we could, for instance, use a *filter*, which we shall call a detector (although strictly speaking it is not a detector but a device to prepare for a measurement). Such a filter could correspond to a pure state, described by a wave function

$$\Psi^{det} = c_1^{det}\,\Psi_1 + c_2^{det}\,\Psi_2$$

where

$$\Psi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \Psi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{6.32}$$

are the wave functions corresponding to the two polarisation states. This pure state corresponds to a 2×2 detector density matrix $\hat{\rho}^{det}$ given by its matrix elements

$$\rho_{ij}^{det} = c_i^{det} \cdot (c_j^{det})^*$$

Find an expression for the probability of a response of a detector described by $\hat{\rho}^{det}$ to a photon in a state described by a density matrix $\hat{\rho}$.

6.3.2 Density operator in the position representation

The density operator in the position representation is defined by

$$\begin{aligned} \rho(x', x) &= \langle x'|\hat{\rho}|x\rangle \\ &= \sum_i p_i\,\Psi_i(x')\,\Psi_i^*(x) \end{aligned} \tag{6.33}$$

which, for a pure state becomes

$$\rho(x', x) = \Psi(x')\,\Psi^*(x) \quad \text{(pure state)} \tag{6.34}$$

The expectation value for an operator $\hat{A}$ is then given by

$$\begin{aligned} \langle \hat{A} \rangle &= \mathrm{Tr}\,\hat{\rho}\,\hat{A} \\ &= \int dx\,\langle x|\hat{\rho}\hat{A}|x\rangle \\ &= \int dx\,\langle x|\hat{\rho}\Big(\int dx'|x'\rangle\langle x'|\Big)\hat{A}|x\rangle \\ &= \int\int dx'\,dx\,\langle x|\hat{\rho}|x'\rangle\,\langle x'|\hat{A}|x\rangle \\ &= \int\int dx'\,dx\,\rho(x, x')\,A(x', x) \end{aligned} \tag{6.35}$$

Problem 5: Show that

(a) When $\hat{A} = \hat{x}$, i.e. the position operator, then

$$\langle \hat{x} \rangle = \int dx\, x\, \rho(x, x)$$

(b) When $\hat{A} = \hat{p}$, i.e. the momentum operator, then

$$\langle \hat{p} \rangle = -\frac{\hbar}{i} \int dx \left[\frac{\partial}{\partial x} \rho(x, x') \right]_{x'=x}$$

Problem 6: Often one is dealing with a system which is part of a larger system. Let x and q denote, respectively, the coordinates of the smaller system and the coordinates of the remainder of the larger system. The larger system will be described by a normalised wave function $\Psi(x, q)$ which cannot necessarily be written as a product of functions depending on x and q only. Let $\hat{A}$ be an operator acting *only on the x variables*, let $\hat{H}$ be the Hamiltonian describing the smaller system, and let the density operator $\hat{\rho}$ be defined in the position representation by the equation

$$\langle x|\hat{\rho}|x' \rangle = \int \Psi^*(q, x')\, \Psi(q, x)\, dq \tag{6.36}$$

where the integration is over all the degrees of freedom of the remainder of the larger system.

(a) Express the expectation value of $\hat{A}$ in terms of $\hat{\rho}$ for the case where the larger system is described by the wave function $\Psi(q, x)$.

(b) What is the normalisation condition for $\hat{\rho}$?

(c) Find the equation of motion for $\hat{\rho}$.

Problem 7: If the wave function $\Psi(q, x)$ of the preceding problem can be written in the form

$$\Psi(q, x) = \Phi(q)\, \chi(x) \tag{6.37}$$

we are dealing with a *pure* state. Prove that the necessary and sufficient condition for the pure state is that $\hat{\rho}$ is idempotent, i.e. that

$$\hat{\rho}^2 = \hat{\rho} \tag{6.38}$$

6.4 Density operator in statistical mechanics

Let $\{\Psi_n\}$ be a complete set of orthonormal functions (eigenstates of the Hamiltonian $\hat{H}$ for the system), that satisfy

$$\hat{H}\Psi_n = E_n\Psi_n \tag{6.39}$$

and let the states be occupied according to the Boltzmann distribution

$$p_n = Ne^{-\beta E_n} \tag{6.40}$$

where p_n is the probability of finding the system in the eigenstate Ψ_n with energy E_n, $\beta = 1/kT$ with k the Boltzmann constant, T the absolute temperature and N a normalisation constant chosen to ensure that

$$\sum_n p_n = 1$$

Condition (6.40) defines **thermal equilibrium** and the corresponding density operator $\hat{\rho}(\beta)$ is (from (6.3)):

$$\hat{\rho}(\beta) = N\sum_n e^{-\beta E_n}|\Psi_n\rangle\langle\Psi_n| \tag{6.41}$$

Since

$$e^{-\beta\hat{H}}|\Psi_n\rangle = e^{-\beta E_n}|\Psi_n\rangle \tag{6.42}$$

this enables us to write

$$\begin{aligned}
\hat{\rho}(\beta) &= N\sum_n e^{-\beta E_n}|\Psi_n\rangle\langle\Psi_n| \\
&= N\sum_n e^{-\beta\hat{H}}|\Psi_n\rangle\langle\Psi_n| \\
&= Ne^{-\beta\hat{H}}\sum_n |\Psi_n\rangle\langle\Psi_n| \\
&= Ne^{-\beta\hat{H}} \qquad (6.43)
\end{aligned}$$

To determine N, we note that

$$\begin{aligned}
\mathrm{Tr}\,\hat{\rho}(\beta) &= N\,\mathrm{Tr}\,e^{-\beta\hat{H}} \\
&= 1 \qquad (6.44)
\end{aligned}$$

which implies

$$N = \frac{1}{\mathrm{Tr}\,e^{-\beta\hat{H}}} \tag{6.45}$$

Hence the density operator under thermal equilibrium is

$$\begin{aligned}\hat{\rho}(\beta) &= \frac{e^{-\beta\hat{H}}}{\mathrm{Tr}\, e^{-\beta\hat{H}}} \\ &= \frac{1}{Z}\, e^{-\beta\hat{H}}\end{aligned} \tag{6.46}$$

where

$$Z = \mathrm{Tr}\, e^{-\beta\hat{H}} \tag{6.47}$$

is known as the **canonical partition function**. (Note that the partition function is a function of absolute temperature, T, the volume V, and the number of particles that make up the system, N.) We see that from the knowledge of the density operator in any representation, one can determine the partition function and therefore all thermodynamic properties of the system. For instance, the average of an observable $\hat{A}$ is given by

$$\begin{aligned}\langle\hat{A}\rangle &= \mathrm{Tr}\,\hat{\rho}\,\hat{A} \\ &= \frac{\mathrm{Tr}\left[e^{-\beta\hat{H}}\,\hat{A}\right]}{\mathrm{Tr}\left[e^{-\beta\hat{H}}\right]}\end{aligned} \tag{6.48}$$

The mean energy of the system (i.e. the internal energy) is given by U where

$$\begin{aligned}U &= \langle\hat{H}\rangle \\ &= \frac{\mathrm{Tr}\left[e^{-\beta\hat{H}}\,\hat{H}\right]}{\mathrm{Tr}\left[e^{-\beta\hat{H}}\right]} \\ &= -\frac{\partial}{\partial\beta}\,\ln\left(\mathrm{Tr}\, e^{-\beta\hat{H}}\right) \\ &= -\frac{\partial}{\partial\beta}\,\ln\, Z(T,V,N)\end{aligned} \tag{6.49}$$

From the partition function, we obtain all thermodynamic observables:

$$\begin{aligned}S &= -k\,\mathrm{Tr}\,(\hat{\rho}\,\ln\hat{\rho}) \qquad \text{(entropy)} \\ &= k\,\beta\,\langle\hat{H}\rangle + k\,\ln\, Z(T,V,N) \\ F &= U - TS \qquad \text{(Helmholtz free energy)}\end{aligned}$$

$$\begin{aligned} &= -kT \ln Z(T,V,N) \\ &= -kT \ln \left[\mathrm{Tr}\left\{ e^{-\beta \hat{H}} \right\} \right] \end{aligned} \tag{6.50}$$

We now calculate the density operator (in various representations) for some concrete cases.

6.4.1 Density operator for a free particle in the momentum representation

We determine the density operator in the momentum representation for a free particle in a box of volume L^3 with periodic boundary conditions. The Hamiltonian is given by $\hat{H} = \hat{p}^2/2m$ and the energy eigenfunction are plane waves;

$$\hat{H}|\psi_{\mathbf{k}}\rangle = E|\psi_{\mathbf{k}}\rangle \tag{6.51}$$

with

$$E = \frac{\hbar^2 \mathbf{k}^2}{2m} \tag{6.52}$$

and $|\psi_{\mathbf{k}}\rangle$ defined by

$$\begin{aligned} \psi_{\mathbf{k}}(\mathbf{r}) &= \frac{1}{\sqrt{V}}\, e^{i\mathbf{k}\cdot\mathbf{r}} \\ \mathbf{k} &= \frac{2\pi}{L}(n_x, n_y, n_z) \\ n_i &= 0, \pm 1, \pm 2, \ldots \end{aligned} \tag{6.53}$$

Note that the energy eigenvalues are discrete but their mutual separation for macroscopic volumes is so small that one may treat them as essentially continuous. The advantage of the formulation using a box and periodic boundary conditions is that one has automatically introduced into the formalism a finite volume for the particles, which is not the case for free plane waves we have used so far (in scattering theory for example). The functions $\psi_{\mathbf{k}}(\mathbf{r})$ are orthonormalized,

$$\begin{aligned} \langle \psi_{\mathbf{k}'} | \psi_{\mathbf{k}} \rangle &= \delta_{\mathbf{k},\mathbf{k}'} \\ &= \delta_{n_{x'},n_x}\, \delta_{n_{y'},n_y}\, \delta_{n_z,n_z} \end{aligned} \tag{6.54}$$

and complete,

$$\sum_{\mathbf{k}} \psi^*_{\mathbf{k}}(\mathbf{r}')\psi_{\mathbf{k}}(\mathbf{r}) = \delta(\mathbf{r}' - \mathbf{r}) \tag{6.55}$$

The canonical partition function is

$$\begin{aligned} Z(T,V,1) &= \operatorname{Tr} e^{-\beta\hat{H}} \\ &= \sum_{\mathbf{k}} \langle \psi_{\mathbf{k}} | e^{-\beta\hat{H}} | \psi_{\mathbf{k}} \rangle \\ &= \sum_{\mathbf{k}} e^{-\frac{\beta\hbar^2}{2m}\mathbf{k}^2} \end{aligned} \tag{6.56}$$

Since the eigenvalues $\mathbf{k}$ are very close together in a large volume, we can replace the sum in (6.56) by an integral.

$$\begin{aligned} Z(T,V,1) &= \frac{V}{(2\pi)^3} \int d\mathbf{k}\, e^{-\frac{\beta\hbar^2}{2m}\mathbf{k}^2} \\ &= \frac{V}{(2\pi)^3} \left(\frac{2m\pi}{\beta\hbar^2} \right)^{3/2} \\ &= \frac{V}{\lambda^3} \end{aligned} \tag{6.57}$$

where λ is called the *thermal wavelength.* The matrix elements of the density operator thus becomes

$$\langle \psi_{\mathbf{k}'} | \hat{\rho} | \psi_{\mathbf{k}} \rangle = \frac{\lambda^3}{V}\, e^{-\frac{\beta\hbar^2}{2m}\mathbf{k}^2}\, \delta_{\mathbf{k},\mathbf{k}'} \tag{6.58}$$

which is a diagonal matrix.

6.4.2 Density operator for a free particle in the position representation

We look for the canonical density operator in the position representation for a free particle in a box of volume V and periodic boundary conditions. We have

$$\begin{aligned} \langle \mathbf{r}' | \hat{\rho} | \mathbf{r} \rangle &= \sum_{\mathbf{k},\mathbf{k}'} \langle \mathbf{r}' | \mathbf{k}' \rangle\, \langle \mathbf{k}' | \hat{\rho} | \mathbf{k} \rangle\, \langle \mathbf{k} | \mathbf{r} \rangle \\ &= \sum_{\mathbf{k},\mathbf{k}'} \psi_{\mathbf{k}'}(\mathbf{r}') \Big\{ \frac{\lambda^3}{V}\, e^{-\frac{\beta\hbar^2}{2m}\mathbf{k}^2}\, \delta_{\mathbf{k},\mathbf{k}'} \Big\} \psi^*_{\mathbf{k}}(\mathbf{r}) \\ &= \frac{\lambda^3}{V} \frac{1}{(2\pi)^3} \int d\mathbf{k}\, \exp\Big\{ -\frac{\beta\hbar^2}{2m}\mathbf{k}^2 + i\mathbf{k}\cdot(\mathbf{r}' - \mathbf{r}) \Big\} \end{aligned} \tag{6.59}$$

Problem 8: Show that Eq. (6.59) reduces to

$$\begin{aligned} \langle \mathbf{r}'|\hat{\rho}|\mathbf{r}\rangle &= \frac{\lambda^3}{V}\frac{1}{(2\pi)^3}\exp\Big\{-\frac{m}{2\beta\hbar^2}(\mathbf{r}'-\mathbf{r})^2\Big\}\Big(\frac{2m\pi}{\beta\hbar^2}\Big)^{3/2} \\ &= \frac{1}{V}\exp\Big\{-\frac{\pi}{\lambda^2}(\mathbf{r}'-\mathbf{r})^2\Big\} \qquad (6.60) \end{aligned}$$

Hence in the position representation, the density matrix is no longer a diagonal matrix, but a Gaussian function in $(\mathbf{r}'-\mathbf{r})$. The diagonal elements of the density matrix in the position representation can be interpreted as the density distribution in position space i.e.

$$\langle \mathbf{r}'|\hat{\rho}|\mathbf{r}\rangle = \rho(\mathbf{r}) = \frac{1}{V} \qquad (6.61)$$

The non-diagonal elements $\mathbf{r} \neq \mathbf{r}'$ can be interpreted as the transition probability of the particle to move from a position $\mathbf{r}$ to a new position $\mathbf{r}'$ (though these transitions are restricted to spatial regions having the size of the thermal wavelengths.) For large temperatures ($\lambda \to 0$) this is hardly observable, but for low temperatures λ may become very large, which implies that quantum effects play an especially large role at low temperatures.

6.4.3 *Density matrix for the harmonic oscillator

Here, we determine the density matrix for the one-dimensional quantum harmonic oscillator in the position representation. This result is of great importance in quantum statistical mechanics and the mathematical steps involved in deriving the final result carry over to other areas of theoretical physics. We shall use the expression for the energy eigenfunction in the position representation derived before:

$$\begin{aligned} \Psi_n(q) &= \Big(\frac{M\omega}{\pi\hbar}\Big)^{1/4}\frac{H_n(x)}{\sqrt{2^n n!}}\exp\Big\{-\frac{1}{2}x^2\Big\} \\ x &= \sqrt{\frac{M\omega}{\hbar}}\,q \qquad (6.62) \end{aligned}$$

and the energy eigenvalues are $E_n = \hbar\omega(n+\frac{1}{2})$, and the Hermite polynomials are defined by

$$\begin{aligned} H_n(x) &= (-1)^n e^{x^2}\Big(\frac{d}{dx}\Big)^n e^{-x^2} \\ &= \frac{e^{x^2}}{\sqrt{\pi}}\int_{-\infty}^{+\infty}(-2iu)^n\,\exp\{-u^2+2ixu\}\,du \qquad (6.63) \end{aligned}$$

The density operator in the energy representation is trivial:

$$\begin{aligned}\langle m|\hat{\rho}|n\rangle &= \rho_n \delta_{mn} \\ \rho_n &= \frac{1}{Z}\exp\Big\{-\beta\,\hbar\omega(n+\frac{1}{2})\Big\} \\ n &= 0,1,2,\ldots \end{aligned} \tag{6.64}$$

where

$$Z(T,V,1) = \left[2\sinh\left(\frac{1}{2}\beta\hbar\omega\right)\right]^{-1} \tag{6.65}$$

Problem 9†: Verify Eq. (6.64)–(6.65).

In the position representation, we have

$$\begin{aligned}\langle q'|\hat{\rho}|q\rangle &= \sum_{nn'} \langle q'|n'\rangle\langle n'|\hat{\rho}|n\rangle\langle q|n\rangle \\ &= \sum_{nn'} \Psi_{n'}(q')\rho_{nn'}\Psi_n^*(q) \\ &= \frac{1}{Z}\sum_n \exp\Big\{-\beta\,\hbar\omega(n+\frac{1}{2})\Big\}\,\Psi_n^*(q)\,\Psi_n(q') \\ &= \frac{1}{Z}\left(\frac{M\omega}{\pi\hbar}\right)^{1/2}\exp\Big\{-\frac{1}{2}(x^2+x'^2)\Big\} \\ &\times \sum_{n=0}^{\infty}\frac{1}{2^n n!}\exp\Big\{-\beta\,\hbar\omega(n+\frac{1}{2})\Big\}\,H_n(x)\,H_n(x') \end{aligned} \tag{6.66}$$

Problem 10†: Verify the steps leading to (6.66).

Hint: We have twice inserted the complete set of energy eigenfunctions.

Using the integral representation of the Hermite polynomials we get:

$$\begin{aligned}\langle q'|\hat{\rho}|q\rangle &= \frac{1}{Z\pi}\left(\frac{M\omega}{\pi\hbar}\right)^{1/2}\exp\Big\{+\frac{1}{2}(x^2+x'^2)\Big\} \\ &\times \int_{-\infty}^{+\infty}du\int_{-\infty}^{+\infty}dv\sum_{n=0}^{\infty}\frac{(-2uv)^n}{n!}\exp\Big\{-\beta\,\hbar\omega(n+\frac{1}{2})\Big\}\exp\{-u^2+2ixu\} \\ &\times \exp\{-v^2+2ix'v\} \end{aligned} \tag{6.67}$$

The summation over n can be carried out as follows:

$$\begin{aligned}
&\sum_{n=0}^{\infty} \frac{(-2uv)^n}{n!} \exp\Big\{ -\beta\,\hbar\omega(n+\frac{1}{2})\Big\} \\
&= \exp\Big\{ -\frac{1}{2}\beta\hbar\omega\Big\{ \sum_{n=0}^{\infty} \frac{1}{n!}\Big[-2uv\exp(-\beta\hbar\omega)\Big]^n \\
&= \exp\Big\{ -\frac{1}{2}\beta\hbar\omega\Big\{ \exp\Big\{ -2uv\, e^{-\beta\hbar\omega}\Big\} \qquad (6.68)
\end{aligned}$$

Then Eq. (6.67) becomes

$$\begin{aligned}
\langle q'|\hat{\rho}|q\rangle &= \frac{1}{Z\pi}\left(\frac{M\omega}{\pi\hbar}\right)^{1/2} \exp\Big\{ +\frac{1}{2}(x^2+x'^2-\beta\hbar\omega)\Big\} \\
&\times \int_{-\infty}^{+\infty} du \int_{-\infty}^{+\infty} dv\, \exp\Big\{ -u^2+2ixu-v^2+2ix'v-2uv\, e^{-\beta\hbar\omega}\Big\} \qquad (6.69)
\end{aligned}$$

The argument in the exponent is a general quadratic form, which can be rewritten in the form

$$-u^2+2ixu-v^2+2ix'v-2uv\, e^{-\beta\hbar\omega} = -\frac{1}{2}\mathbf{w}^T\cdot\mathbf{A}\cdot\mathbf{w}+i\mathbf{b}\cdot\mathbf{w} \qquad (6.70)$$

where

$$\begin{aligned}
\mathbf{A} &= 2\begin{pmatrix} 1 & e^{-\beta\hbar\omega} \\ e^{-\beta\hbar\omega} & 1 \end{pmatrix} \\
\mathbf{b} &= 2\begin{pmatrix} x \\ x' \end{pmatrix} \\
\mathbf{w} &= \begin{pmatrix} u \\ v \end{pmatrix} \qquad (6.71)
\end{aligned}$$

We now use the general formula

$$\int d^n\mathbf{w}\, \exp\Big\{ -\frac{1}{2}\mathbf{w}^T\cdot\mathbf{A}\cdot\mathbf{w}+i\mathbf{b}\cdot\mathbf{w}\Big\} = \frac{(2\pi)^{n/2}}{[det\mathbf{A}]^{\frac{1}{2}}} \exp\Big\{ -\frac{1}{2}\mathbf{b}^T\cdot\mathbf{A}^{-1}\cdot\mathbf{b}\Big\} \qquad (6.72)$$

which holds if **A** is an invertible symmetric matrix.

Problem 11†: Verify (6.72).

Using (6.72) we get

$$\langle q'|\hat{\rho}|q\rangle = \frac{1}{Z}\left(\frac{M\omega}{\pi\hbar}\right)^{1/2} \frac{e^{-\frac{1}{2}\beta\hbar\omega}}{[1-e^{-2\beta\hbar\omega}]^{\frac{1}{2}}}$$

$$
\begin{aligned}
&\times \quad \exp\left\{\frac{1}{2}(x^2+x'^2) - [1-e^{-2\beta\hbar\omega}]^{-1}\left(x^2+x'^2-2xx'e^{-\beta\hbar\omega}\right)\right\} \\
&= \quad \frac{1}{Z}\left[\frac{M\omega}{2\pi\hbar\sinh(\beta\hbar\omega)}\right]^{\frac{1}{2}} \exp\left\{-\frac{1}{2}(x^2+x'^2)\coth(\beta\hbar\omega)+\frac{xx'}{\sinh(\beta\hbar\omega)}\right\} \qquad (6.73)
\end{aligned}
$$

Using the identity

$$\tanh\left(\frac{1}{2}\beta\hbar\omega\right) = \frac{\cosh(\beta\hbar\omega)-1}{\sinh(\beta\hbar\omega)} = \frac{\sinh(\beta\hbar\omega)}{1+\cosh(\beta\hbar\omega)} \qquad (6.74)$$

one finally gets

$$
\begin{aligned}
\langle q'|\hat{\rho}|q\rangle \quad &= \quad \frac{1}{Z}\left[\frac{M\omega}{2\pi\hbar\sinh(\beta\hbar\omega)}\right]^{\frac{1}{2}} \\
&\times \quad \exp\left\{-\frac{M\omega}{4\hbar}\left[(q+q')^2\tanh\left(\frac{1}{2}\beta\hbar\omega\right)+(q-q')^2\coth\left(\frac{1}{2}\beta\hbar\omega\right)\right]\right\} \qquad (6.75)
\end{aligned}
$$

The diagonal elements of the density matrix in the position representation yield directly the average density distribution of a quantum mechanical oscillator (at temperature T):

$$\rho(q) = \left[\frac{M\omega}{\pi\hbar}\tanh\left(\frac{1}{2}\beta\hbar\omega\right)\right]^{\frac{1}{2}} \exp\left\{-\frac{M\omega}{\hbar}\tanh\left(\frac{1}{2}\beta\hbar\omega\right)q^2\right\} \qquad (6.76)$$

which is a Gaussian distribution with width

$$\sigma_q = \left[\frac{\hbar}{2M\omega\tanh\left(\frac{1}{2}\beta\hbar\omega\right)}\right]^{\frac{1}{2}} \qquad (6.77)$$

Problem 12†: Show that in the limit of high temperatures, $\beta\hbar\omega \ll 1$,

$$\rho(q) \approx \left(\frac{m\omega^2}{2\pi kT}\right)^{\frac{1}{2}} \exp\left\{-\frac{M\omega^2q^2}{2kT}\right\} \qquad (6.78)$$

and at low temperature $\beta\hbar\omega \gg 1$,

$$\rho(q) \approx \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{2}} \exp\left\{-\frac{M\omega q^2}{\hbar}\right\} \qquad (6.79)$$

Therefore the density matrix thus contains, for high temperatures, the classical limit, and for very low temperatures, the quantum mechanical ground state density.
