Preface

In this course, we cover some more advanced and mathematical 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 such as scattering theory. A course outline is provided below.

Course Outline

- **Introduction/Revision:** Mathematical foundations of non-relativistic quantum mechanics. Vector spaces. Operator methods for discrete and continuous eigenspectra. Generalized form of the uncertainty principle. Dirac delta function and delta-function potential.
- **Quantum Dynamics:** Time development operator. Schrödinger, Heisenberg and interaction pictures. Canonical quantisation and constants of motion. Coordinate and momentum representations. Free particle and simple harmonic oscillator propagators. Introduction to path integral formulation.
- **Approximate Methods:** Variational methods and their application to problems of interest. The JWKB method and connection formulae, with applications to bound states and barrier penetration. The anharmonic oscillator. Asymptotic expansions.
- **Scattering Theory:** Scattering amplitudes and differential cross section. Partial wave analysis and the optical theorem. Green functions, weak scattering and the Born approximation. Relation between Born approximation and partial wave expansions. Beyond the Born approximation.
- **Density Matrices:** Pure and mixed states. The density operator and its properties. Position and momentum representation of the density operator. Spin density matrix and polarisation. Density matrix for the harmonic oscillator. Applications in statistical mechanics.
- **Lie Groups:** Rotation group, SO(3) and SU(2). SO(4) and the hydrogen atom.

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

While we have tried to make these notes as self-contained as possible, you are encouraged to deepen your understanding by reading the relevant sections of the recommended texts listed below. Merzbacher gives a very careful treatment of many of the main topics. Liboff is at about the right level and it is particularly strong on applications. Sakurai is more demanding mathematically although he makes a lot of effort to explain the concepts clearly. Jones covers most of the group theory at the right level. Bender and Orszag is a mine of information on techniques that are extremely useful to a practising theorist and hard to find elsewhere.

At about the level of the course:
E Merzbacher, *Quantum Mechanics*, 3^{rd} edn., Wiley, 1998
RL Liboff, *Introductory Quantum Mechanics*, 3^{rd} edn., Addison-Wesley, 1998
HF Jones, *Groups, Representations and Physics*, 2^{nd} edn., IoP, 1998

At a more advanced level:
JJ Sakurai, *Modern Quantum Mechanics*, 2^{nd} edn., Addison-Wesley, 1994
C Bender & SA Orszag, *Adv. Mathematical Methods for Scientists and Engineers*, Springer, 1999

Contents

Chapter 1

Introduction/Revision

1.1 Postulates of quantum mechanics

The purpose of this chapter is twofold: first to review the mathematical formalism of elementary non-relativistic quantum mechanics, especially the terminology, and second to present the basic tools of operator methods, commutation relations, etc. 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

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

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

$$\int_{all\ space} \Phi^* \Phi \, d\mathbf{r} = 1$$

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.

1.2 Vector spaces

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.

1.2.1 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

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

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

   $$\alpha(|a\rangle + |b\rangle) = \alpha|a\rangle + \alpha|b\rangle \tag{1.4}$$

   $$(\alpha + \beta)|a\rangle = \alpha|a\rangle + \beta|a\rangle \tag{1.5}$$

   $$(\alpha\beta)|a\rangle = \alpha(\beta|a\rangle) \tag{1.6}$$

   $$1|a\rangle = |a\rangle \tag{1.7}$$

2. A *scalar product* (or *inner 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

   $$(|a\rangle, \lambda|b\rangle) = \lambda(|a\rangle, |b\rangle) \tag{1.8}$$

   $$(|a\rangle, |b\rangle + |c\rangle) = (|a\rangle, |b\rangle) + (|a\rangle, |c\rangle) \tag{1.9}$$

   $$(|a\rangle, |b\rangle) = (|b\rangle, |a\rangle)^* \tag{1.10}$$

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

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

It follows from (1.11) that the scalar product of any vector with itself, $\langle a|a\rangle$, is a real number. To qualify as a true scalar product, the following requirements must also be satisfied:

$$\begin{aligned} \langle a|a\rangle &\geq 0 && (1.16)\\ \langle a|a\rangle &= 0 \text{ iff } |a\rangle = |\text{null}\rangle\ . && (1.17) \end{aligned}$$

The *norm* of a vector can then be defined by

$$\|a\| = \sqrt{\langle a|a\rangle} \tag{1.18}$$

which corresponds to the "length" of a vector. The norm of a vector is a real number ≥ 0, and only the vector $|\text{null}\rangle$ has norm zero.

1.2.2 The Schwartz inequality

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

$$\|a\|\ \ \|b\| \geq |\langle a|b\rangle| \tag{1.19}$$

with the equality only being valid for the case

$$|a\rangle = \lambda|b\rangle \tag{1.20}$$

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

Proof: Define a $|c\rangle$ such that

$$|c\rangle = |a\rangle + \lambda|b\rangle \tag{1.21}$$

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

$$\langle c|c\rangle = \langle a|a\rangle + \lambda\langle a|b\rangle + \lambda^*\langle b|a\rangle + \lambda\lambda^*\langle b|b\rangle \tag{1.22}$$

$$\geq 0 \tag{1.23}$$

Choose for λ the value

$$\lambda = -\frac{\langle b|a\rangle}{\langle b|b\rangle} \tag{1.24}$$

and substitute into the above equation, which reduces to

$$\langle a|a\rangle - \frac{\langle a|b\rangle\langle b|a\rangle}{\langle b|b\rangle} \geq 0 \tag{1.25}$$

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

$$\langle a|a\rangle\langle b|b\rangle \geq \langle a|b\rangle\langle b|a\rangle \tag{1.26}$$

$$\geq |\langle a|b\rangle|^2 \tag{1.27}$$

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*).

1.2.3 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

$$\||a\rangle - |a_n\rangle\| \leq \epsilon \tag{1.28}$$

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

$$\lim_{n\to\infty} \||a\rangle - |a_n\rangle\| = 0 \tag{1.29}$$

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.

1.2.4 Orthonormal systems

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

$$\langle a|b\rangle = 0 \tag{1.30}$$

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

$$\langle a_n|a_m\rangle = \delta_{n,m} \tag{1.31}$$

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

$$|a\rangle = \sum_n \alpha_n |a_n\rangle \tag{1.32}$$

where in general α_n are complex numbers whose values are

$$\alpha_m = \langle a_m|a\rangle \tag{1.33}$$

Proof:

$$\begin{aligned} \langle a_m|a\rangle &= \langle a_m|\Big(\sum_n \alpha_n|a_n\rangle\Big) \\ &= \sum_n \alpha_n \langle a_m|a_n\rangle \\ &= \sum_n \alpha_n \delta_{m,n} \\ &= \alpha_m \end{aligned} \tag{1.34}$$

Thus we can write

$$|a\rangle = \sum_n |a_n\rangle\langle a_n|a\rangle \tag{1.35}$$

Note that this implies

$$\hat{I} = \sum_n |a_n\rangle\langle a_n| \tag{1.36}$$

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 $\{|a_n\rangle\}$.

1.3 Operators on Hilbert space

1.3.1 Definitions

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

$$\hat{A}(\alpha|a\rangle + \beta|b\rangle) = \alpha\hat{A}|a\rangle + \beta\hat{A}|b\rangle \tag{1.37}$$

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

$$\|\hat{A}|a\rangle\| \leq C\||a\rangle\| \tag{1.38}$$

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

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

$$|a_n\rangle \to |a\rangle \tag{1.39}$$

then it follows that

$$\hat{A}|a_n\rangle \to \hat{A}|a\rangle \tag{1.40}$$

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

$$\hat{A}|a\rangle = \hat{B}|a\rangle \tag{1.41}$$

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

Unit operator, $\hat{I}$

$$\hat{I}|a\rangle = |a\rangle \tag{1.42}$$

Zero operator, $\hat{0}$

$$\hat{0}|a\rangle = |\text{null}\rangle \tag{1.43}$$

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

$$(\hat{A} + \hat{B})|a\rangle = \hat{A}|a\rangle + \hat{B}|a\rangle \tag{1.44}$$

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

$$(\hat{A}\hat{B})|a\rangle = \hat{A}(\hat{B}|a\rangle) \tag{1.45}$$

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

$$(|b\rangle, \hat{A}|a\rangle) = (\hat{A}^\dagger|b\rangle, |a\rangle) \tag{1.46}$$

or

$$\langle b|\hat{A}|a\rangle = \langle a|\hat{A}^\dagger|b\rangle^* \tag{1.47}$$

The adjoint of an operator has the following properties:

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

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

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger \tag{1.50}$$

$$(\hat{A}^\dagger)^\dagger = \hat{A} \tag{1.51}$$

Hermitian operator : If $\hat{A}$ is self-adjoint it is said to be *Hermitian*. Then

$$\begin{aligned} \hat{A} &= \hat{A}^\dagger \\ \langle b|\hat{A}|b\rangle &= \langle b|\hat{A}^\dagger|b\rangle \\ &= \langle b|\hat{A}^\dagger|b\rangle^* \\ &= \langle b|\hat{A}|b\rangle^* \\ &= \text{real} \end{aligned} \tag{1.52}$$

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

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I} \tag{1.53}$$

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 $|a\rangle$

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

We write this symbolically as

$$\hat{P} = |a\rangle\langle a| \tag{1.55}$$

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

$$\hat{P}^2 = |a\rangle\langle a|a\rangle\langle a| = |a\rangle\langle a| \tag{1.56}$$

since $|a\rangle$ is normalised. Note that the resolution of the identity (1.36) is a sum of projection operators.

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

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

Note that in general

$$\hat{A}\hat{B} \neq \hat{B}\hat{A} \tag{1.58}$$

Properties of commutators:

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

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

$$[\hat{A}, \hat{B}\hat{C}] = [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}] \tag{1.61}$$

$$[\hat{A}, [\hat{B}, \hat{C}]] + [\hat{B}, [\hat{C}, \hat{A}]] + [\hat{C}, [\hat{A}, \hat{B}]] = \hat{0} \tag{1.62}$$

$$[\hat{A}, \hat{B}]^\dagger = [\hat{B}^\dagger, \hat{A}^\dagger] \tag{1.63}$$

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

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

1.3.2 Eigenvectors and eigenvalues

If

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

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 are true:

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

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

$$\hat{A}|a'\rangle = a'|a'\rangle \tag{1.66}$$

and if $a \neq a'$, then

$$\langle a|a'\rangle = 0 \tag{1.67}$$

- In addition, the normalised eigenvectors of a **bounded** Hermitian operator give rise to a *countable, complete orthonormal system.* The eigenvalues form a discrete spectrum.

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 $\hat{A}$:

$$|\psi\rangle = \sum_a |a\rangle\langle a|\psi\rangle \tag{1.68}$$

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

Problem 1: The operator $\hat{Q}$ satisfies the equations

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

$$\hat{Q}\hat{Q}^\dagger + \hat{Q}^\dagger \hat{Q} = \hat{I} \tag{1.69}$$

The Hamiltonian for the system is given by

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

where α is a real constant.

- 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 2 : 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**

  $$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger \hat{A}^\dagger \tag{1.70}$$

  and in general

  $$(\hat{A}\hat{B}\hat{C}\ldots)^\dagger = \ldots \hat{C}^\dagger \hat{B}^\dagger \hat{A}^\dagger \tag{1.71}$$

- **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 $\hat{A}^2$ is 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

  $$[\hat{A}, \hat{B}] = \hat{B} + 2\hat{B}\hat{A}^2 \tag{1.72}$$

  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:

$$\begin{aligned}
[\hat{A},\hat{B}^n] &= \hat{A}\hat{B}^n - \hat{B}^n\hat{A} \qquad (1.73)\\
&= \hat{A}\hat{B}\hat{B}^{n-1} - \hat{B}\hat{A}\hat{B}^{n-1} + \hat{B}(\hat{A}\hat{B})\hat{B}^{n-2} - \hat{B}(\hat{B}\hat{A})\hat{B}^{n-3} + \cdots \hat{B}^{n-1}\hat{A}\hat{B} - \hat{B}^{n-1}\hat{B}\hat{A}\\
&= [\hat{A},\hat{B}]\hat{B}^{n-1} + \hat{B}[\hat{A},\hat{B}]\hat{B}^{n-2} + \cdots + \hat{B}^{n-1}[\hat{A},\hat{B}]
\end{aligned}$$

  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

$$\begin{aligned}
[\hat{A},f(\hat{B})] &= [\hat{A},\sum_n a_n\hat{B}^n] \qquad (1.74)\\
&= \sum_n a_n[\hat{A},\hat{B}^n]\\
&= [\hat{A},\hat{B}]\sum_n n\ a_n\ \hat{B}^{n-1}\\
&= [\hat{A},\hat{B}]f'(\hat{B})
\end{aligned}$$

- (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:

$$\begin{aligned} \frac{d\hat{F}}{ds} &= \left(\frac{d}{ds}e^{s\hat{A}}\right)e^{s\hat{B}} + e^{s\hat{A}}\left(\frac{d}{ds}e^{s\hat{B}}\right) \\ &= \hat{A}e^{s\hat{A}}e^{s\hat{B}} + e^{s\hat{A}}\hat{B}e^{s\hat{B}} \\ &= \hat{A}e^{s\hat{A}}e^{s\hat{B}} + e^{s\hat{A}}\hat{B}e^{-s\hat{A}}e^{s\hat{A}}e^{s\hat{B}} \\ &= \left[\hat{A} + e^{s\hat{A}}\hat{B}e^{-s\hat{A}}\right]\hat{F}(s) \end{aligned} \tag{1.75}$$

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 $\hat{B}$:

$$e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + [\hat{A}, \hat{B}] + \frac{1}{2!}[\hat{A}, [\hat{A}, \hat{B}]] + \frac{1}{3!}[\hat{A}, [\hat{A}, [\hat{A}, \hat{B}]]] + \cdots \tag{1.76}$$

Answer: To show this, define

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

where λ is a real parameter. Then,

$$\begin{aligned} f(0) &= \hat{B} \\ f(1) &= e^{\hat{A}}\hat{B}e^{-\hat{A}} \end{aligned} \tag{1.77}$$

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

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.

1.3.3 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

$$|\psi\rangle = \sum_{n=1}^{D} |\psi_n\rangle\langle\psi_n|\psi\rangle \tag{1.78}$$

(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

$$P(a_n) = |\langle\psi_n|\psi\rangle|^2 \tag{1.79}$$

(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

$$\sum_{n=1}^{D} P(a_n) = 1 \tag{1.80}$$

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

$$\hat{A} = \sum_n a_n |\psi_n\rangle\langle\psi_n| \tag{1.81}$$

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 (1.79) and (1.81), we have

$$\begin{aligned} \langle\hat{A}\rangle &= \sum_n a_n P(a_n) = \sum_n a_n |\langle\psi_n|\psi\rangle|^2 \\ &= \sum_n a_n \langle\psi|\psi_n\rangle\langle\psi_n|\psi\rangle = \langle\psi|\hat{A}|\psi\rangle \end{aligned} \tag{1.82}$$

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

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

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

$$\begin{aligned} \langle 1|\hat{V}|1\rangle &= 0 \\ \langle 1|\hat{V}|2\rangle &= \langle 2|\hat{V}|1\rangle = V \\ \langle 2|\hat{V}|2\rangle &= 0 \end{aligned} \tag{1.84}$$

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

Answer: Eigenvalues are $E_0 \pm V$, with corresponding eigenvectors $(|1\rangle \pm |2\rangle)/\sqrt{2}$.

1.3.4 Generalised uncertainty principle

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

$$[\hat{A}, \hat{B}] = i\hat{C} \tag{1.85}$$

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

$$\Delta A \, \Delta B \geq \frac{1}{2} \left| \langle \hat{C} \rangle \right| \tag{1.86}$$

where

$$\Delta A = \left[\langle (\hat{A} - \langle \hat{A} \rangle)^2 \rangle \right]^{\frac{1}{2}} \tag{1.87}$$

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 4: Prove (1.86).

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

$$\langle a_{n'}, b_{p'}, c_{q'}, \ldots | a_n, b_p, c_q, \ldots \rangle = \delta_{n'n}\delta_{p'p}\delta_{q'q} \ldots \tag{1.88}$$

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.

Given a set of observables $\hat{A}, \hat{B}, \ldots$, any one of the following conditions implies 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).

1.3.5 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

$$|\phi_n\rangle = \hat{U}|\psi_n\rangle \tag{1.89}$$

where

$$\hat{U} = \sum_i |\phi_i\rangle\langle\psi_i| \tag{1.90}$$

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 5:

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

  $$\hat{A}|\psi_n\rangle = a_n|\psi_n\rangle \tag{1.91}$$

  show that the operator

  $$\hat{A}' = \hat{U}\hat{A}\hat{U}^\dagger \tag{1.92}$$

  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.

1.3.6 Matrix representation of operators

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

$$\hat{A} = \hat{I}\hat{A}\hat{I} = \sum_n \sum_{n'} |n'\rangle\langle n'|\hat{A}|n\rangle\langle n| \tag{1.93}$$

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	$\lvert A\rvert = 0$	

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

Problem 6:

- 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.

1.3.7 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

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\delta(x)dx &= f(0) \\ \int_{-\infty}^{\infty} \delta(x)dx &= 1 \\ \int_{-\infty}^{\infty} f(x')\delta(x - x')dx' &= f(x) \\ \int_{-\infty}^{\infty} \delta(x - x')dx' &= 1 \end{aligned} \tag{1.94}$$

Note that

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

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

In three dimensions, the above definition is generalised as follows

$$\int_{all\ space} f(\mathbf{r})\delta(\mathbf{r}-\mathbf{a})\,d\mathbf{r} = f(\mathbf{a}) \tag{1.95}$$

An important property is the Fourier representation

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx}\,dk \tag{1.96}$$

or in three dimensions

$$\delta(\mathbf{r}) = \frac{1}{(2\pi)^3}\int_{all\ k-space} e^{i\mathbf{k}\cdot\mathbf{r}}\,d\mathbf{k} \tag{1.97}$$

Some useful properties

$$\begin{aligned}
\delta(x) &= \delta(-x) \\
\delta'(x) &= -\delta'(-x) \\
x\,\delta(x) &= 0 \\
\delta(a\,x) &= \frac{1}{|a|}\,\delta(x) \\
\delta(x^2-a^2) &= \frac{1}{|2a|}\Big[\delta(x-a)-\delta(x+a)\Big] \\
\int_{-\infty}^{\infty}\delta(a-x)\,\delta(x-b)dx &= \delta(a-b) \\
f(x)\,\delta(x-a) &= f(a)\,\delta(x-a) \\
x\,\delta'(x) &= -\delta(x) \\
\int g(x)\,\delta\Big[f(x)-a\Big]\,dx &= \sum_{x_0}\frac{g(x)}{|df/dx|}\Big|_{x=x_0,\,f(x_0)=a}
\end{aligned} \tag{1.98}$$

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

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

we proceed as follows

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)x\ \delta^{'}(x)dx &= \int_{-\infty}^{\infty}\frac{d}{dx}\Big(f\ x\ \delta\Big)dx - \int_{-\infty}^{\infty}\delta\ \frac{d}{dx}\Big(f\ x\ \Big)dx \\ &= -\int_{-\infty}^{\infty}\delta(x)\Big(x\frac{df}{dx}+f\Big)dx \\ &= -\int_{-\infty}^{\infty}\delta(x)f(x)dx \end{aligned} \tag{1.99}$$

where we have used integration by parts.

1.3.8 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:

$$\langle a|a'\rangle = \delta(a-a') \tag{1.100}$$

The resolution of the identity (or closure relation) becomes

$$\int da\,|a\rangle\langle a| = \hat{I} \tag{1.101}$$

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

$$|\psi\rangle = \int da'|a'\rangle\langle a'|\psi\rangle \tag{1.102}$$

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

$$\begin{aligned}\langle\psi|\phi\rangle &= \int da'\langle\psi|a'\rangle\langle a'|\phi\rangle \\ &= \int \psi^*(a')\phi(a')da' \end{aligned} \tag{1.103}$$

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

$$|\psi\rangle = \sum_{a'} |a'\rangle\langle a'|\psi\rangle + \int |a'\rangle\langle a'|\psi\rangle da' \tag{1.104}$$

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

1.3.9 Example: 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:

$$\begin{aligned} V(x) &= -V_0 \quad \text{for } |x| < \tfrac{1}{2}a \\ V(x) &= 0 \quad \text{elsewhere} \end{aligned} \tag{1.105}$$

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:

$$V(x) = -\mathcal{V}\,\delta(x) \tag{1.106}$$

(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

$$\psi(x) = Ae^{-\kappa|x|} \tag{1.107}$$

where

$$\kappa = \frac{\sqrt{-2mE}}{\hbar} \tag{1.108}$$

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

$$E\psi = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} - \mathcal{V}\,\delta(x)\psi \tag{1.109}$$

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

$$E\int_{-\epsilon}^{+\epsilon} \psi\, dx = -\frac{\hbar^2}{2m}\left[\left(\frac{\partial\psi}{\partial x}\right)_{x=+\epsilon} - \left(\frac{\partial\psi}{\partial x}\right)_{x=-\epsilon}\right] - \mathcal{V}\psi(0) \tag{1.110}$$

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

$$\lim_{\epsilon\to 0}\left[\left(\frac{\partial\psi}{\partial x}\right)_{x=+\epsilon} - \left(\frac{\partial\psi}{\partial x}\right)_{x=-\epsilon}\right] = -\frac{2m\mathcal{V}}{\hbar^2}\psi(0) \tag{1.111}$$

Inserting the form (1.107) for the solution, we find that

$$\kappa = \frac{m\mathcal{V}}{\hbar^2} \tag{1.112}$$

and hence

$$E = -\frac{1}{2}m\left(\frac{\mathcal{V}}{\hbar}\right)^2 \tag{1.113}$$

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 (1.111) 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

$$\psi(x) = C\sin kx \tag{1.114}$$

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

$$\psi(x) = C\cos(k|x| + \phi) \tag{1.115}$$

where, from equation (1.111), the phase shift ϕ satisfies the equation

$$\tan\phi = \frac{m\mathcal{V}}{\hbar^2 k} = \frac{\kappa}{k} \tag{1.116}$$

The presence of the phase shift guarantees that the positive-energy even-parity eigenstates (1.115) are orthogonal to the bound state (1.107). To see this, consider the overlap integral

$$\begin{aligned}
\int_{-\infty}^{\infty}\cos(k|x|+\phi)e^{-\kappa|x|}dx &= 2\int_0^{\infty}\cos(kx+\phi)e^{-\kappa x}dx \\
&= \int_0^{\infty}\left(e^{ikx+i\phi-\kappa x} + e^{-ikx-i\phi-\kappa x}\right)dx \\
&= \frac{e^{i\phi}}{\kappa - ik} + \frac{e^{-i\phi}}{\kappa + ik} \\
&= \frac{2}{\kappa^2+k^2}(\kappa\cos\phi - k\sin\phi) \\
&= 0
\end{aligned} \tag{1.117}$$

on account of equation (1.116). Hence any unbound state of the system can be Fourier decomposed into sine waves and the "kinked cosine" waves (1.115). Conversely, the square modulus of the overlap between any normalised state and the bound state (1.107) gives the probability that the particle is bound in the potential well.

Problem 7: Find the probability P_b that a particle with wave function

$$\psi(x) = \sqrt{c}\, e^{-c|x|} \tag{1.118}$$

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 8†[1]: Confirm that

$$\int_0^\infty P(k)\, dk = 1 - P_b \tag{1.119}$$

---

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

Chapter 2

Quantum Dynamics

2.1 Time development operator

So far, we have presented state vectors as fixed objects in a Hilbert space, but in general we expect them to change with time, in accordance with Postulate 6 of quantum mechanics. Suppose that the state vector at some initial time t_0 is $|\Psi(t_0)\rangle$ and that some other time t it is $|\Psi(t)\rangle$. Then we may define the *time development operator* $\hat{T}(t, t_0)$ which relates the two:

$$|\Psi(t)\rangle = \hat{T}(t, t_0)|\Psi(t_0)\rangle \tag{2.1}$$

The *principle of superposition* means that any linear superposition of initial states $\alpha|\Psi_a(t_0)\rangle + \beta|\Psi_b(t_0)\rangle$ will evolve into the corresponding linear superposition $\alpha|\Psi_a(t)\rangle + \beta|\Psi_b(t)\rangle$, in other words that $\hat{T}(t, t_0)$ is a linear operator. Clearly, it has the property

$$\hat{T}(t_0, t_0) = \hat{I} \tag{2.2}$$

Furthermore, if the system evolves from t_0 to t_1 and then from t_1 to t_2, we have

$$|\Psi(t_2)\rangle = \hat{T}(t_2, t_1)|\Psi(t_1)\rangle = \hat{T}(t_2, t_1)\hat{T}(t_1, t_0)|\Psi(t_0)\rangle \tag{2.3}$$

and therefore, since this is true for any initial state,

$$\hat{T}(t_2, t_0) = \hat{T}(t_2, t_1)\hat{T}(t_1, t_0) \tag{2.4}$$

In particular, we have

$$\hat{T}(t_0, t_0) = \hat{I} = \hat{T}(t_0, t)\hat{T}(t, t_0) \tag{2.5}$$

from which it follows that

$$\hat{T}(t_0, t) = \left[\hat{T}(t, t_0)\right]^{-1} \tag{2.6}$$

To find the explicit form of the time development operator, we note from Postulate 6 that

$$\begin{aligned}\frac{\partial}{\partial t}|\Psi(t)\rangle &= \left[\frac{\partial}{\partial t}\hat{T}(t,t_0)\right]|\Psi(t_0)\rangle \\ &= -\frac{i}{\hbar}\hat{H}|\Psi(t)\rangle \\ &= -\frac{i}{\hbar}\hat{H}\hat{T}(t,t_0)|\Psi(t_0)\rangle \end{aligned} \tag{2.7}$$

Again, this must be true for any $|\Psi(t_0)\rangle$ and therefore the operator equation

$$\frac{\partial}{\partial t}\hat{T}(t,t_0) = -\frac{i}{\hbar}\hat{H}\hat{T}(t,t_0) \tag{2.8}$$

must hold for all t and t_0. Taking into account the initial condition (2.2), and assuming the Hamiltonian operator does not depend explicitly on time, the solution of this differential equation is

$$\hat{T}(t,t_0) = \exp\left[-\frac{i}{\hbar}\hat{H}(t-t_0)\right] \tag{2.9}$$

Since the Hamiltonian operator is Hermitian, the time development operator is unitary:

$$\begin{aligned}\hat{T}^\dagger(t,t_0) &= \exp\left[+\frac{i}{\hbar}\hat{H}^\dagger(t-t_0)\right] \\ &= \exp\left[+\frac{i}{\hbar}\hat{H}(t-t_0)\right] \\ &= \left[\hat{T}(t,t_0)\right]^{-1} \end{aligned} \tag{2.10}$$

It follows that the norm of states is preserved under time development:

$$\|\Psi(t)\|^2 = \langle\Psi(t)|\Psi(t)\rangle = \langle\Psi(t_0)|\hat{T}^\dagger(t,t_0)\hat{T}(t,t_0)|\Psi(t_0)\rangle = \langle\Psi(t_0)|\Psi(t_0)\rangle = \|\Psi(t_0)\|^2 \tag{2.11}$$

The time-dependence of observables is controlled by the same time development operator. Denoting the expectation value of observable $\hat{A}$ at time t by $\langle\hat{A}\rangle_t$, we have

$$\begin{aligned}\langle\hat{A}\rangle_t &= \langle\Psi(t)|\hat{A}(t)|\Psi(t)\rangle \\ &= \langle\Psi(t_0)|\hat{T}^\dagger(t,t_0)\hat{A}(t)\hat{T}(t,t_0)|\Psi(t_0)\rangle \\ &= \langle\hat{T}^\dagger(t,t_0)\hat{A}(t)\hat{T}(t,t_0)\rangle_{t_0} \end{aligned} \tag{2.12}$$

Here we have allowed for the possibility that the operator $\hat{A}$ depends explicitly on time. If it does not, and if it commutes with the Hamiltonian operator, then it also commutes with the time

development operator and we have

$$\langle \hat{A} \rangle_t = \langle \hat{T}^\dagger(t,t_0)\hat{A}\hat{T}(t,t_0) \rangle_{t_0} = \langle \hat{T}^\dagger(t,t_0)\hat{T}(t,t_0)\hat{A} \rangle_{t_0} = \langle \hat{A} \rangle_{t_0} \tag{2.13}$$

so that the observable $\hat{A}$ is a *constant of the motion*.

Problem 1: Derive the differential equation

$$\frac{d}{dt}\langle \hat{A} \rangle = \frac{i}{\hbar}\langle [\hat{H}, \hat{A}] \rangle + \langle \frac{\partial}{\partial t}\hat{A} \rangle \ , \tag{2.14}$$

which again shows that $\langle \hat{A} \rangle$ is constant if $\hat{A}$ does not depend explicitly on time and commutes with $\hat{H}$.

2.2 Schrödinger, Heisenberg and interaction pictures

The formulation of quantum dynamics that we have used so far, in the above discussion and in lecture courses, treats the time evolution of a system as a property of the state vectors; operators representing observables that do not depend explicitly on time, such as position and momentum, are regarded as time-independent. This is called the *Schrödinger picture* of quantum dynamics. It seems very far from the classical picture, in which for example the position and momentum of a particle are continuously changing under the influence of forces.

The Schrödinger picture is in fact only one of an infinite range of equivalent formulations of quantum dynamics, related by unitary transformations of the Hilbert space as discussed in Section 1.3.5. We saw there that all the fundamental properties of the system are unaffected by a transformation of the state vectors of the form

$$|\Psi\rangle \rightarrow |\Psi'\rangle = \hat{U}|\Psi\rangle \tag{2.15}$$

where $\hat{U}$ is unitary, provided the observables of the system are also transformed according to

$$\hat{A} \rightarrow \hat{A}' = \hat{U}\hat{A}\hat{U}^\dagger \tag{2.16}$$

If the unitary operator $\hat{U}$ is time-dependent, the transformation (2.16) will introduce time-dependence into the operator $\hat{A}'$ representing *the same observable* in the new formulation as was originally represented by the time-independent operator $\hat{A}$ in the Schrödinger picture.

Consider in particular the effect of choosing the transformation $\hat{U}$ to be the adjoint (that is, the inverse) of the time development operator. The transformation of the state vectors is

$$|\Psi(t)\rangle \rightarrow \hat{T}^\dagger(t,t_0)|\Psi(t)\rangle = T(t_0,t)T(t,t_0)|\Psi(t_0)\rangle = |\Psi(t_0)\rangle \tag{2.17}$$

Thus, in this picture, all the time dependence of the state vectors has been removed: they remain static in the Hilbert space. The observables, on the other hand, have absorbed all the time dependence: they are represented by operators of the form

$$\hat{A}'(t) = \hat{T}^\dagger(t,t_0)\hat{A}\hat{T}(t,t_0) = \hat{T}(t_0,t)\hat{A}\hat{T}(t,t_0) \tag{2.18}$$

This formulation is called the *Heisenberg picture*. In many respects it looks closer to classical mechanics, because the operators representing observables obey equations of motion similar to those of the corresponding classical observables. Differentiating eq. (2.18) and using eq. (2.8), we find

$$\frac{d}{dt}\hat{A}' = \frac{i}{\hbar}\left[\hat{H}', \hat{A}'\right] + \frac{\partial}{\partial t}\hat{A}' \tag{2.19}$$

When the Schrödinger operator $\hat{A}$ is time-independent and commutes with the Hamiltonian, then so also is the Heisenberg operator $\hat{A}'$, in fact $\hat{A}' = \hat{A}$, since it represents a constant of the motion. In particular, the Hamiltonian itself (when not explicitly time-dependent) is constant and represents the energy of the system.

On the other hand, when $\hat{A}'$ does not commute with the Hamiltonian then eq. (2.19) is the analogue of the classical equation of motion. Consider for example the motion of a particle in a potential $V(\mathbf{r})$. The Heisenberg momentum operator $\hat{\mathbf{p}}'$ satifies the equation

$$\frac{d}{dt}\hat{\mathbf{p}}' = \frac{i}{\hbar}\left[\hat{H}', \hat{\mathbf{p}}'\right] \tag{2.20}$$

where the Hamiltonian is

$$\hat{H}' = \frac{1}{2m}\hat{\mathbf{p}}'^2 + V(\hat{\mathbf{r}}') \tag{2.21}$$

The Schrödinger and Heisenberg position and momentum operators satisfy the same commutation relations:

$$\left[\hat{r}'_j, \hat{p}'_k\right] = [\hat{r}_j, \hat{p}_k] = i\hbar\,\delta_{jk} \tag{2.22}$$

and hence

$$\left[\hat{H}', \hat{\mathbf{p}}'\right] = i\hbar\nabla V(\hat{\mathbf{r}}') \tag{2.23}$$

so that

$$\frac{d}{dt}\hat{\mathbf{p}}' = -\nabla V(\hat{\mathbf{r}}') = \hat{\mathbf{F}}(\hat{\mathbf{r}}') \tag{2.24}$$

where $\hat{\mathbf{F}}$ is the force operator. This is just Newton's second law, but now it involves quantum mechanical operators in the Heisenberg picture.

Problem 2: Prove the results (2.22) and (2.23).

A third formulation of quantum dynamics that is often useful is the *interaction picture*. Here we partition the Hamiltonian into an interaction-free part $\hat{H}_0$ ($\hat{\mathbf{p}}^2/2m$ for a single particle) and an interaction term $\hat{V}$. Then we make a unitary transformation that involves only the interaction-free part of the Hamiltonian:

$$|\Psi(t)\rangle \to |\Psi_i(t)\rangle = \hat{U}(t,t_0)|\Psi(t)\rangle \tag{2.25}$$

and correspondingly

$$\hat{A} \to \hat{A}_i = \hat{U}(t,t_0)\hat{A}\hat{U}^\dagger(t,t_0) \tag{2.26}$$

where

$$\hat{U}(t,t_0) = \exp\left[\frac{i}{\hbar}(t-t_0)\hat{H}_0\right] \tag{2.27}$$

The advantage of the interaction picture is that the only time dependence remaining in the state vectors is that associated with the interaction:

$$\frac{\partial}{\partial t}|\Psi_i(t)\rangle = -\frac{i}{\hbar}\hat{V}_i|\Psi_i(t)\rangle \tag{2.28}$$

where $\hat{V}_i$ is the transformed interaction term,

$$\hat{V}_i = \hat{U}(t,t_0)\hat{V}\hat{U}^\dagger(t,t_0)\ . \tag{2.29}$$

On the other hand, the time dependence of the transformed operators is given entirely by the interaction-free part of the Hamiltonian:

$$\frac{d}{dt}\hat{A}_i = \frac{i}{\hbar}\left[\hat{H}_0, \hat{A}_i\right] + \frac{\partial}{\partial t}\hat{A}_i \tag{2.30}$$

Notice that by setting $\hat{H}_0 = 0$ or $\hat{H}_0 = \hat{H}$ in the interaction picture we recover the Schrödinger or Heisenberg picture, respectively.

Problem 3: Prove the results (2.28) and (2.30).

2.3 Canonical quantisation and constants of motion

We can generalize the relation between the classical equations of motion and the Heisenberg picture of quantum dynamics, and this leads to a general rule for quantising a classical system. We recall

first[1] that if q is a classical generalized coordinate and p is its conjugate momentum, then Hamilton's equations of motion in classical mechanics are

$$\frac{dq}{dt} = \frac{\partial H}{\partial p} \ , \quad \frac{dp}{dt} = -\frac{\partial H}{\partial q} \tag{2.31}$$

Therefore for any classical observable $A(q,p,t)$ we have

$$\begin{aligned} \frac{dA}{dt} &= \frac{\partial A}{\partial q}\frac{dq}{dt} + \frac{\partial A}{\partial p}\frac{dp}{dt} + \frac{\partial A}{\partial t} \\ &= \frac{\partial A}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial A}{\partial p}\frac{\partial H}{\partial q} + \frac{\partial A}{\partial t} \\ &= [A,H]_{PB} + \frac{\partial A}{\partial t} \end{aligned} \tag{2.32}$$

where $[\cdots]_{PB}$ denotes the classical *Poisson bracket*

$$[A,B]_{PB} \equiv \frac{\partial A}{\partial q}\frac{\partial B}{\partial p} - \frac{\partial A}{\partial p}\frac{\partial B}{\partial q} \tag{2.33}$$

Comparing with the quantum mechanical equation of motion in the Heisenberg picture, eq. (2.19), we can formulate the following rules for the *canonical quantisation* of a classical system:

1. Replace the classical observables by quantum mechanical observables *in the Heisenberg picture*, $A \to \hat{A}'$;
2. Replace the classical Poisson brackets by quantum mechanical commutators according to the prescription
$$[A,B]_{PB} \to \frac{1}{i\hbar}[\hat{A}', \hat{B}'] \tag{2.34}$$

Then we see immediately that $[q,p]_{PB} = 1$ implies $[\hat{q}', \hat{p}'] = i\hbar$.

In defining the quantum Hamiltonian $\hat{H}$, care must be taken in ordering of operators, so as to obtain a Hermitian Hamiltonian operator. For example, if the term pq appears in the classical Hamiltonian, it becomes $\frac{1}{2}(\hat{p}'\hat{q}' + \hat{q}'\hat{p}')$. Where there are ambiguities in this procedure, they represent genuinely different quantum systems with the same classical limit, since the differences vanish in the limit $\hbar \to 0$.

Problem 4: Suppose the classical Hamiltonian involves p^2q^2, which can be written as $\frac{1}{2}(p^2q^2 + q^2p^2)$, pq^2p or qp^2q, for example. Show that canonical quantisation of these three expressions yields Hermitian operators, and evaluate the differences between them.

[1] Those not familiar with the Hamiltonian formulation of classical mechanics (covered in the TP1 course) can find a good concise exposition in Kibble and Berkshire, *Classical Mechanics*, Ch. 12.

2.4 Position and momentum representations

For a system consisting of a single particle, we can use the eigenstates of the position operator or the momentum operator as a basis. The states of the system, and operators acting on them, are then said to be represented in the position or momentum representation, respectively. We can find the relationship between them as follows.

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

$$\begin{aligned} \hat{x}|x'\rangle &= x'|x'\rangle \\ \hat{p}|p'\rangle &= p'|p'\rangle \\ \langle x|x'\rangle &= \delta(x-x') \\ \langle p|p'\rangle &= \delta(p-p') \end{aligned} \tag{2.35}$$

These definitions, the fundamental commutator

$$[\hat{x},\hat{p}] = i\hbar \tag{2.36}$$

and the properties of the Dirac delta function in Section 1.3.7 can be used to determine the following matrix elements:

$$\begin{aligned} \langle x'|\hat{p}|x''\rangle &= \frac{\hbar}{i}\frac{\partial}{\partial x'}\delta(x'-x'') \\ \langle p'|\hat{x}|p''\rangle &= -\frac{\hbar}{i}\frac{\partial}{\partial p'}\delta(p'-p'') \\ \langle x'|\hat{p}^2|x''\rangle &= \left(-i\hbar\frac{\partial}{\partial x'}\right)^2\delta(x'-x'') \\ \langle p'|\hat{x}^2|p''\rangle &= \left(i\hbar\frac{\partial}{\partial p'}\right)^2\delta(p'-p'') \end{aligned} \tag{2.37}$$

Problem 5: Verify the formulae (2.37)

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

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

then we have

$$\begin{aligned}
\langle x'|\hat{p}|p'\rangle &= \int dx''\langle x'|\hat{p}|x''\rangle\langle x''|p'\rangle \\
&= \int dx''\Big(-i\hbar\frac{\partial}{\partial x'}\delta(x'-x'')\Big)\langle x''|p'\rangle \\
&= -i\hbar\frac{\partial}{\partial x'}\int dx''\delta(x'-x'')\langle x''|p'\rangle \\
&= -i\hbar\frac{\partial}{\partial x'}\langle x'|p'\rangle \qquad (2.38)
\end{aligned}$$

On the other hand, we also have

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

Therefore

$$-i\hbar\frac{\partial}{\partial x'}\langle x'|p'\rangle = p'\langle x'|p'\rangle \qquad (2.39)$$

which implies

$$\langle x'|p'\rangle = \frac{1}{\sqrt{2\pi\hbar}}\exp\left(\frac{ip'x'}{\hbar}\right) \qquad (2.40)$$

where we have chosen the normalisation such that

$$\begin{aligned}
\langle p''|p'\rangle &= \int dx'\langle p''|x'\rangle\langle x'|p'\rangle \\
&= \int dx'\langle x'|p''\rangle^*\langle x'|p'\rangle \\
&= \frac{1}{(2\pi\hbar)}\int dx'\exp\left(\frac{i(p'-p'')x'}{\hbar}\right) \\
&= \delta(p''-p') \qquad (2.41)
\end{aligned}$$

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

$$\begin{aligned}
|\mathbf{r}\rangle &= |x,y,z\rangle \\
\hat{\mathbf{r}}|\mathbf{r}\rangle &= \mathbf{r}|\mathbf{r}\rangle \\
\langle\mathbf{r}'|\mathbf{r}''\rangle &= \delta(\mathbf{r}'-\mathbf{r}'')
\end{aligned}$$

$$
\begin{aligned}
|\mathbf{p}\rangle &= |p_x, p_y, p_z\rangle \\
\hat{\mathbf{p}}|\mathbf{p}\rangle &= \mathbf{p}|\mathbf{p}\rangle \\
\langle \mathbf{p}'|\mathbf{p}''\rangle &= \delta(\mathbf{p}' - \mathbf{p}'') \\
\langle \mathbf{r}'|\hat{\mathbf{p}}|\mathbf{r}''\rangle &= -i\hbar\nabla_{\mathbf{r}'}\delta(\mathbf{r}' - \mathbf{r}'') \\
\langle \mathbf{p}'|\hat{\mathbf{r}}|\mathbf{p}''\rangle &= i\hbar\nabla_{\mathbf{p}'}\delta(\mathbf{p}' - \mathbf{p}'') \\
\langle \mathbf{r}|\mathbf{p}\rangle &= \frac{1}{(2\pi\hbar)^{3/2}} \exp\left(i\mathbf{r}\cdot\mathbf{p}/\hbar\right)
\end{aligned} \tag{2.42}
$$

The wave function of a state $|\psi\rangle$ in the position representation is defined as

$$\psi(\mathbf{r}) = \langle \mathbf{r}|\psi\rangle \tag{2.43}$$

and similarly in the momentum representation

$$\phi(\mathbf{p}) = \langle \mathbf{p}|\psi\rangle \tag{2.44}$$

Using the above results we obtain the familiar Fourier relation

$$
\begin{aligned}
\phi(\mathbf{p}) &= \int d\mathbf{r}\, \langle \mathbf{p}|\mathbf{r}\rangle\, \langle \mathbf{r}|\psi\rangle \\
&= \frac{1}{(2\pi\hbar)^{3/2}} \int d\mathbf{r}\, \exp\left(-i\mathbf{r}\cdot\mathbf{p}/\hbar\right) \psi(\mathbf{r})
\end{aligned} \tag{2.45}
$$

and similarly for the inverse relation

$$\psi(\mathbf{r}) = \frac{1}{(2\pi\hbar)^{3/2}} \int d\mathbf{p}\, \exp\left(i\mathbf{r}\cdot\mathbf{p}/\hbar\right) \phi(\mathbf{p}) \tag{2.46}$$

2.5 The propagator in the position representation

A very useful concept in quantum dynamics is the *propagator*, which is defined in the position representation as the amplitude for a transition to position $\mathbf{r}$ at time t from an initial position $\mathbf{r}'$ at time t'. In terms of the time development operator introduced in Section 2.1, we have

$$K(\mathbf{r}, \mathbf{r}'; t, t') = \langle \mathbf{r}|\hat{T}(t, t')|\mathbf{r}'\rangle \tag{2.47}$$

and since $\hat{T}(t, t) = \hat{I}$ the initial condition is

$$K(\mathbf{r}, \mathbf{r}'; t, t) = \langle \mathbf{r}|\mathbf{r}'\rangle = \delta(\mathbf{r} - \mathbf{r}') \tag{2.48}$$

Using the definition (2.43) of the position-space wave function and the resolution of the identity operator

$$\hat{I} = \int d\mathbf{r}' \, |\mathbf{r}'\rangle\langle\mathbf{r}'| \tag{2.49}$$

we can write

$$\begin{aligned} \psi(\mathbf{r},t) &= \langle\mathbf{r}|\psi(t)\rangle = \langle\mathbf{r}|\hat{T}(t,t')|\psi(t')\rangle \\ &= \int d\mathbf{r}' \, \langle\mathbf{r}|\hat{T}(t,t')|\mathbf{r}'\rangle\langle\mathbf{r}'|\psi(t')\rangle = \int d\mathbf{r}' \, K(\mathbf{r},\mathbf{r}';t,t')\psi(\mathbf{r}',t') \end{aligned} \tag{2.50}$$

In other words, the propagator is a *Green function* for the time-dependent Schrödinger equation, enabling us to express the wave function at time t in terms of its form at some initial time t'.

Problem 6: Derive the following properties of the propagator from those of the time development operator

$$\begin{aligned} K(\mathbf{r}',\mathbf{r};t',t) &= K^*(\mathbf{r},\mathbf{r}';t,t') \\ K(\mathbf{r},\mathbf{r}';t,t') &= \int d\mathbf{r}'' \, K(\mathbf{r},\mathbf{r}'';t,t'')K(\mathbf{r}'',\mathbf{r}';t'',t') \end{aligned} \tag{2.51}$$

In the case that the Hamiltonian does not depend explicitly on time, the time development operator is given by eq. (2.9) and so

$$K(\mathbf{r},\mathbf{r}';t,t') = \langle\mathbf{r}| \exp\left[-\frac{i}{\hbar}\hat{H}(t-t')\right] |\mathbf{r}'\rangle \tag{2.52}$$

Now suppose that we know the energy eigenstates of the system, $|\psi_n\rangle$, with energies E_n. Inserting the resolution of the identity operator

$$\hat{I} = \sum_n |\psi_n\rangle\langle\psi_n| \tag{2.53}$$

we have

$$K(\mathbf{r},\mathbf{r}';t,t') = \sum_n \psi_n(\mathbf{r})\psi_n^*(\mathbf{r}') \exp\left[-\frac{i}{\hbar}E_n(t-t')\right] \tag{2.54}$$

2.5.1 Free particle propagator

For a free particle the Hamiltonian is simply

$$\hat{H} = \hat{\mathbf{p}}^2/2m \tag{2.55}$$

and momentum eigenstates are also energy eigenstates. Therefore the sum in eq. (2.54) becomes an integral over momenta and we can write

$$\begin{aligned} K(\mathbf{r},\mathbf{r}';t,t') &= \int d\mathbf{p}\,\langle\mathbf{r}|\mathbf{p}\rangle\langle\mathbf{p}|\mathbf{r}'\rangle \exp\left[-\frac{i\mathbf{p}^2}{2m\hbar}(t-t')\right] \\ &= \frac{1}{(2\pi\hbar)^3}\int d\mathbf{p}\,\exp\left\{\frac{i}{\hbar}\left[\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')-\frac{\mathbf{p}^2}{2m}(t-t')\right]\right\} \end{aligned} \tag{2.56}$$

This can be reduced to a standard integral by completing the square in the exponent. Defining

$$\mathbf{p}' = \mathbf{p} - m\frac{\mathbf{r}-\mathbf{r}'}{t-t'} \tag{2.57}$$

we have

$$\begin{aligned} K(\mathbf{r},\mathbf{r}';t,t') &= \frac{1}{(2\pi\hbar)^3}\int d\mathbf{p}'\,\exp\left\{\frac{i}{\hbar}\left[\frac{m(\mathbf{r}-\mathbf{r}')^2}{2(t-t')}-\frac{\mathbf{p}'^2}{2m}(t-t')\right]\right\} \\ &= \left[\frac{-im}{2\pi\hbar(t-t')}\right]^{3/2}\exp\left[\frac{im(\mathbf{r}-\mathbf{r}')^2}{2\hbar(t-t')}\right] \end{aligned} \tag{2.58}$$

This expression has a number of important properties that carry over to more general cases. First we note that it is a function of $\mathbf{r}-\mathbf{r}'$ and $t-t'$, because the Hamiltonian is invariant under translation of the origin in space and time. Since the Hamiltonian is also invariant under time reversal, the propagator also has the property

$$K(\mathbf{r},\mathbf{r}';t,t') = K^*(\mathbf{r},\mathbf{r}';t',t) \tag{2.59}$$

It follows from this and the fundamental property

$$K(\mathbf{r}',\mathbf{r};t',t) = K^*(\mathbf{r},\mathbf{r}';t,t') \tag{2.60}$$

derived earlier, that the propagator is symmetric under the interchange of $\mathbf{r}$ and $\mathbf{r}'$:

$$K(\mathbf{r},\mathbf{r}';t,t') = K(\mathbf{r}',\mathbf{r};t,t') \tag{2.61}$$

Problem 7: Show that the propagator (2.58) is indeed a Green function for the time-dependent Schrödinger equation, i.e. it satisfies that equation and the boundary condition (2.48). Show further that it gives the expected wave function at $t \neq 0$ when the wave function at $t = 0$ is a plane wave $\psi(\mathbf{r},0) = \exp(i\mathbf{k}\cdot\mathbf{r})$.

2.5.2 Simple harmonic oscillator propagator

A more complicated case is that of a particle moving in a potential $V(\mathbf{r})$. For simplicity we consider motion in one dimension, for which the Hamiltonian is

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}) \tag{2.62}$$

Noting that this expression is invariant under time translation (but not under spatial translation, unless V is constant), we expect the propagator to be a function of $t - t'$,

$$K(x, x'; t, t') = K(x, x'; t - t', 0) \tag{2.63}$$

Therefore, without loss of generality, we can set $t' = 0$. Let us try a solution of the form

$$K(x, x'; t, 0) = \exp\left[\frac{i}{\hbar} S(x, x'; t)\right] \tag{2.64}$$

Since the propagator satisfies the time-dependent Schrödinger equation

$$i\hbar \frac{\partial K}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2 K}{\partial x^2} + V(x)K \tag{2.65}$$

the function S satisfies the non-linear differential equation

$$\frac{\partial S}{\partial t} + \frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2 - \frac{i\hbar}{2m}\frac{\partial^2 S}{\partial x^2} + V(x) = 0 \tag{2.66}$$

Eq. (2.66) is difficult to solve in general. In the case of a free particle, $V(x) = 0$, we see from eq. (2.58) that $S(x, x'; t)$ is a quadratic function of x and x'. In the simple harmonic oscillator potential

$$V(x) = \frac{1}{2}m\omega^2 x^2 \tag{2.67}$$

the extra term in the Hamiltonian is quadratic, and therefore we guess that $S(x, x'; t)$ remains quadratic. Furthermore the Hamiltonian is time-reversal invariant and so $S(x, x'; t)$ must be symmetric in x and x'. We therefore try a solution of the form

$$S(x, x'; t) = a(t)(x^2 + x'^2) + b(t)xx' + c(t) \tag{2.68}$$

which gives

$$\dot{a}(x^2 + x'^2) + \dot{b}xx' + \dot{c} + \frac{1}{2m}(2ax + bx')^2 - \frac{i\hbar a}{m} + \frac{1}{2}m\omega^2 x^2 = 0 \tag{2.69}$$

Since this equation must be true everywhere, the coefficient of each power of x and x' must be zero, giving the following coupled ordinary differential equations

$$\dot{a} = -\frac{2}{m}a^2 - \frac{1}{2}m\omega^2 = -\frac{b^2}{2m}\,, \quad \dot{b} = -\frac{2}{m}ab\,, \quad \dot{c} = \frac{i\hbar}{m}a \tag{2.70}$$

It is not difficult to see that the solution of these equations, satisfying the appropriate initial condition, is

$$a(t) = \frac{1}{2}m\omega \cot \omega t \ , \quad b(t) = -\frac{m\omega}{\sin \omega t} \ , \quad c(t) = \frac{1}{2}i\hbar \log \sin \omega t + \text{constant} \tag{2.71}$$

so that

$$K(x, x'; t, 0) = \frac{C}{\sqrt{\sin \omega t}} \exp \left\{ \frac{im\omega}{2\hbar \sin \omega t} \left[(x^2 + x'^2) \cos \omega t - 2xx' \right] \right\} \tag{2.72}$$

where C is a constant. Notice that in the limit $\omega \to 0$ the potential vanishes and we must recover the free-particle propagator. This fixes the constant as

$$C = \sqrt{\frac{m\omega}{2\pi i\hbar}} \tag{2.73}$$

Problem 8: Recall that the ground-state wave function of the harmonic oscillator has the form

$$\psi_0(x) = N \exp \left[-\frac{m\omega}{2\hbar} x^2 \right] \tag{2.74}$$

where N is a normalization constant. Suppose that the initial state is represented by a displaced ground-state

$$\psi(x, t = 0) = \psi_0(x - x_0) \tag{2.75}$$

Show that $|\psi(x,t)|^2$ oscillates without any change of shape, i.e.

$$|\psi(x,t)|^2 = |\psi_0(x - x_0 \cos \omega t)|^2 \tag{2.76}$$

2.6 Introduction to path integral formulation

Up to now, everything that you have learnt in quantum mechanics has been formulated in terms of vector spaces and operators acting on those spaces. Even simple wave mechanics has been presented in the framework of a vector space of wave functions. In this approach, the Hamiltonian operator plays a central rôle as the generator of time development, culminating in expressions like eq. (2.52) for the propagator.

Surprisingly, there is a completely different but equivalent formulation of quantum mechanics, which does not involve vector spaces or operators and is based on the Lagrangian rather than the Hamiltonian version of of classical mechanics. This is the *path integral formulation* invented by

Feynman. In many ways it is more intuitive, but it involves the mathematics of functional integrals, which is less familiar.

We shall illustrate the path integral formulation in the simple case of a particle moving in a one-dimensional potential $V(x)$. Consider the propagator $K(x_T, x_0; T)$, which gives the probability amplitude for propagation of the particle from position x_0 at time $t = 0$ to position x_T at time T. In the operator formulation we write this as

$$K(x_T, x_0; T) = \langle x_T | e^{-i\hat{H}(\hat{p},\hat{x})T/\hbar} | x_0 \rangle \tag{2.77}$$

where $\hat{H}(\hat{p}, \hat{x})$ is the Hamiltonian operator, given in this case by

$$\hat{H} = \frac{\hat{p}^2}{2m} + \hat{V}(\hat{x}) \tag{2.78}$$

In classical mechanics, the propagation of the particle from x_0 to x_T takes place along the unique trajectory $x(t)$ which is the *path of least action* between these two points. Recall that the action S is defined as the time integral of the Lagrangian,

$$S[x(t)] = \int_0^T L(x, \dot{x})\, dt \tag{2.79}$$

where in this case

$$L(x, \dot{x}) = \frac{1}{2} m\dot{x}^2 - V(x) \tag{2.80}$$

We write $S[x(t)]$ with square brackets to indicate that S is a *functional* of the path $x(t)$. Technically, a functional is a mapping from functions to numbers. Thus any definite integral is a functional of its integrand. The Dirac delta function is also, strictly speaking, a functional, since it only has meaning as a mapping from a function $f(x)$ to its value at a particular point:

$$f(x) \to \int f(x)\, \delta(x - a)\, dx = f(a) \tag{2.81}$$

As should be familiar by now, minimizing (or maximizing) the action functional leads to the Lagrange equation of motion,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = \frac{\partial L}{\partial x} \tag{2.82}$$

which in this case is just Newton's second law:

$$m\ddot{x} = -\frac{dV}{dx} \tag{2.83}$$

In quantum mechanics, particles have wave-like properties and therefore they do not propagate along a unique trajectory. Instead, waves propagate in all directions from the source at x_0, and can

interfere constructively or destructively at the detection point x_T. To find the amplitude at x_T we have to sum the complex amplitudes associated with all possible paths from x_0 to x_T. The path integral formulation asserts that the amplitude for the path $x(t)$ is simply

$$A[x(t)] = e^{iS[x(t)]/\hbar} \tag{2.84}$$

We can see that classical mechanics will follow from this formula in the limit $\hbar \to 0$. For in that case the phases of neighbouring paths will be so rapidly varying that their contributions to the sum will cancel, except for the paths around the trajectory where $S[x(t)]$ is an extremum. In that region of *stationary phase* the contributions will all combine in phase and, in the limit $\hbar \to 0$, the particle will certainly be found there. Thus we recover the result that the classical trajectory is an extremum of the action.

In the quantum-mechanical case we have to sum over all possible paths, and in general there will be an infinite number of them. The expression for the propagator therefore becomes a *functional integral*

$$K(x_T, x_0; T) = \int \mathcal{D}x(t)\, e^{iS[x(t)]/\hbar} \tag{2.85}$$

The notation $\int \mathcal{D}x(t)$ indicates that we have to sum over all paths starting at x_0 and ending at x_T. We shall do this by discretizing the problem: divide the time T into N intervals of duration $\epsilon = T/N$ and specify the path $x(t)$ by the points $x_k = x(k\epsilon)$, $k = 1, \ldots, N$, with $x_N = x_T$. Then we have to integrate over all possible values of $x_1, \ldots, x_{N-1}$, and take the limit $\epsilon \to 0$, $N \to \infty$ with $\epsilon N = T$ fixed, at the end. Thus we define

$$\int \mathcal{D}x(t) = \mathrm{Lim}\, [C(\epsilon)]^N \, dx_1 \ldots dx_{N-1} \tag{2.86}$$

where $C(\epsilon)$ is a constant factor in the contribution from each segment of the path, to be determined.

To justify eq. (2.85), and to determine $C(\epsilon)$, we go back to the expression (2.77) for the propagator in the operator formulation and write

$$e^{-i\hat{H}T/\hbar} = \mathrm{Lim} \left(1 - \frac{i\hat{H}T}{\hbar N}\right)^N \tag{2.87}$$

Now we insert the resolution of the identity

$$\hat{I} = \int dx_k \, |x_k\rangle\langle x_k| \tag{2.88}$$

between every factor, so that

$$K(x_T, x_0; T) = \mathrm{Lim} \int \left(\prod_{k=1}^{N-1} dx_k\right) \prod_{k=0}^{N-1} \langle x_{k+1}|(1 - i\hat{H}\epsilon/\hbar)|x_k\rangle \tag{2.89}$$

Each factor in the product is now easy to evaluate. Introducing

$$\hat{I} = \int dp_k \, |p_k\rangle\langle p_k| \tag{2.90}$$

we have

$$\langle x_{k+1}|\hat{p}^2|x_k\rangle = \int dp_k \, p_k^2 \langle x_{k+1}|p_k\rangle\langle p_k|x_k\rangle = \int \frac{dp_k}{2\pi\hbar} \, p_k^2 \, \exp\left[ip_k(x_{k+1} - x_k)\right] \tag{2.91}$$

and

$$\langle x_{k+1}|\hat{V}(\hat{x})|x_k\rangle = V(x_k) \, \delta(x_{k+1} - x_k) \tag{2.92}$$

We choose to write the delta function as a Fourier integral,

$$\delta(x_{k+1} - x_k) = \int \frac{dp_k}{2\pi\hbar} \, \exp\left[ip_k(x_{k+1} - x_k)/\hbar\right] \tag{2.93}$$

so that, collecting all the terms,

$$\langle x_{k+1}|(1 - i\hat{H}\epsilon/\hbar)|x_k\rangle = \int \frac{dp_k}{2\pi\hbar} \, \left[1 - iH(p_k, x_k)\epsilon/\hbar\right] \exp\left[ip_k(x_{k+1} - x_k)\right] \tag{2.94}$$

Notice that the *operator* $\hat{H}(\hat{p}, \hat{x})$ on the left-hand side has been replaced on the right-hand side by the *function* $H(p_k, x_k)$. Furthermore, since we shall take the limit $\epsilon \to 0$, we can equally well write $[1 - iH\epsilon/\hbar]$ as $\exp[-iH\epsilon/\hbar]$. Putting everything together, we have the following non-operator expression for the propagator:

$$K(x_T, x_0; T) = \mathrm{Lim} \int \left(\prod_{k=1}^{N-1} dx_k \right) \prod_{k=0}^{N-1} \frac{dp_k}{2\pi\hbar} \exp\left\{ \frac{i\epsilon}{\hbar} \left[p_k(x_{k+1} - x_k)/\epsilon - H(p_k, x_k) \right] \right\} \tag{2.95}$$

Taking the limit, we finally obtain

$$K(x_T, x_0; T) = \int \mathcal{D}p \, \mathcal{D}x \exp\left\{ \frac{i}{\hbar} \int_0^T dt \, \left[p\dot{x} - H(p, x) \right] \right\} \tag{2.96}$$

where the functional integration here is defined by

$$\int \mathcal{D}p \, \mathcal{D}x = \mathrm{Lim} \int \left(\prod_{k=1}^{N-1} dx_k \right) \prod_{k=0}^{N-1} \int \frac{dp_k}{2\pi\hbar} \tag{2.97}$$

Noting that the Hamiltonian is defined as

$$H(p, x) = p\dot{x} - L(x, \dot{x}) \tag{2.98}$$

where $p = \partial L/\partial\dot{x}$, we see that eq. (2.96) is now very close to the required form (2.85). The difference lies in the integration over all intermediate positions and momenta rather than just the

positions. However, the integration over momenta can be performed explicitly, since the exponent is a quadratic function of the relevant variables:

$$\begin{aligned}\int dp_k \, \exp\left\{\frac{i\epsilon}{\hbar}\left[p_k(x_{k+1}-x_k)/\epsilon - \frac{p_k^2}{2m}\right]\right\} &= \int dp'_k \, \exp\left\{\frac{i\epsilon m}{2\hbar}(x_{k+1}-x_k)^2/\epsilon^2 - \frac{i\epsilon}{2m\hbar}(p'_k)^2\right\} \\ &= \sqrt{\frac{2\pi m\hbar}{i\epsilon}} \exp\left\{\frac{i\epsilon m}{2\hbar}(x_{k+1}-x_k)^2/\epsilon^2\right\} \qquad (2.99)\end{aligned}$$

where we have used the change of variable $p'_k = p_k - m(x_{k+1}-x_k)/\epsilon$ to perform a Gaussian integration. Consequently an expression equivalent to (2.95) is

$$K(x_T, x_0; T) = \mathrm{Lim}\left(\frac{-im}{2\pi\hbar\epsilon}\right)^{\frac{N}{2}} \int \left(\prod_{k=1}^{N-1} dx_k\right) \prod_{k=0}^{N-1} \exp\left\{\frac{i\epsilon}{\hbar}\left[\frac{1}{2}m(x_{k+1}-x_k)^2/\epsilon^2 - V(x_k)\right]\right\} \qquad (2.100)$$

Taking the limit, we now obtain

$$K(x_T, x_0; T) = \int \mathcal{D}x(t) \exp\left\{\frac{i}{\hbar}\int_0^T \left[\frac{1}{2}m\dot{x}^2 - V(x)\right] dt\right\} \qquad (2.101)$$

which is exactly eq. (2.85), with the functional integration defined as in eq. (2.86) with

$$C(\epsilon) = \sqrt{\frac{-im}{2\pi\hbar\epsilon}} \qquad (2.102)$$

In summary, the path integral formulation allows us to proceed directly from classical to quantum mechanics, without having to introduce the machinery of vector spaces and operators. It is easily extended to any system that can be represented by a classical Lagrangian $L(\{q_i\}, \{\dot{q}_i\})$, where $\{q_i\}$ are generalized coordinates, no matter how many. This makes it particularly convenient for quantum field theory, where the field strength at every point in space-time has to be treated as a separate generalized coordinate. The price to be paid for the compact formalism is the introduction of functional integration, which in practice can only be carried out in closed form for Gaussian integrands like those encountered above.

Problem 9: Show that the exponential factor in the propagator for the simple harmonic oscillator, eq. (2.72), is just $\exp(iS_{cl}/\hbar)$ where S_{cl} is the action of the classical trajectory. (The non-exponential factor, which depends only on the elapsed time and not on the initial or final position, is the contribution of deviations from the classical path, i.e. *quantum fluctuations*.)

Chapter 3

Approximate Methods

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

We assume here that perturbative methods are familiar from the Advanced Quantum Physics course, and so we concentrate on the other two approaches. Variational methods should also be somewhat familiar, but we develop them further here. The JWKB method will be completely new. At the end of the chapter we compare and contrast these methods with perturbation theory, using as an example the anharmonic oscillator.

3.2 Variational methods

Variational methods (as applied 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 $\psi_{trial}(\mathbf{r}, \{\alpha, \beta, \ldots\})$)

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

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 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

$$\langle\hat{H}\rangle = \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} \tag{3.12}$$

Note that this is a "functional" of the state vector $|\Psi\rangle$. Consider an infinitesimally small change to $|\Psi\rangle$:

$$|\Psi\rangle \rightarrow |\Psi\rangle + |\delta\Psi\rangle \tag{3.13}$$

We need to determine the corresponding change to $\langle\hat{H}\rangle$. From (3.12) we have

$$\langle\Psi|\Psi\rangle\langle\hat{H}\rangle = \langle\Psi|\hat{H}|\Psi\rangle \tag{3.14}$$

Inserting (3.13) into the above equation we have:

$$(\langle\Psi| + \langle\delta\Psi|)(|\Psi\rangle + |\delta\Psi\rangle)[\langle\hat{H}\rangle + \delta\langle\hat{H}\rangle] = (\langle\Psi| + \langle\delta\Psi|)\hat{H}(|\Psi\rangle + |\delta\Psi\rangle) \tag{3.15}$$

which when expanded out gives:

$$(\langle\Psi|\delta\Psi\rangle + \langle\Psi|\Psi\rangle + \langle\delta\Psi|\Psi\rangle + \langle\delta\Psi|\delta\Psi\rangle)[\langle\hat{H}\rangle + \delta\langle\hat{H}\rangle]$$

$$= \langle\Psi|\hat{H}|\delta\Psi\rangle + \langle\Psi|\hat{H}|\Psi\rangle + \langle\delta\Psi|\hat{H}|\delta\Psi\rangle + \langle\delta\Psi|\hat{H}|\Psi\rangle \tag{3.16}$$

Using (3.13), dropping all terms of second order, after some algebra we finally get:

$$\delta\langle\hat{H}\rangle\langle\Psi|\Psi\rangle = \langle\Psi|(\hat{H} - \langle\hat{H}\rangle)|\delta\Psi\rangle + \langle\delta\Psi|(\hat{H} - \langle\hat{H}\rangle)|\Psi\rangle \tag{3.17}$$

Thus $\langle\hat{H}\rangle$ is stationary, i.e. $\delta\langle\hat{H}\rangle = 0$, provided the right hand side of (3.17) is zero:

$$\langle\Psi|(\hat{H} - \langle\hat{H}\rangle)|\delta\Psi\rangle + \langle\delta\Psi|(\hat{H} - \langle\hat{H}\rangle)|\Psi\rangle = 0 \tag{3.18}$$

Suppose we define $|\Phi\rangle = (\hat{H} - \langle\hat{H}\rangle)|\Psi\rangle$. Then (3.18) becomes:

$$\langle\Phi|\delta\Psi\rangle + \langle\delta\Psi|\Phi\rangle = 0 \tag{3.19}$$

This must be satisfied for **any** $|\delta\Psi\rangle$. Let us choose a $|\delta\Psi\rangle$ such that:

$$|\delta\Psi\rangle = \delta\lambda|\Phi\rangle \tag{3.20}$$

where $\delta\lambda$ is a number of first order in small quantities. This implies that (3.19) becomes:

$$2\delta\lambda\langle\Phi|\Phi\rangle = 0 \tag{3.21}$$

i.e. that the *norm* of $|\Phi\rangle$ equals zero. Therefore, $|\Phi\rangle$ must be zero, i.e.

$$(\hat{H} - \langle\hat{H}\rangle)|\Psi\rangle = 0 \tag{3.22}$$

This implies that:

$$\hat{H}|\Psi\rangle = \langle\hat{H}\rangle|\Psi\rangle \tag{3.23}$$

Therefore the right hand side of (3.12) 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 $\hat{H}$.

3.2.3 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:

$$|\Psi\rangle = \sum_{i=1}^{N} c_i \, |i\rangle \tag{3.24}$$

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

$$\langle\Psi|\hat{H}|\Psi\rangle - E\langle\Psi|\Psi\rangle = 0 \tag{3.25}$$

(We set $E = \langle\hat{H}\rangle$). Substituting (3.24) into (3.25) yields:

$$\sum_{i=1}^{N}\sum_{j=1}^{N} c_j^* \, c_i \, H_{ji} - E\sum_{i=1}^{N}\sum_{j=1}^{N} c_j^* \, c_i \, S_{ji} = 0 \tag{3.26}$$

where

$$\begin{aligned} H_{ji} &= \langle j|\hat{H}|i\rangle \\ S_{ji} &= \langle j|i\rangle \end{aligned} \tag{3.27}$$

Differentiating (3.26) with respect to c_l^* gives:

$$\sum_{i=1}^{N}(H_{li} - E \, S_{li}) \, c_i = 0 \tag{3.28}$$

Cramer's rule tells us that all the c_i's are zero unless the determinant of the coefficients vanishes:

$$|\mathbf{H} - E\,\mathbf{S}| = 0 \tag{3.29}$$

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 (3.29) 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

$$E_0 \leq \tilde{E}_0 \tag{3.30}$$

Moreover, it can be proved[1] that

$$E_1 \leq \tilde{E}_1, \quad E_2 \leq \tilde{E}_2, \ \ldots, \ E_{N-1} \leq \tilde{E}_{N-1} \tag{3.31}$$

Thus the linear variation method provides upper bounds to the energies of the lowest N eigenstates of the system. The roots of Eq. (3.29) are used as approximations to the energies of the lowest eigenstates. Increasing the value of N in (3.24) (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 (3.28); 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. (3.28) 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

$$|\Psi\rangle = \sum_{i=1}^{N} c_i^0 \ |i\rangle$$

as an approximate ground state wave function. Use of the higher roots of Eq. (3.24) in Eq. (3.28) 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 6†: Let $V(x) = 0$ for $-1 \leq x \leq +1$ and ∞ otherwise (the "particle in a box" problem). Use

$$f_1(x) \ = \ (1 - x^2)$$

$$f_2(x) \ = \ (1 - x^4)$$

[1] J.K. MacDonald, Phys. Rev. **43**, p.830 (1933); R.H. Young, Int. J. Quantum Chem. **6**, p.596 (1972).

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 (**S**) and Hamiltonian matrix (**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 7: 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

$$\Psi(x) = \sum_{i=1}^{4} c_i \, f_i(x)$$

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 $\approx$ 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 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

$$\psi(x) = A\, e^{\pm ikx}$$

$$k = \frac{\sqrt{2m(E-V)}}{\hbar}$$

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.

3.3.1 Derivation

We seek to solve

$$\begin{aligned} \frac{d^2\psi}{dx^2} + k^2(x)\psi(x) = 0 \\ k^2(x) = \frac{2m}{\hbar^2}(E - V(x)) \end{aligned} \tag{3.32}$$

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,

$$\begin{aligned} \frac{d\psi}{dx} &= iS'e^{iS} \\ \frac{d^2\psi}{dx^2} &= (iS'' - S'^2)e^{iS} \end{aligned} \tag{3.33}$$

and the Schrödinger equation reduces to $(iS'' - S'^2 + k^2)e^{iS} = 0$, or

$$\begin{aligned} S' &= \pm\sqrt{k^2(x) + iS''(x)} \\ &= \pm k(x)\sqrt{1 + iS''(x)/k^2} \end{aligned} \tag{3.34}$$

(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:

$$\begin{aligned} S' &= \pm k\sqrt{1 \pm ik'(x)/k^2} \\ &\approx \pm k\left(1 \pm \frac{i}{2}\frac{k'(x)}{k^2}\right) \\ &\approx \pm k + \frac{i}{2}\frac{k'(x)}{k} \end{aligned} \tag{3.35}$$

where we have assumed that the corrections are small. Then, we have

$$\begin{aligned} \frac{dS}{dx} &= \pm k + \frac{i}{2}\frac{k'}{k} \\ S(x) &\sim \pm\int^x k(x)dx + \frac{i}{2}\int^x \frac{k'}{k}dx + c \end{aligned} \tag{3.36}$$

The second integral is a perfect differential ($d \ln k$), so

$$\begin{aligned} S(x) &= \pm\int^x k(x)dx + \frac{i}{2}\ln k + c \\ \psi &= e^{iS} \\ &= C\, e^{\pm i\int^x k(x)dx} e^{-\frac{1}{2}\ln k} \\ &= \frac{C}{\sqrt{k(x)}} e^{\pm i\int^x k(x)dx} \end{aligned} \tag{3.37}$$

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

$$\begin{aligned} \psi(x) &= \frac{C}{\sqrt{K(x)}} e^{\pm\int^x K(x)dx} \\ K^2 &= \frac{2m}{\hbar^2}(V - E) > 0 \end{aligned} \tag{3.38}$$

3.3.2 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

$$k^2(x) = \frac{2m}{\hbar^2}\left(-\frac{\partial V}{\partial x}\right)_{x=a}(x-a) \tag{3.39}$$

(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

$$\psi'' - \left(\frac{\partial V}{\partial x}\right)_{x=a}\frac{2m}{\hbar^2}\,(x-a)\psi = 0 \tag{3.40}$$

This is a linear potential problem which is solved in terms of **Airy** functions. If we let

$$\begin{aligned} y &= \alpha(a-x) \geq 0 \\ \alpha^3 &= \frac{2m}{\hbar^2}\,\frac{\partial V}{\partial x} \geq 0 \end{aligned} \tag{3.41}$$

then the above differential equation becomes

$$\psi''(y) + y\,\psi(y) = 0 \tag{3.42}$$

whose solutions are:

$$\begin{aligned} \psi &= A\sqrt{y}J_{-\frac{1}{3}}(z) + B\sqrt{y}J_{+\frac{1}{3}}(z) \\ z &= \frac{2}{3}\,y^{\frac{3}{2}} \\ &= \int_x^a k(x)\,dx \end{aligned} \tag{3.43}$$

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

$$\begin{aligned} J_\nu &\to \sqrt{\frac{2}{\pi z}}\,\cos\left(z - \frac{\nu\pi}{2} - \frac{\pi}{4}\right) \\ \psi &\to \sqrt{\frac{2}{\pi z}}\left[A\sqrt{y}\,\cos\left(z + \frac{\pi}{6} - \frac{\pi}{4}\right) + B\sqrt{y}\,\cos\left(z - \frac{\pi}{6} - \frac{\pi}{4}\right)\right] \end{aligned} \tag{3.44}$$

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

$$\begin{aligned} J_\nu &\sim \frac{\left(\frac{z}{2}\right)^\nu}{\Gamma(\nu+1)} \\ \psi &\to \frac{A\sqrt{y}(\frac{1}{2}z)^{-1/3}}{\Gamma(\frac{2}{3})} + \frac{B\sqrt{y}(\frac{1}{2}z)^{1/3}}{\Gamma(\frac{4}{3})} \end{aligned} \tag{3.45}$$

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

$$\begin{aligned} \psi_{left}(y \to \infty; x \ll a) &\to \sqrt{\frac{3}{\pi}}\frac{\alpha^{1/4}}{\sqrt{k(x)}}\Big[A\cos\Big(\int_x^a k(x)\,dx - \frac{\pi}{12}\Big) \\ &+ B\cos\Big(\int_x^a k(x)\,dx - \frac{5\pi}{12}\Big)\Big] \\ \psi_{left}(\text{near } x = a) &\to \frac{A\,3^{1/3}}{\Gamma(\frac{2}{3})} + \frac{By}{\Gamma(\frac{4}{3})\,3^{1/3}} \end{aligned} \tag{3.46}$$

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:

$$\begin{aligned} \psi &= C\,\sqrt{y}\,I_{\frac{1}{3}}(z) + D\,\sqrt{y}\,I_{-\frac{1}{3}}(z) \\ z &= \int_a^x K(x)\,dx \\ K^2(x) &= \frac{2m}{\hbar^2}\Big(V(x) - E\Big) \end{aligned} \tag{3.47}$$

The $I's$ are Bessel functions of imaginary argument.

When the same steps described in detail above are followed, we find

$$\psi_{right}(x \gg a) \to \sqrt{\frac{3}{4\pi}}\,\frac{C\,\alpha^{1/4}}{\sqrt{K(x)}}\Big[e^{\int_a^x K(x)\,dx} - \frac{\sqrt{3}}{2}e^{-\int_a^x K(x)\,dx}\Big]$$

$$+ \quad \sqrt{\frac{3}{4\pi}}\,\frac{D\,\alpha^{1/4}}{\sqrt{K(x)}}\left[e^{\int_a^x K(x)\,dx} + \frac{\sqrt{3}}{2}e^{-\int_a^x K(x)\,dx}\right]$$

$$\psi_{right}(\text{near } x=a) \quad \rightarrow \quad \frac{D\,3^{1/3}}{\Gamma(\frac{2}{3})} + \frac{Cy}{\Gamma(\frac{4}{3})\,3^{1/3}} \tag{3.48}$$

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:

$$\begin{aligned}
\frac{2}{\sqrt{k(x)}}\,\cos\left[\int_x^a k(x)\,dx - \frac{\pi}{4}\right] \quad &\Longleftrightarrow \quad \frac{1}{\sqrt{K(x)}}\,e^{-\int_a^x K(x)\,dx} \\
\frac{1}{\sqrt{k(x)}}\,\sin\left[\int_x^a k(x)\,dx - \frac{\pi}{4}\right] \quad &\Longleftrightarrow \quad -\frac{1}{\sqrt{K(x)}}\,e^{\int_a^x K(x)\,dx}
\end{aligned} \tag{3.49}$$

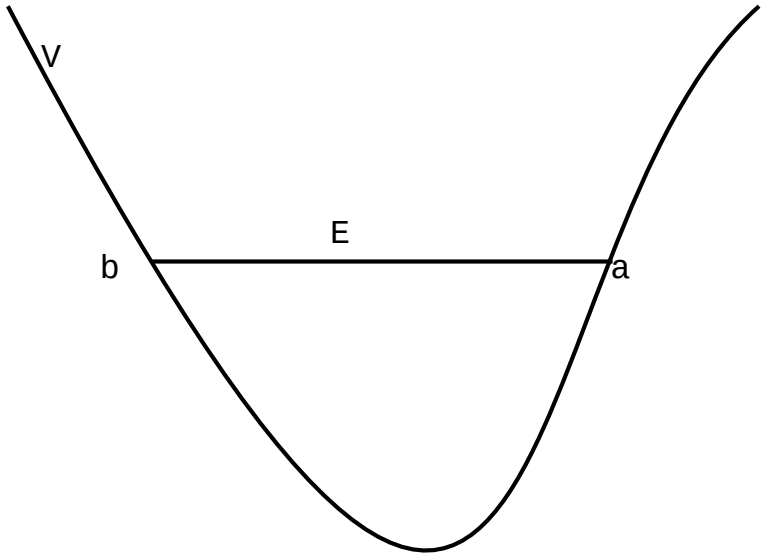
This allow a continuation of solutions. A sine solution on the left matches into a $-e^{+\int_a^x K(x)\,dx}$ solution on the right. Similar formulae for the reverse situation with the turning point on the left ($x < b$ classically forbidden) give

$$\begin{aligned}
\frac{1}{\sqrt{K(x)}}\,e^{-\int_x^b K(x)\,dx} \quad &\Longleftrightarrow \quad \frac{2}{\sqrt{k(x)}}\,\cos\left[\int_b^x k(x)\,dx - \frac{\pi}{4}\right] \\
\frac{1}{\sqrt{K(x)}}\,e^{\int_x^b K(x)\,dx} \quad &\Longleftrightarrow \quad -\frac{1}{\sqrt{k(x)}}\,\sin\left[\int_b^x k(x)\,dx - \frac{\pi}{4}\right]
\end{aligned} \tag{3.50}$$

Problem 8: Verify Eq. (3.49).

3.3.3 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 $x \to -\infty$):

$$\psi_{left} = \frac{\text{const}}{\sqrt{K(x)}}\, e^{\int_b^x K(x)\, dx} \to 0 \tag{3.51}$$

Inside the potential well, this matches to:

$$\psi_{inside} = \frac{2 \cdot \text{const}}{\sqrt{k(x)}}\, \cos\left(\int_b^x k(x)\, dx - \frac{\pi}{4} \right) \tag{3.52}$$

We now move along to the next turning point $x = a$. Rewriting ψ inside (for easy matching) as:

$$\psi_{inside} = \frac{2 \cdot \text{const}}{\sqrt{k(x)}}\, \cos\left(\int_b^a k(x)\, dx - \int_x^a k(x)\, dx - \frac{\pi}{4} \right) \tag{3.53}$$

We define

$$\phi = \int_b^a k(x)\, dx \tag{3.54}$$

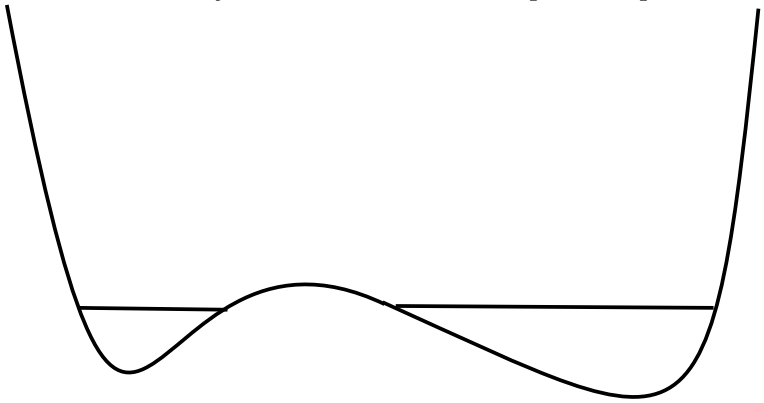
and applying the connection formulae developed in the previous section, we get

$$\begin{aligned} \psi_{inside} &= \frac{2 \cdot \text{const}}{\sqrt{k(x)}}\, \cos\left[-\int_x^a k(x)\, dx + \phi - \frac{\pi}{2} + \frac{\pi}{4} \right] \\ &= \frac{2 \cdot \text{const}}{\sqrt{k(x)}}\, \left[\cos\left(\phi - \frac{\pi}{2} \right)\, \cos\left(\int_x^a k(x)\, dx - \frac{\pi}{4} \right) \right. \\ &- \left. \sin\left(\frac{\pi}{2} - \phi \right)\, \sin\left(\int_x^a k(x)\, dx - \frac{\pi}{4} \right) \right] \end{aligned} \tag{3.55}$$

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

$$\int_b^a k(x)\, dx = (n + \frac{1}{2})\pi \tag{3.56}$$

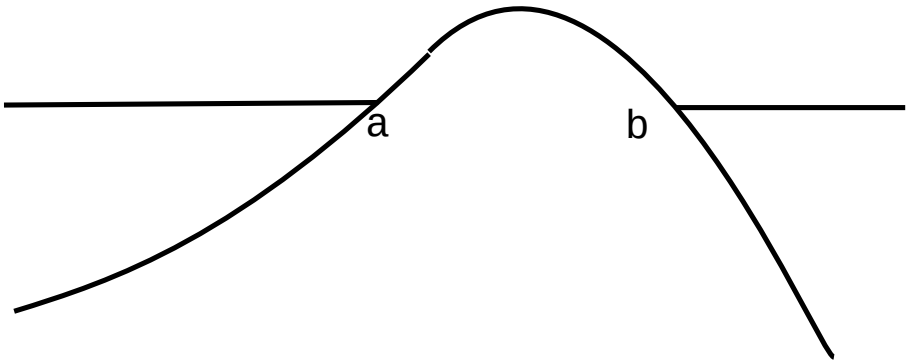
This is similar to the **Bohr-Sommerfeld Quantization Rule**, but with n shifted by $\frac{1}{2}$. This matching may need to be done many times if we have a complicated potential such as:



Problem 9: Show that Eq. (3.56) gives the exact energy levels of a simple harmonic oscillator.

3.3.4 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

$$\psi_{left} = \frac{A}{\sqrt{k}}\, e^{i\int_a^x k(x)\,dx} + \frac{B}{\sqrt{k}}\, e^{-i\int_a^x k(x)\,dx}$$

$$\psi_{middle} = \frac{C}{\sqrt{K}}\, e^{-\int_a^x K(x)\,dx} + \frac{D}{\sqrt{K}}\, e^{\int_a^x K(x)\,dx}$$

$$\psi_{right} = \frac{F}{\sqrt{k}}\, e^{i\int_b^x k(x)\,dx} + \frac{G}{\sqrt{k}}\, e^{-i\int_b^x k(x)\,dx} \tag{3.57}$$

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:

$$\begin{aligned}\psi_{left} &= (A\, e^{-i\frac{\pi}{4}} + B\, e^{i\frac{\pi}{4}})\, \frac{1}{\sqrt{k}}\, \cos\left(\int_x^a k(x)\,dx - \frac{\pi}{4}\right) \\ &+ i\,(-A\, e^{-i\frac{\pi}{4}} + B\, e^{i\frac{\pi}{4}})\, \frac{1}{\sqrt{k}}\, \sin\left(\int_x^a k(x)\,dx - \frac{\pi}{4}\right)\end{aligned} \tag{3.58}$$

On applying the connection formula, we find:

$$\begin{aligned}A\, e^{-i\frac{\pi}{4}} + B\, e^{i\frac{\pi}{4}} &= 2\,C \\ -i\,B\, e^{i\frac{\pi}{4}} + i\,A\, e^{-i\frac{\pi}{4}} &= D\end{aligned} \tag{3.59}$$

We do the same at the next turning point ($x = b$). Rewrite ψ_{middle} as

$$\begin{aligned}\psi_{middle} &= \frac{C}{\sqrt{K}}\, \Theta^{-1}\, e^{\int_x^b K(x)\,dx} + \frac{D}{\sqrt{K}}\, \Theta\, e^{-\int_x^b K(x)\,dx} \\ \Theta &= e^{\int_a^b K(x)\,dx}\end{aligned} \tag{3.60}$$

which on matching yields:

$$\begin{aligned}F\, e^{i\frac{\pi}{4}} + G\, e^{-i\frac{\pi}{4}} &= 2\,D\,\Theta \\ i\,(F\, e^{i\frac{\pi}{4}} - G\, e^{-i\frac{\pi}{4}}) &= -\frac{C}{\Theta}\end{aligned} \tag{3.61}$$

These relations can be written in a simple matrix notation as:

$$\begin{aligned}\begin{pmatrix} C \\ D \end{pmatrix} &= \frac{1}{2}\begin{pmatrix} e^{-i\frac{\pi}{4}} & e^{i\frac{\pi}{4}} \\ i\,2\,e^{-i\frac{\pi}{4}} & -i\,2\,e^{i\frac{\pi}{4}} \end{pmatrix}\begin{pmatrix} A \\ B \end{pmatrix} \\ \begin{pmatrix} C \\ D \end{pmatrix} &= \begin{pmatrix} -i\,\Theta\, e^{i\frac{\pi}{4}} & i\,\Theta\, e^{-i\frac{\pi}{4}} \\ \frac{1}{2\Theta}\, e^{i\frac{\pi}{4}} & \frac{1}{2\Theta}\, e^{-i\frac{\pi}{4}} \end{pmatrix}\begin{pmatrix} F \\ G \end{pmatrix}\end{aligned} \tag{3.62}$$

After some matrix algebra we get

$$\begin{aligned}\begin{pmatrix} A \\ B \end{pmatrix} &= \mathbf{T}\begin{pmatrix} F \\ G \end{pmatrix} \\ \mathbf{T} &= \frac{1}{2}\begin{pmatrix} 2\,\Theta + \frac{1}{2\Theta} & i\,(2\,\Theta - \frac{1}{2\Theta}) \\ -i\,(2\,\Theta - \frac{1}{2\Theta}) & 2\,\Theta + \frac{1}{2\Theta} \end{pmatrix}\end{aligned} \tag{3.63}$$

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

$$\begin{aligned} \left|\frac{F}{A}\right|^2 &= \left|\frac{1}{\Theta + \frac{1}{4\Theta}}\right|^2 \\ &\sim \Theta^{-2} \text{ for large } \Theta. \end{aligned} \tag{3.64}$$

Therefore for large Θ the transmission coefficient will be

$$T \equiv \left|\frac{F}{A}\right|^2 \sim \Theta^{-2} = e^{-2\int_a^b K(x)\,dx} \tag{3.65}$$

Problem 10: Verify Eq. (3.64).

Problem 11: Consider a particle with total energy $E = 0$ moving in the *anharmonic oscillator* potential

$$V(x) = \frac{1}{2}m\omega^2(x^2 - \epsilon x^4)$$

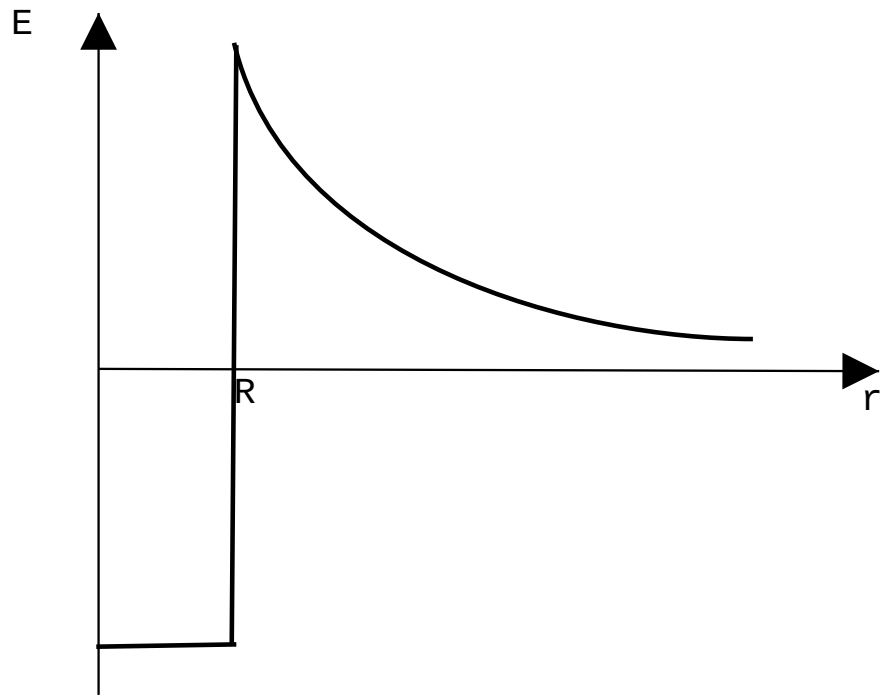
Show that, for small ϵ, the transmission coefficient of the potential barrier between $x = 0$ and $x = 1/\sqrt{\epsilon}$ is

$$T \sim e^{-2m\omega/3\hbar\epsilon}\ .$$

(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.)

3.3.5 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:

$$V(r) = \frac{2Ze^2}{r} \quad r > R \tag{3.66}$$

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

$$\begin{aligned} I &= \frac{1}{\sqrt{E}} \int_R^{r_0} \left(\frac{2Ze^2}{r} - E \right)^{\frac{1}{2}} dr \\ &= \int_R^{\gamma R} \left(\frac{\gamma R}{r} - 1 \right)^{\frac{1}{2}} dr \\ \gamma &= \frac{2Ze^2}{ER} \end{aligned} \tag{3.67}$$

Substitute

$$\begin{aligned} \frac{r}{\gamma R} &= \sin^2 \theta \\ dr &= 2\gamma R \sin\theta \cos\theta \, d\theta \\ \theta_0 &= \sin^{-1} \frac{1}{\sqrt{\gamma}} \end{aligned} \tag{3.68}$$

into the above integral to get

$$\begin{aligned} I &= 2\gamma R \int_{\theta_0}^{\frac{\pi}{2}} \cos^2\theta \, d\theta \\ &= \gamma R\Big[\frac{\pi}{2} - \sin^{-1}\frac{1}{\sqrt{\gamma}} - \frac{(\gamma-1)^2}{\gamma}\Big] \end{aligned} \tag{3.69}$$

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

$$\ln T = -2\frac{\sqrt{2m}}{\hbar}\, I\, \sqrt{E} \tag{3.70}$$

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

$$\begin{aligned} T &= A\, e^{-\beta} \\ \beta &= \pi\frac{\sqrt{2m}}{\hbar}\Big(\frac{2Ze^2}{\sqrt{E}}\Big) \\ A &= e^{-\frac{4}{\hbar}\sqrt{Ze^2mR}} \end{aligned} \tag{3.71}$$

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.

$$\Gamma = \omega\, A\, e^{-\beta} \tag{3.72}$$

We estimate ω as follows:

$$\begin{aligned} \omega &\simeq \frac{v}{R} \\ &\cong \frac{\Delta p}{mR} \\ &\cong \frac{\hbar}{2mR^2} \end{aligned} \tag{3.73}$$

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

$$\Gamma \sim 0.6\, e^{-182/\sqrt{E}} \times 10^{40}\ \text{sec}^{-1} \quad (E \text{ in MeV}) \tag{3.74}$$

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!

3.4 Example: the anharmonic oscillator

A good place to study the strengths and weaknesses of the various approximate methods is the *anharmonic oscillator*, by which we mean a simple harmonic oscillator with an additional quartic term in the potential:

$$E\psi = -\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \frac{1}{2}m\omega^2 x^2\psi + \lambda x^4\psi \tag{3.75}$$

(You will see shortly why we do not allow a cubic term.) It will be convenient to recast the Schrödinger equation in dimensionless form by defining

$$y = \sqrt{\frac{2m\omega}{\hbar}}\,x\ , \quad \mathcal{E} = \frac{E}{\hbar\omega}\ , \quad g = \frac{\lambda\hbar}{m^2\omega^3} \tag{3.76}$$

so that

$$\mathcal{E}\psi = -\frac{d^2\psi}{dy^2} + \frac{1}{4}y^2(1+gy^2)\psi \tag{3.77}$$

3.4.1 Perturbation theory

The obvious approximation scheme to try first is perturbation theory. Recall that if $|n\rangle$ denotes the nth excited state of the SHO, the perturbation series for the energy of that state will be

$$\mathcal{E}_n = n + \frac{1}{2} + \frac{1}{4}g\langle n|y^4|n\rangle + \frac{1}{16}g^2\sum_{n'\neq n}\frac{\langle n|y^4|n'\rangle\langle n'|y^4|n\rangle}{n-n'} + \cdots \tag{3.78}$$

The simplest way to evaluate the matrix elements is by means of the *ladder operators* for the SHO, which in our rescaled notation are

$$\hat{a} = \frac{1}{2}y + \frac{d}{dy}\ , \quad \hat{a}^\dagger = \frac{1}{2}y - \frac{d}{dy} \tag{3.79}$$

so that $[\hat{a},\hat{a}^\dagger] = 1$ and the unperturbed ($g = 0$) Schrödinger equation becomes

$$\mathcal{E}_n|n\rangle = \left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)|n\rangle = \left(\hat{a}\hat{a}^\dagger - \frac{1}{2}\right)|n\rangle \tag{3.80}$$

The properties of the ladder operators should be familiar to you. They change the level of excitation of the SHO by one unit:

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle\ , \quad \hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle\ . \tag{3.81}$$

Therefore, starting from the ground state $|0\rangle$ and using $y = \hat{a} + \hat{a}^\dagger$, we find

$$y|0\rangle = |1\rangle\ , \quad y^2|0\rangle = |0\rangle + \sqrt{2}|2\rangle\ , \quad y^3|0\rangle = 3|1\rangle + \sqrt{6}|3\rangle\ , \quad y^4|0\rangle = 3|0\rangle + 6\sqrt{2}|2\rangle + 2\sqrt{6}|4\rangle \tag{3.82}$$

Thus

$$\mathcal{E}_0 \equiv \mathcal{E}(g) = \frac{1}{2} + \frac{3}{4}g - \frac{1}{16}g^2\left(\frac{72}{2} + \frac{24}{4} = 42\right) + \cdots \tag{3.83}$$

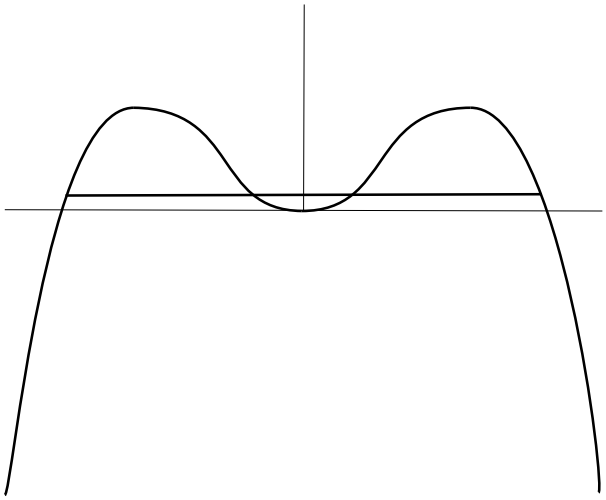
Including the next term, one finds

$$\mathcal{E}(g) = \frac{1}{2} + \frac{3}{4}g - \frac{21}{8}g^2 + \frac{333}{16}g^3 - \cdots \tag{3.84}$$

Notice that the coefficients in this series seem to be getting rather large; the fourth term is already equal to the first when $g = 0.024$. So we start to worry about the radius of convergence – and in fact it is *zero*! The series does not converge for *any* finite value of g.

3.4.2 JWKB method

The reason for the non-convergence of the perturbation series becomes obvious when we consider $\mathcal{E}(g)$ for *negative* values of g. In that case the potential is as shown below:



No matter how small the value of g, eventually the negative x^4 term wins at large $|x|$, and the potential goes down to $-\infty$. This means that the system has *no stable bound states* for $g < 0$.

What will happen if $g < 0$ and we put a particle in the potential well at small x is that it will eventually tunnel through the potential barrier and escape to $x = \pm\infty$. In fact, we already computed the transmission coefficient for this potential barrier using the JWKB approximation in Problem 11. In the notation used there, $\epsilon = -2m\omega g/\hbar$ and therefore

$$T \sim \exp\left(\frac{1}{3g}\right) \qquad (g < 0) \tag{3.85}$$

It follows that the probability density $|\Psi|^2$, for any of the states that would be bound states when $g = 0$, acquires a time dependence when $g < 0$, of the form

$$|\Psi(x,t)|^2 = |\Psi(x,0)|^2 e^{-\Gamma t} \tag{3.86}$$

where Γ is the decay rate due to tunnelling. As in the case of alpha decay, we estimate

$$\Gamma \sim \omega T \sim \omega \exp\left(\frac{1}{3g}\right) \tag{3.87}$$

A more careful evaluation[2] gives for the ground state

$$\Gamma \approx \omega\sqrt{\frac{-2}{\pi g}} \exp\left(\frac{1}{3g}\right) \tag{3.88}$$

Correspondingly, the energy of the state acquires an imaginary part $-i\hbar\Gamma/2$:

$$\Psi(x,t) = \psi(x)e^{-iEt/\hbar} \to \psi(x)e^{-iEt/\hbar - \Gamma t/2} = \psi(x)e^{-i(E - i\hbar\Gamma/2)t/\hbar} \tag{3.89}$$

and hence for the ground state

$$\mathrm{Im}\,\mathcal{E}(g) \approx \frac{-1}{\sqrt{-2\pi g}} \exp\left(\frac{1}{3g}\right) \qquad (g < 0) \tag{3.90}$$

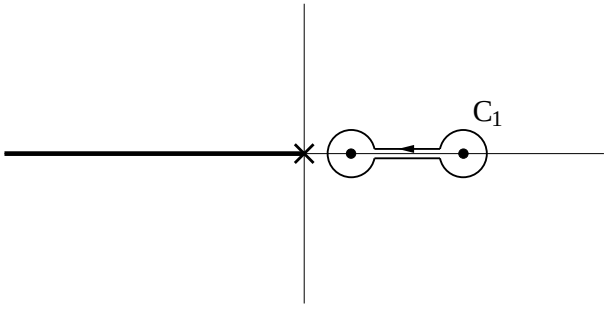
Thus $\mathcal{E}(g)$, considered as a function of the complex variable g, has a *branch cut* running from the origin along the negative real axis. The existence of this branch cut means that there is no region around the origin within which $\mathcal{E}(g)$ is an analytic function of g, and hence it has no convergent Taylor expansion about the origin, i.e. *no convergent perturbation series.*

3.4.3 Dispersion theory

We can see precisely how the perturbation series diverges, and find a better way to evaluate $\mathcal{E}(g)$, by using *dispersion theory.* This is a way of reconstructing an analytic function from its imaginary part, given that the function has no singularities other than branch cuts on the real axis. In the case at hand, consider the function

$$f(g) = \frac{\mathcal{E}(g)}{(g - g_0)(g - g_1)} \tag{3.91}$$

where $g_{0,1}$ are real and $0 < g_0 < g_1$. This function has poles at $g = g_0$ and g_1 in addition to the branch cut of $\mathcal{E}(g)$, as depicted below.

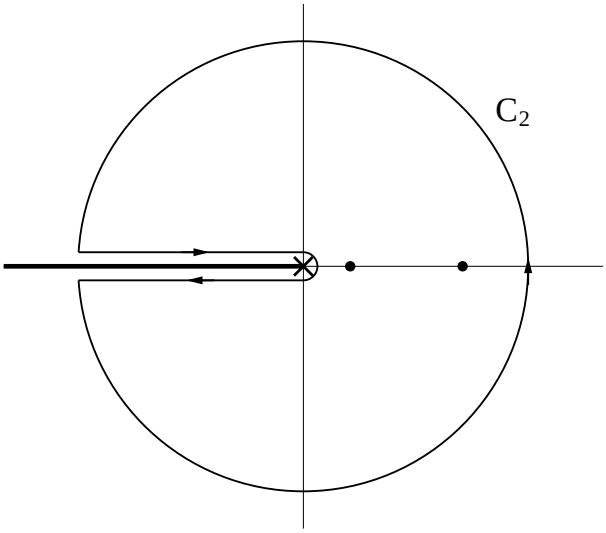


[2] C.M. Bender and T.T. Wu, Phys. Rev. D7 (1973) 1620.

Integrating around the contour C_1 enclosing the two poles, we have by the residue theorem

$$\int_{C_1} f(g)\,dg = 2\pi i\,\frac{\mathcal{E}(g_1)-\mathcal{E}(g_0)}{g_1-g_0} \tag{3.92}$$

Now let us open up the contour into the form C_2 shown below.



We have

$$\int_{C_2} f(g)\,dg = \int_{C_\infty} f(g)\,dg + \lim_{\epsilon\to 0}\int_{-\infty}^{0} dg\,\frac{\mathcal{E}(g+i\epsilon)-\mathcal{E}(g-i\epsilon)}{(g-g_0)(g-g_1)} \tag{3.93}$$

where C_∞ is the circular portion of the contour. To compute its contribution we must know how $\mathcal{E}(g)$ behaves at large values of g. This is easily done using a scaling argument. Let's make a change of variable $y = g^p z$, where p is to be determined. The Schrödinger equation becomes

$$\mathcal{E}(g)\psi = -g^{-2p}\frac{d^2\psi}{dz^2} + \frac{1}{4}g^{2p}z^2\psi + \frac{1}{4}g^{4p+1}z^4\psi \tag{3.94}$$

If we choose $-2p = 4p+1$, i.e. $p=-1/6$, the equation can be rearranged as

$$g^{-1/3}\mathcal{E}(g)\psi = -\frac{d^2\psi}{dz^2} + \frac{1}{4}g^{-2/3}z^2\psi + \frac{1}{4}z^4\psi \tag{3.95}$$

Therefore at large g we can drop the term involving z^2 and it follows that $\mathcal{E}(g)\sim Ag^{1/3}$ where A is an eigenvalue of the operator

$$\hat{A} = -\frac{d^2}{dz^2} + \frac{1}{4}z^4 \tag{3.96}$$

Since $\mathcal{E}(g)$ behaves like $g^{1/3}$ at large g, the contribution from C_∞ in eq. (3.93) is negligible. For if the circle has radius $|g|=R$ then the integrand behaves like $R^{-5/3}$ and the length of the contour is $2\pi R$, so its contribution vanishes like $R^{-2/3}$ as $R\to\infty$. Equating the integrals on C_1 and C_2 and rearranging terms, we have

$$\mathcal{E}(g_1) = \mathcal{E}(g_0) + \frac{g_1-g_0}{2\pi i}\lim_{\epsilon\to 0}\int_{-\infty}^{0} dg\,\frac{\mathcal{E}(g+i\epsilon)-\mathcal{E}(g-i\epsilon)}{(g-g_0)(g-g_1)} \tag{3.97}$$

Now from the fact that the Schrödinger equation has real coefficients we can deduce that

$$\mathcal{E}(g - i\epsilon) = \mathcal{E}^*(g + i\epsilon) \tag{3.98}$$

and the correct definition of $\mathcal{E}(g)$ at negative g is

$$\mathcal{E}(g) = \lim_{\epsilon \to 0} \mathcal{E}(g - i\epsilon) \qquad (g < 0) \tag{3.99}$$

Thus

$$\mathcal{E}(g_1) = \mathcal{E}(g_0) - \frac{g_1 - g_0}{\pi} \int_{-\infty}^{0} dg \frac{\operatorname{Im} \mathcal{E}(g)}{(g - g_0)(g - g_1)} \tag{3.100}$$

Finally, it is convenient to take the limit $g_0 \to 0^+$ and to relabel g as $-g'$ and g_1 as g, to obtain the *dispersion relation*

$$\mathcal{E}(g) = \frac{1}{2} - \frac{g}{\pi} \int_0^\infty dg' \, \frac{\operatorname{Im} \mathcal{E}(-g')}{g'(g' + g)} \tag{3.101}$$

Now we can see what goes wrong with the perturbation series. To get a power series in g from eq. (3.101) we must expand the factor of $(g' + g)^{-1}$ in the integrand,

$$(g' + g)^{-1} = \sum_{n=0}^{\infty} (-g)^n (g')^{-n-1} \tag{3.102}$$

to obtain

$$\mathcal{E}(g) = \frac{1}{2} + \sum_{n=1}^{\infty} a_n g^n \tag{3.103}$$

where

$$a_n = \frac{(-1)^n}{\pi} \int_0^\infty dg' \, (g')^{-n-1} \operatorname{Im} \mathcal{E}(-g') \tag{3.104}$$

However, this procedure does not make much sense, because the region of convergence of the series (3.102) is $|g| < g'$ and we are going to integrate over g' from zero to infinity. So there is always a part of the integration region, $g' < g$, that is outside the region of convergence of the series. That is the fundamental reason why the perturbation series for $\mathcal{E}(g)$ cannot converge.

Let us nevertheless press ahead with evaluating the coefficients a_n, which remain perfectly well defined even though the series (3.103) does not converge. At large values of n the integral (3.104) will be dominated by small values of g', where the approximation (3.90) is most valid. Therefore we expect at large n

$$\begin{aligned} a_n &\sim \frac{(-1)^{n+1}}{\sqrt{2\pi^3}} \int_0^\infty dg' \, (g')^{-n-\frac{3}{2}} \exp\left(-\frac{1}{3g'}\right) \\ &= -(-3)^n \sqrt{\frac{3}{2\pi^3}} \int_0^\infty du \, u^{n-\frac{1}{2}} e^{-u} \\ &= -(-3)^n \sqrt{\frac{3}{2\pi^3}} \, \Gamma\left(n + \frac{1}{2}\right) \end{aligned} \tag{3.105}$$

Recalling that $\Gamma(z)$ grows factorially for large positive z, we see that in high orders the terms of the perturbation series will grow like $n!\,(-3g)^n$, a truly horrible rate of divergence.

Even at small n the approximation (3.105) for the perturbative coefficients is not too bad: we have

$$\begin{aligned}\mathcal{E}(g) &\approx \frac{1}{2} + \frac{\sqrt{6}}{\pi}\left(\frac{3}{4}g - \frac{27}{8}g^2 + \frac{405}{16}g^3 - \cdots\right) \\ &\approx 0.500 + 0.585g - 2.631g^2 + 19.736g^3 - \cdots \end{aligned} \tag{3.106}$$

to be compared with the true series

$$\mathcal{E}(g) = 0.500 + 0.750g - 2.625g^2 + 20.812g^3 - \cdots \tag{3.107}$$

This suggests a better way of estimating $\mathcal{E}(g)$ at positive g. We know exactly what the sum of the divergent series (3.106) should be: the dispersion relation (3.101) with the approximation (3.90) substituted for $\operatorname{Im}\mathcal{E}$. Let us call this $\mathcal{E}_d(g)$:

$$\begin{aligned}\mathcal{E}_d(g) &= \frac{1}{2} + \frac{g}{\pi}\int_0^\infty \frac{dg'}{\sqrt{2\pi g'}}\frac{1}{g'(g'+g)}\exp\left(-\frac{1}{3g'}\right) \\ &= \frac{1}{2} + \frac{3g}{\pi}\sqrt{\frac{3}{2\pi}}\int_0^1 du\,\frac{\sqrt{-\ln u}}{1-3g\ln u}\end{aligned} \tag{3.108}$$

where we have made the change of variable $u = \exp(-1/3g')$ to facilitate numerical evaluation of the integral. This integral is perfectly well-defined and convergent, as long as we do not try to expand it as a power series in g. Now we can subtract the divergent series (3.106) from the perturbation series and replace it by its integral representation (3.108), to obtain a much better approximation:

$$\mathcal{E}(g) \approx \mathcal{E}_d(g) + 0.165g + 0.006g^2 + \cdots \tag{3.109}$$

3.4.4 Variational method

A simple variational approach to the anharmonic oscillator is to use a trial function of the same form as the SHO ground-state wave function but with a variable width:

$$\Psi = N\,e^{-\alpha y^2} \tag{3.110}$$

This is similar to Problem 3, but here we have a quadratic as well as a quartic term in the potential:

$$\begin{aligned}\langle\hat{H}\rangle = \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} &= \frac{\int_{-\infty}^{\infty} dy\; e^{-2\alpha y^2}\left[2\alpha - 4\alpha^2 y^2 + \frac{1}{4}y^2(1+gy^2)\right]}{\int_{-\infty}^{\infty} dy\; e^{-2\alpha y^2}} \\ &= \alpha + \frac{1}{16\alpha} + \frac{3g}{64\alpha^2}\end{aligned} \tag{3.111}$$

This is minimized when

$$\frac{\partial}{\partial\alpha}\langle\hat{H}\rangle = 1 - \frac{1}{16\alpha^2} - \frac{3g}{32\alpha^3} = 0 \tag{3.112}$$

i.e. when

$$g = \frac{2}{3}\alpha(16\alpha^2 - 1) \tag{3.113}$$

For any $\alpha > \frac{1}{4}$, we can find g from eq. (3.113) and, substituting in (3.111), an approximation to the ground state energy $\mathcal{E}$ for that value of g.

3.4.5 Linear variation functions

A more laborious but potentially more accurate approach is to apply the linear variation method using the lowest N SHO eigenstates $\{|n\rangle, n = 0, \ldots, N-1\}$ as a basis. The necessary matrix elements $Y_{nn'} = \langle n|y^4|n'\rangle$ can be evaluated as follows. Using the SHO ladder operators we have

$$y^2|n\rangle = \sqrt{n(n-1)}|n-2\rangle + (2n+1)|n\rangle + \sqrt{(n+1)(n+2)}|n+2\rangle \tag{3.114}$$

Taking the inner product of $y^2|n\rangle$ and $y^2|n'\rangle$ and using $\langle n|n'\rangle = \delta_{n,n'}$ then gives

$$\begin{aligned} Y_{nn'} &= 3[1+2n(n+1)]\delta_{n,n'} + \Big[2(2n-1)\sqrt{n(n-1)}\delta_{n,n'+2} \\ &+ \sqrt{n(n-1)(n-2)(n-3)}\delta_{n,n'+4} + (n \leftrightarrow n')\Big] \end{aligned} \tag{3.115}$$

The $N \times N$ secular equation to be solved is then

$$|\mathbf{H}^{(0)} + \frac{1}{4}g\mathbf{Y} - \mathcal{E}\mathbf{I}| = 0 \tag{3.116}$$

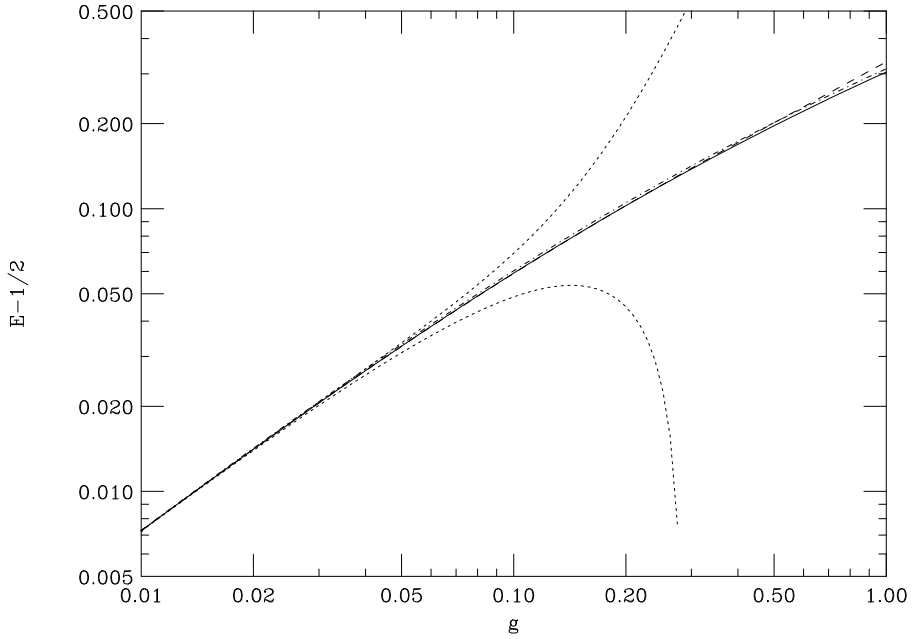
where $\mathbf{H}^{(0)}$ represents the SHO Hamiltonian matrix

$$H^{(0)}_{nn'} = \left(n + \frac{1}{2}\right)\delta_{n,n'} \tag{3.117}$$

We know that the smallest root of this equation will be an upper bound on the ground-state energy $\mathcal{E}(g)$, the bound improving as N increases.

3.4.6 Numerical results

The plot below shows the results of the various approaches discussed above.



The solid curve shows the result for $\mathcal{E} - \frac{1}{2}$ from linear variation, using the $n = 0, \ldots, 20$ SHO eigenstates as a basis. This method gives essentially the exact result (within the width of the line) over the range shown, $0 < g < 1$. The short-dashed curves show second-order (below) and third-order (above) perturbation theory. The long-dashed curve is eq. (3.109), while the nearby dot-dashed curve is the simple variational approximation obtained from eqs. (3.111) and (3.113).

We see that the integral representation used in eq. (3.109) gives a very good approximation up to $g \sim 0.4$. The simple variational approximation is less good in this range, but somewhat better at higher values of g. Perturbation theory fails badly above $g = 0.1$, due to the factorial divergence identified above. Nevertheless, the truncated perturbation series still provides a useful approximation over a limited range. This is because the perturbation series is in fact an *asymptotic expansion*.

3.5 Asymptotic expansions

As we have seen, the perturbation series is certainly 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 $f(g)$,

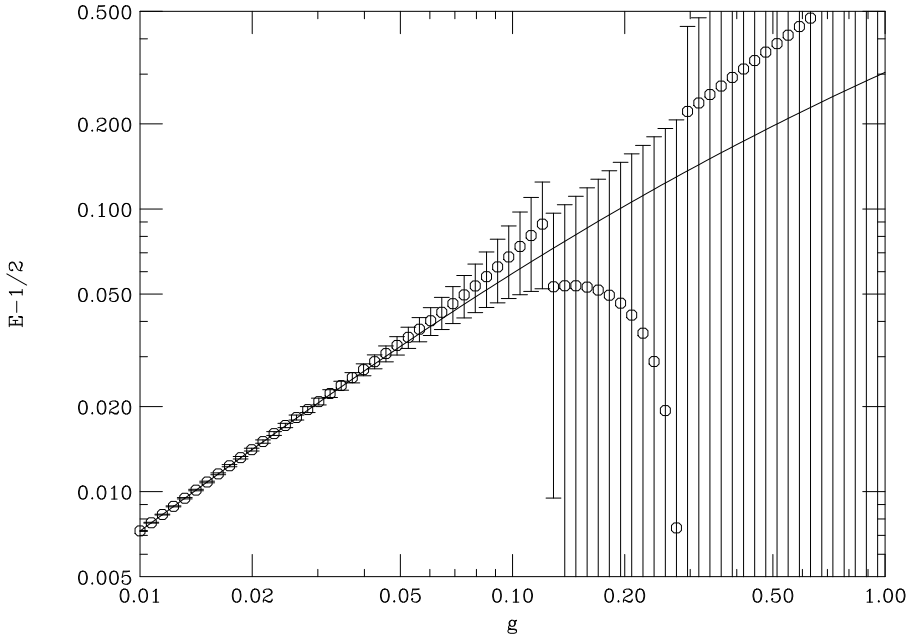
$$f(g) = \sum_{k=0}^{m} a_k \, g^k + R_m(g) \tag{3.118}$$

is characterised by the following behaviour of the remainder:

$$\begin{aligned} \lim_{g \to 0} \frac{R_m(g)}{g^m} &= 0 \\ \lim_{m \to \infty} R_m(g) &= \infty \end{aligned} \tag{3.119}$$

It follows that there exists a range of g for which the remainder $R_m(g)$ is smaller than the last term in the truncated series, $a_m \, g^m$. It therefore makes sense always to truncate the series at its smallest term. Because of the second property above, there is a limit to the accuracy we can achieve, which is of the order of the smallest term.

In the case of the anharmonic oscillator, for example, this procedure yields the estimates of the ground-state energy shown below.



The points show the partial sum of the perturbation series, up to and including the smallest term, and the error bars show the magnitude of the smallest term. We see that the true value, shown by the curve, does indeed always lie within the error bars, although the estimate is not of much use above $g \sim 0.1$.

Problem 12: Derive the perturbation series for the energy of the *first excited state* of the anharmonic oscillator, up to second order in g. What is the largest value of g for which you would expect

this expression to provide an estimate of the true energy? At that value of g, what is the estimate and its uncertainty?

Answer: $\mathcal{E}_1 = \frac{3}{2} + \frac{15}{4}g - \frac{165}{8}g^2\,; \quad g = \frac{2}{11}\,; \quad \mathcal{E}_1 = \frac{3}{2} \pm \frac{15}{22}$.

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:

$$V(r) = \begin{cases} -V_0 & r \leq a; \\ 0, & r > a. \end{cases} \tag{4.1}$$

where $V_0 > 0$. An incident particle will feel an attractive potential when it is within the spherical region of radius $r \leq a$. Note that this potential is spherically symmetric i.e. it depends only on the distance from the origin. The time independent Schrödinger equation is:

$$\Big[-\frac{\hbar^2}{2m}\nabla^2 + V(r) \Big] \Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \tag{4.2}$$

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:

$$\Big[-\frac{\hbar^2}{2m}\Big(\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2}\Big) + V(r)\Big] R_l(r) = E\, R_l(r) \tag{4.4}$$

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:

$$\Big[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2} + k^2\Big] R_l(r) = 0 \tag{4.5}$$

Substituting $\rho = kr$ into (4.5) gives

$$\Big[\frac{d^2}{d\rho^2} + \frac{2}{\rho}\frac{d}{d\rho} - \frac{l(l+1)}{\rho^2} + 1\Big] R_l(\rho) = 0 \tag{4.6}$$

The general solution to (4.6) are the **spherical Bessel functions**, $j_l(\rho)$ and the **spherical Neumann functions** $n_l(\rho)$:

$$R_l(\rho) = A j_l(\rho) + B n_l(\rho) \tag{4.7}$$

(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:

$$R_l(r) = A j_l(kr) \tag{4.8}$$

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) takes the form:

$$R_l(r) = \begin{cases} A j_l(k'r) & r < a; \\ B j_l(kr) + C n_l(kr), & r > a. \end{cases}$$

where A, B and C are constants to be determined from the boundary conditions and

$$\begin{aligned} k &= \sqrt{2mE}/\hbar \\ k' &= \sqrt{2m(E+V_0)}/\hbar \end{aligned} \tag{4.10}$$

From the continuity conditions (i.e. both $R_l(r)$ and its first derivative must be continuous at $r = a$), we have

$$k' \frac{j_l'(k'r)}{j_l(k'r)}\bigg|_{r=a} = k\left[\frac{Bj_l'(kr) + Cn_l'(kr)}{Bj_l(kr) + Cn_l(kr)}\right]_{r=a} \tag{4.11}$$

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 k').

Problem 1: Show that for $l = 0$ i.e. *s-wave scattering*, the value of the phase shift, $\delta_0(k)$ is determined by:

$$k' \cot k'a = k\cot(ka + \delta_0(k)) \tag{4.15}$$

and hence

$$\delta_0(k) = \arctan\left(\frac{k}{k'}\tan k'a\right) - ka \tag{4.16}$$

4.3 Mathematical preliminaries

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 $\mathrm{Res} f(z_0)$ 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 $\mathrm{Res} f(0) = 1$.

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):

$$\begin{aligned} j_l(\rho) &= (-\rho)^l \left(\frac{1}{\rho}\frac{d}{d\rho}\right)^l \left(\frac{\sin\rho}{\rho}\right) \\ n_l(\rho) &= -(-\rho)^l \left(\frac{1}{\rho}\frac{d}{d\rho}\right)^l \left(\frac{\cos\rho}{\rho}\right) \end{aligned} \tag{4.18}$$

where $l = 0, 1, 2, \ldots$. They can be expressed by the ascending power series

$$\begin{aligned} j_l(\rho) &= \frac{\rho^l}{(2l+1)!!}\left[1 - \frac{\frac{1}{2}\rho^2}{1!(2l+3)} + \frac{(\frac{1}{2}\rho^2)^2}{2!(2l+3)(2l+5)} - \cdots\right] \\ n_l(\rho) &= -\frac{(2l-1)!!}{\rho^{l+1}}\left[1 - \frac{\frac{1}{2}\rho^2}{1!(1-2l)} + \frac{(\frac{1}{2}\rho^2)^2}{2!(1-2l)(3-2l)} - \cdots\right] \end{aligned} \tag{4.19}$$

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:

$$\begin{aligned} h_l^{(1)}(\rho) &= j_l(\rho) + i n_l(\rho) \\ h_l^{(2)}(\rho) &= [h_l^{(1)}(\rho)]^* \end{aligned} \tag{4.20}$$

The explicit forms for the first few values of l are:

$$j_0(\rho) = \frac{\sin\rho}{\rho}$$

$$
\begin{aligned}
n_0(\rho) &= -\frac{\cos\rho}{\rho} \\
h_0^{(1)}(\rho) &= \frac{e^{i\rho}}{i\rho} \\
j_1(\rho) &= \frac{\sin\rho}{\rho^2} - \frac{\cos\rho}{\rho} \\
n_1(\rho) &= -\frac{\cos\rho}{\rho^2} - \frac{\sin\rho}{\rho} \\
h_1^{(1)}(\rho) &= -\frac{e^{i\rho}}{\rho}\left(1 + \frac{i}{\rho}\right)
\end{aligned} \tag{4.21}
$$

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**:

$$
\begin{aligned}
\frac{2l+1}{\rho} z_l(\rho) &= z_{l-1}(\rho) + z_{l+1}(\rho) \\
z_l^{'}(\rho) &= \frac{1}{(2l+1)}\Big[l z_{l-1}(\rho) - (l+1) z_{l+1}(\rho)\Big]
\end{aligned} \tag{4.22}
$$

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**:

$$
j_l(\rho) = \frac{(-i)^l}{2} \int_{-1}^{1} dz P_l(z) e^{i\rho z} \tag{4.23}
$$

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* ($\rho \ll 1, l$):

$$
\begin{aligned}
j_l(\rho) &\rightarrow \frac{\rho^l}{(2l+1)!!} \\
n_l(\rho) &\rightarrow -\frac{(2l-1)!!}{\rho^{l+1}}
\end{aligned} \tag{4.24}
$$

For *large arguments* ($\rho \gg l$):

$$
j_l(\rho) \rightarrow \frac{1}{\rho} \sin\left(\rho - \frac{l\pi}{2}\right)
$$

$$\begin{aligned} n_l(\rho) &\rightarrow -\frac{1}{\rho}\cos\left(\rho - \frac{l\pi}{2}\right) \\ h_l^{(1)}(\rho) &\rightarrow -\frac{i}{\rho}e^{i(\rho - l\pi/2)} \end{aligned} \tag{4.25}$$

This means that, for large r, $h_l^{(1)}(kr)$ represents a purely **outgoing** spherical wave, whereas $h_l^{(2)}(kr) = [h_l^{(1)}(kr)]^*$ represents a purely **incoming** wave.

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

$$\begin{aligned} c_{lm}(\mathbf{k}) &\rightarrow A_l \\ Y_{lm}(\theta,\varphi) &\rightarrow \left[\frac{2l+1}{4\pi}\right]^{\frac{1}{2}} P_l(\cos\theta) \end{aligned} \tag{4.27}$$

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:

$$\frac{d\sigma}{d\Omega} = |f(\theta,\varphi)|^2 \tag{4.38}$$

4.5 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:

$$\begin{aligned}
&[\nabla^2 + k^2 - U(r)]\psi(\mathbf{r}) = 0 \\
&U(r) = \frac{2mV(r)}{\hbar^2} \\
&k^2 = \frac{2mE}{\hbar^2}
\end{aligned} \tag{4.40}$$

4.5.1 Partial wave expansion

We now decompose the $\psi(\mathbf{r})$ into spherical **partial 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

$$\begin{aligned} e^{i\mathbf{k}\cdot\mathbf{r}} &= \sum_{l=0}^{\infty} i^l \ (2l+1) j_l(kr) \ P_l(\cos\theta) \\ &= \frac{1}{2}\sum_{l=0}^{\infty} i^l (2l+1) \left[h_l^{(2)}(kr) + h_l^{(1)}(kr)\right] P_l(\cos\theta) \end{aligned} \tag{4.45}$$

Not surprisingly, the expansion of the plane wave representing the beam consists of equal amounts of incoming ($h_l^{(2)}$) and outgoing ($h_l^{(1)}$) spherical waves. The effect of the potential is to cause scattering and hence to modify the amplitude of each of the **outgoing** spherical waves only. 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. 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 affects only the outgoing spherical waves. Equation (4.46) can now be rewritten as:

$$\begin{aligned} \psi(\mathbf{r}) \quad \underset{r \to \infty}{\longrightarrow} \quad & \frac{1}{2}\sum_{l=0}^{\infty} i^l \ (2l+1) \left[h_l^{(1)}(kr) + h_l^{(2)}(kr) + [S_l(k) - 1] h_l^{(1)}(kr)\right] P_l(\cos\theta) \\ = \quad & \sum_{l=0}^{\infty} i^l (2l+1) \left\{ j_l(kr) + \frac{1}{2}[S_l(k) - 1] h_l^{(1)}(kr) \right\} P_l(\cos\theta) \\ = \quad & e^{i\mathbf{k}\cdot\mathbf{r}} + \sum_{l=0}^{\infty} i^l \ (2l+1) \ \frac{1}{2} \ [S_l(k) - 1] \ h_l^{(1)}(kr) \ P_l(\cos\theta) \end{aligned} \tag{4.47}$$

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 $S_l(k)$.

As we discussed at the beginning of this chapter, in general the identities of the scattered particles (and the target) may change in a scattering process – this is called inelastic scattering. In some circumstances, however, only elastic scattering is possible. In that case, by conservation of the number of particles, the *flux* of outgoing particles cannot be changed by the scattering process. Furthermore, by angular momentum conservation, this must be true for *each partial wave separately.* Consequently the only thing that can be changed is the *phase* of each partial wave. Therefore, for purely elastic scattering, 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

$$\sigma_l = \frac{4\pi}{k^2}(2l+1)\sin^2\delta_l \tag{4.55}$$

It follows that for elastic scattering

$$\sigma_l \leq \frac{4\pi}{k^2}(2l+1) \tag{4.56}$$

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.

4.5.2 The optical theorem

From eqs. (4.49) and (4.51), the imaginary part of the scattering amplitude is given by

$$\begin{aligned} \mathrm{Im}f(\theta) &= \frac{1}{2k}\sum_{l=0}^{\infty}(2l+1)\,\mathrm{Re}[1-S_l]\;P_l(\cos\theta) \\ &= \frac{1}{2k}\sum_{l=0}^{\infty}(2l+1)\;[2\sin^2\delta_l]\;P_l(\cos\theta) \end{aligned} \tag{4.58}$$

Setting $\theta = 0$ and making use of $P_l(1) = 1$, we get

$$\mathrm{Im}f(\theta=0) = \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)\;\sin^2\delta_l \tag{4.59}$$

Comparing this with (4.53), we obtain the **optical theorem**:

$$\sigma_{tot} = \frac{4\pi}{k}\,\mathrm{Im}f(\theta=0) \tag{4.60}$$

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 $ka \ll 1$.

In general, both elastic and inelastic scattering may occur, and then the outgoing flux of particles of the same type as the incoming beam may be reduced by scattering. In that case the amplitudes

$S_l(k)$ of the outgoing elastically-scattered waves are reduced and the corresponding phase shifts become complex. However, it can be shown that the optical theorem (4.60) remains valid, with σ_{tot} being to the total cross section summed over all types of scattering (elastic plus inelastic) and $f(\theta = 0)$ the purely elastic forward scattering amplitude.

4.6 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.

4.6.1 Integral form of the Schrödinger equation

The time-independent Schrödinger equation

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{4.61}$$

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 $\delta(\mathbf{r}-\mathbf{r}')$ may be written as

$$\delta(\mathbf{r}-\mathbf{r}') = \frac{1}{(2\pi)^3}\int_{all\ space} e^{i\mathbf{s}\cdot(\mathbf{r}-\mathbf{r}')}d\mathbf{s} \tag{4.65}$$

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 the **Green function** for the equation (4.62). In general, the Green 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

$$\begin{aligned} G(\mathbf{r}) &= \frac{i}{8\pi^2 r}[I_1 - I_2] \\ I_1 &= \int_{-\infty}^{\infty} \frac{s e^{isr}}{(s-k)(s+k)} ds \\ I_2 &= \int_{-\infty}^{\infty} \frac{s e^{-isr}}{(s-k)(s+k)} ds \end{aligned} \tag{4.76}$$

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.

4.6.2 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). Then 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 proportional to the *Fourier transform* of the scattering potential.

For a spherically symmetric potential, $V(\mathbf{r}) = V(r)$, there is no φ dependence and the Born approximation reduces to

$$f(\theta) = -\frac{2m}{q\hbar^2}\int_0^\infty r\sin(qr)V(r)dr \tag{4.88}$$

where q is the momentum transfer,

$$q = |\mathbf{k}' - \mathbf{k}| = 2k\sin(\theta/2) \tag{4.89}$$

To see this we use spherical polar coordinates with the polar axis along $\mathbf{k}' - \mathbf{k}$, so that

$$(\mathbf{k}' - \mathbf{k}) \cdot \mathbf{r}' = qr' \cos\theta' \tag{4.90}$$

Then (4.87) becomes

$$f(\theta) = -\frac{m}{2\pi\hbar^2} \int_0^\infty dr' \int_{-1}^1 d\cos\theta' \int_0^{2\pi} d\varphi'\, r'^2 e^{iqr'\cos\theta'} V(r') \tag{4.91}$$

Now the $\cos\theta'$ and φ' integrations can be performed simply and, writing the remaining variable of integration as r instead of r', we obtain (4.88).

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

$$f(\theta) = \frac{2Am}{\hbar^2}\left(\frac{r_0^2}{1 + q^2 r_0^2}\right)$$

The differential cross section in the limit $r_0 \to \infty$ (i.e. for the Coulomb potential) is thus

$$\frac{d\sigma}{d\Omega} = \frac{4A^2m^2}{\hbar^4 q^4}$$

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

$$f(\theta) = \frac{2mV_0}{\hbar^2 q^3}(\sin qa - qa\cos qa)$$

Problem 9: Determine the scattering amplitude in the first Born approximation for the Gaussian potential well:

$$V(r) = -Ae^{-br^2}$$

where $A,\ b > 0$.

Answer: The scattering amplitude is given by

$$f(\theta) = \frac{Am\sqrt{\pi}}{2\hbar^2 b^{3/2}} e^{-q^2/4b}$$

4.6.3 Low-energy scattering

In general, the scattering amplitude in the Born approximation has a complicated angular dependence, as seen in the above examples through their dependence on $q = 2k\sin(\theta/2)$. This corresponds to a large number of terms contributing significantly in the partial wave expansion. However, 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 (4.87) simplifies to

$$f(\theta,\varphi) \cong -\frac{m}{2\pi\hbar^2}\int_{all\ space} V(\mathbf{r}')d\mathbf{r}' \tag{4.92}$$

which is independent of θ and φ and thus represents *pure s-wave scattering*. Now from Eq. (4.51) we see that the scattering amplitude for pure s-wave scattering is

$$f = \frac{1}{k}e^{i\delta_0(k)}\sin\delta_0(k) \tag{4.93}$$

where $\delta_0(k)$ is the s-wave phase shift. In the Born approximation we assume that the scattering potential is weak and therefore the phase shift is also small, so we can write

$$f \cong \frac{1}{k}\delta_0(k) \tag{4.94}$$

Comparing with (4.92) we see that the s-wave phase shift is given to first order in V by

$$\delta_0(k) \cong -\frac{km}{2\pi\hbar^2}\int_{all\ space} V(\mathbf{r}')d\mathbf{r}' \tag{4.95}$$

Example: For the square well potential (4.1) we find from (4.95) that

$$\delta_0(k) \cong -\frac{km}{2\pi\hbar^2}\left(-\frac{4}{3}\pi a^3 V_0\right) = \frac{2km}{3\hbar^2}a^3V_0 \tag{4.96}$$

We saw in Problem 1 that the exact s-wave phase shift is

$$\delta_0(k) = \arctan\left(\frac{k}{k'}\tan k'a\right) - ka \tag{4.97}$$

where $k' = \sqrt{2m(E+V_0)}/\hbar$. At low energy, k is small and so

$$\delta_0(k) \cong \frac{k}{k'}(\tan k'a - k'a) \tag{4.98}$$

where $k' \cong \sqrt{2mV_0}/\hbar$. If, in addition, the potential is weak, then k' is also small and $\tan k'a \cong k'a + (k'a)^3/3$. Thus

$$\delta_0(k) \cong \frac{1}{3}kk'^2a^3 \cong \frac{2kmV_0}{3\hbar^2}a^3 \tag{4.99}$$

in agreement with (4.96).

4.7 Beyond the (first) Born approximation

To see how to extend the Born approximation to higher orders in the scattering potential, we must develop the formal theory of scattering a little more. Assume that the Hamiltonian for the scattering problem is

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{4.100}$$

where $\hat{H}_0$ is the free-particle Hamiltonian,

$$\hat{H}_0 = \frac{\hat{p}^2}{2m} \tag{4.101}$$

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

$$\hat{H}_0|\mathbf{p}\rangle = \frac{p^2}{2m}|\mathbf{p}\rangle \tag{4.102}$$

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. (4.100) with the *same* energy eigenvalue as Eq. (4.102).

Let $|\phi\rangle$ be an energy eigenstate of $\hat{H}_0$

$$\hat{H}_0|\phi\rangle = E|\phi\rangle \tag{4.103}$$

where we choose $|\phi\rangle$ to be the incident plane wave state. We want to solve

$$\hat{H}|\psi\rangle = \left(\hat{H}_0 + \hat{V}\right)|\psi\rangle = E|\psi\rangle \tag{4.104}$$

with both $\hat{H}_0$ and $\hat{H}$ exhibiting **continuous** energy spectra. We look for solutions to Eq. (4.104) such that

$$|\psi\rangle \to |\phi\rangle \quad \text{as} \quad \hat{V} \to 0 \tag{4.105}$$

where $|\phi\rangle$ is the solution to the free particle Schrödinger equation Eq. (4.103) with the *same* energy eigenvalue.

The desired solution (at least in a formal sense) is

$$|\psi\rangle = |\phi\rangle + \left(E - \hat{H}_0\right)^{-1}\hat{V}|\psi\rangle \tag{4.106}$$

although there are complications arising from the singular nature of the operator $(E - \hat{H}_0)^{-1}$.

Problem 10: Verify that Eq. (4.106) is indeed a solution to Eq. (4.104)

4.7.1 The Lippmann-Schwinger equation

The correct 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.

$$|\psi\rangle = |\phi\rangle + \frac{1}{(E - \hat{H}_0 + i\epsilon)}\hat{V}|\psi\rangle \tag{4.107}$$

Equation (4.107) is called the **Lippmann-Schwinger equation** and is a relation between state vectors, independent of any particular representation. However, it will be convenient here to write it in the position representation, which enables us to obtain the scattering amplitude in the Born approximation.

In the position representation (4.107) becomes

$$\langle \mathbf{r}|\psi\rangle = \langle \mathbf{r}|\phi\rangle + \int d\mathbf{r}'\langle \mathbf{r}|\frac{1}{(E - \hat{H}_0 + i\epsilon)}|\mathbf{r}'\rangle\langle \mathbf{r}'|\hat{V}|\psi\rangle \tag{4.108}$$

which is an **integral equation** for scattering because the unknown ket $|\psi\rangle$ appears under the integral sign. The quantity

$$G(\mathbf{r}, \mathbf{r}') = \frac{\hbar^2}{2m}\langle \mathbf{r}|\frac{1}{(E - \hat{H}_0 + i\epsilon)}|\mathbf{r}'\rangle \tag{4.109}$$

is just the Green function (4.77) for the inhomogeneous Helmholtz equation derived earlier. 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 $E = \hbar^2 k^2/2m$. Equation (4.107) 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 $\mathbf{k}'$:

$$f(\mathbf{k}', \mathbf{k}) = -\frac{m}{2\pi\hbar^2}\langle \mathbf{k}'|\hat{V}|\psi\rangle \tag{4.110}$$

which is *exact*.

4.7.2 The Born series

We now define the **transition operator** $\hat{T}$, such that

$$\hat{V}|\psi\rangle = \hat{T}|\phi\rangle \tag{4.111}$$

Multiplying the Lippmann-Schwinger equation (4.107) by $\hat{V}$, we obtain

$$\hat{T}|\phi\rangle = \hat{V}|\phi\rangle + \hat{V}\frac{1}{(E - \hat{H}_0 + i\epsilon)}\hat{T}|\phi\rangle \tag{4.112}$$

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:

$$\hat{T} = \hat{V} + \hat{V}\frac{1}{(E - \hat{H}_0 + i\epsilon)}\hat{T} \tag{4.113}$$

The scattering amplitude can now be written as

$$f(\mathbf{k}', \mathbf{k}) = -\frac{m}{2\pi\hbar^2}\langle \mathbf{k}'|\hat{T}|\mathbf{k}\rangle \tag{4.114}$$

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:

$$\hat{T} = \hat{V} + \hat{V}\frac{1}{(E - \hat{H}_0 + i\epsilon)}\hat{V} + \hat{V}\frac{1}{(E - \hat{H}_0 + i\epsilon)}\hat{V}\frac{1}{(E - \hat{H}_0 + i\epsilon)}\hat{V} + \cdots \tag{4.115}$$

Correspondingly, we can expand $f(\mathbf{k}', \mathbf{k})$ as follows:

$$f(\mathbf{k}', \mathbf{k}) = \sum_{n=1}^{\infty} f^{(n)}(\mathbf{k}', \mathbf{k}) \tag{4.116}$$

where n is the number of times the $\hat{V}$ operator enters. We have

$$\begin{aligned} f^{(1)}(\mathbf{k}', \mathbf{k}) &= -\frac{m}{2\pi\hbar^2}\langle \mathbf{k}'|\hat{V}|\mathbf{k}\rangle \\ f^{(2)}(\mathbf{k}', \mathbf{k}) &= -\frac{m}{2\pi\hbar^2}\langle \mathbf{k}'|\hat{V}\frac{1}{(E - \hat{H}_0 + i\epsilon)}\hat{V}|\mathbf{k}\rangle \end{aligned} \tag{4.117}$$

etc., which are just the first, second etc. order Born approximations.

Equation (4.116) 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 function is called the **momentum-space 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 momentum-space 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

Density Matrices

5.1 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.

5.2 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 said to be in **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 $\hat{A}$ 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

$$\langle\hat{A}\rangle = \langle\Psi|\hat{A}|\Psi\rangle \tag{5.1}$$

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:

$$\langle\hat{A}\rangle = \sum_i p_i \langle\Psi_i|\hat{A}|\Psi_i\rangle \tag{5.2}$$

where we have again assumed that the $|\Psi_n\rangle$ are normalised. Note that statistics enter into Eq. (5.2) 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.

5.3 Properties of the Density Operator

The density operator is defined by

$$\hat{\rho} = \sum_i p_i |\Psi_i\rangle\langle\Psi_i| \tag{5.3}$$

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

$$0 \leq p_i \leq 1, \quad \sum_i p_i = 1, \quad \sum_i p_i^2 \leq 1 \tag{5.4}$$

For a *pure state* there is just one p_i (which is equal to unity) and all the rest are zero. In that case

$$\hat{\rho} = |\Psi\rangle\langle\Psi| \quad \text{(pure state)} \tag{5.5}$$

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

$$|\Psi_i\rangle = \sum_n c_{ni} |\psi_n\rangle \tag{5.6}$$

and from the orthonormality of the $\{|\psi_i\rangle\}$,

$$c_{ni} = \langle\psi_n|\Psi_i\rangle \tag{5.7}$$

We now construct the **density matrix** which consists of the matrix elements of the density operator in the $\{|\psi_i\rangle\}$ basis:

$$\begin{aligned}\langle\psi_n|\hat{\rho}|\psi_m\rangle &= \sum_i p_i\langle\psi_n|\Psi_i\rangle\langle\Psi_i|\psi_m\rangle \\ &= \sum_i p_i\, c_{ni}\, c^*_{mi} \end{aligned} \tag{5.8}$$

which characterises $\hat{\rho}$ as a Hermitian operator since

$$\langle\psi_n|\hat{\rho}|\psi_m\rangle = \langle\psi_m|\hat{\rho}|\psi_n\rangle^* \tag{5.9}$$

(given that the p_i are real), i.e we have

$$\hat{\rho} = \hat{\rho}^\dagger \tag{5.10}$$

From Eq. (5.8), the probability of finding the system in the state $|\psi_n\rangle$ is given by the diagonal element

$$\langle\psi_n|\hat{\rho}|\psi_n\rangle = \sum_i p_i\,|c_{ni}|^2 \tag{5.11}$$

which gives a physical interpretation of the diagonal elements of the density operator. Because probabilities are positive numbers, we have

$$\langle\psi_n|\hat{\rho}|\psi_n\rangle \geq 0 \tag{5.12}$$

The trace of $\hat{\rho}$ (i.e. the sum of the diagonal matrix elements) is

$$\begin{aligned}\mathrm{Tr}\,\hat{\rho} &= \sum_n \langle\psi_n|\hat{\rho}|\psi_n\rangle \\ &= \sum_i \sum_n p_i\,\langle\psi_n|\Psi_i\rangle\,\langle\Psi_i|\psi_n\rangle \\ &= \sum_i p_i\langle\Psi_i|\Psi_i\rangle \\ &= \sum_i p_i \\ &= 1 \end{aligned} \tag{5.13}$$

(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. (5.8) it follows that

$$0 \leq \langle\psi_n|\hat{\rho}|\psi_n\rangle \leq 1 \tag{5.14}$$

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 $\hat{\rho}^2$:

$$\begin{aligned}\langle\psi_n|\hat{\rho^2}|\psi_m\rangle &= \sum_k \langle\psi_n|\hat{\rho}|\psi_k\rangle\,\langle\psi_k|\hat{\rho}|\psi_m\rangle \\ &= \sum_i\sum_j\sum_k p_i\,p_j\,\langle\psi_n|\Psi_i\rangle\langle\Psi_i|\psi_k\rangle\,\langle\psi_k|\Psi_j\rangle\langle\Psi_j|\psi_m\rangle \qquad (5.15)\end{aligned}$$

where we have used Eq. (5.3).

Problem 1: Using (5.15), show that

$$\mathrm{Tr}\,\hat{\rho}^2 \leq 1 \qquad (5.16)$$

For a pure state, there is only one p_i and it is equal to unity. Therefore

$$\mathrm{Tr}\,\hat{\rho}^2 = 1 \quad \text{(pure state)} \qquad (5.17)$$

and

$$\begin{aligned}\hat{\rho}^2 &= |\Psi\rangle\langle\Psi|\Psi\rangle\langle\Psi| \\ &= |\Psi\rangle\langle\Psi| \\ &= \hat{\rho} \quad \text{(pure state)} \qquad (5.18)\end{aligned}$$

i.e. $\hat{\rho}$ is *idempotent* for a pure state. Thus whether a state is pure or not can be established by testing whether (5.17) or (5.18) is satisfied or not.

We now derive the expectation value of an operator $\hat{A}$ for pure as well as mixed states. Let

$$\langle\hat{A}\rangle_i = \langle\Psi_i|\hat{A}|\Psi_i\rangle \qquad (5.19)$$

and

$$\langle\hat{A}\rangle = \sum_i p_i\,\langle\hat{A}\rangle_i \qquad (5.20)$$

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 (5.20), 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

$$\langle\hat{A}\rangle = \langle\hat{A}\rangle_i \quad \text{(pure state)} \qquad (5.21)$$

Now consider the operator $\hat{\rho}\hat{A}$. From (5.3) we have

$$\hat{\rho}\hat{A} = \sum_i p_i \, |\Psi_i\rangle\langle\Psi_i| \, \hat{A} \tag{5.22}$$

In the $\{|\psi_i\rangle\}$ basis,

$$\langle\psi_n|\hat{\rho}\hat{A}|\psi_m\rangle = \sum_i p_i \, \langle\psi_n|\Psi_i\rangle\langle\Psi_i|\hat{A}|\psi_m\rangle \tag{5.23}$$

Taking the trace of $\hat{\rho}\hat{A}$,

$$\begin{aligned} \operatorname{Tr}\hat{\rho}\hat{A} &= \sum_n \langle\psi_n|\hat{\rho}\hat{A}|\psi_n\rangle \\ &= \sum_i \sum_n p_i \, \langle\psi_n|\Psi_i\rangle\langle\Psi_i|\hat{A}|\psi_n\rangle \\ &= \sum_i p_i \, \langle\Psi_i|\hat{A}|\Psi_i\rangle \\ &= \langle\hat{A}\rangle \end{aligned} \tag{5.24}$$

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: $\operatorname{Tr}\hat{\rho} = 1$.
- $\hat{\rho}$ 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 $\langle\hat{A}\rangle = \operatorname{Tr}\hat{\rho}\hat{A}$.

5.3.1 Density operator for spin states

Suppose the spin state of an electron is given by

$$|\Psi\rangle = |\uparrow\rangle \tag{5.25}$$

so that the density operator is

$$\hat{\rho} = |\uparrow\rangle\langle\uparrow| \tag{5.26}$$

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

$$\hat{\rho} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{5.27}$$

Problem 2: Verify (5.27) 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

$$|\Psi\rangle = a_1|\uparrow\rangle + a_2|\downarrow\rangle \tag{5.28}$$

with

$$|a_1|^2 + |a_2|^2 = 1$$

the density operator is

$$\hat{\rho} = \begin{pmatrix} |a_1|^2 & a_1 a_2^* \\ a_2 a_1^* & |a_2|^2 \end{pmatrix} \tag{5.29}$$

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

$$\hat{\rho} = c_0\, I + c_1\, \hat{S}_x + c_2\, \hat{S}_y + c_3\, \hat{S}_z \tag{5.30}$$

where I is the unit 2×2 matrix and the c_i's are real numbers. The density matrix becomes

$$\hat{\rho} = \begin{pmatrix} c_0 + \frac{1}{2}c_3 & \frac{1}{2}(c_1 - ic_2) \\ \frac{1}{2}(c_1 + ic_2) & c_0 - \frac{1}{2}c_3 \end{pmatrix} \tag{5.31}$$

(where we have set $\hbar = 1$).

Problem 3: Verify (5.31) 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

$$\Psi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \Psi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{5.32}$$

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 $\hat{\rho}$.

5.3.2 Density operator in the position representation

The density operator in the position representation is defined by

$$\begin{aligned} \rho(x', x) &= \langle x'|\hat{\rho}|x\rangle \\ &= \sum_i p_i\,\Psi_i(x')\,\Psi_i^*(x) \end{aligned} \tag{5.33}$$

which, for a pure state becomes

$$\rho(x', x) = \Psi(x')\,\Psi^*(x) \quad \text{(pure state)} \tag{5.34}$$

The expectation value for an operator $\hat{A}$ is then given by

$$\begin{aligned} \langle \hat{A} \rangle &= \mathrm{Tr}\,\hat{\rho}\,\hat{A} \\ &= \int dx\,\langle x|\hat{\rho}\hat{A}|x\rangle \\ &= \int dx\,\langle x|\hat{\rho}\Big(\int dx'|x'\rangle\langle x'|\Big)\hat{A}|x\rangle \\ &= \int\int dx'\,dx\,\langle x|\hat{\rho}|x'\rangle\,\langle x'|\hat{A}|x\rangle \\ &= \int\int dx'\,dx\,\rho(x, x')\,A(x', x) \end{aligned} \tag{5.35}$$

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

$$\langle x|\hat{\rho}|x'\rangle = \int \Psi^*(q, x')\, \Psi(q, x)\, dq \tag{5.36}$$

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

$$\Psi(q, x) = \Phi(q)\, \chi(x) \tag{5.37}$$

we are dealing with a *pure* state. Prove that in this case $\hat{\rho}$ is idempotent, i.e. that

$$\hat{\rho}^2 = \hat{\rho} \tag{5.38}$$

in agreement with Eq. (5.18).

5.4 Density operator in statistical mechanics

A very important application of the density operator is to describe a system in **thermal equilibrium**. Let $\{\Psi_n\}$ be the complete set of eigenstates of the Hamiltonian $\hat{H}$ for the system, which satisfy

$$\hat{H}\Psi_n = E_n\Psi_n \tag{5.39}$$

Then in thermal equilibrium at absolute temperature T the states are occupied according to the Boltzmann distribution

$$p_n = Ne^{-\beta E_n} \tag{5.40}$$

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, and N is a normalisation constant chosen to ensure that

$$\sum_n p_n = 1$$

The corresponding density operator $\hat{\rho}(\beta)$ is (from (5.3)):

$$\hat{\rho}(\beta) = N\sum_n e^{-\beta E_n}|\Psi_n\rangle\langle\Psi_n| \tag{5.41}$$

Since

$$e^{-\beta\hat{H}}|\Psi_n\rangle = e^{-\beta E_n}|\Psi_n\rangle \tag{5.42}$$

this enables us to write

$$\begin{aligned}
\hat{\rho}(\beta) &= N\sum_n e^{-\beta E_n}|\Psi_n\rangle\langle\Psi_n| \\
&= N\sum_n e^{-\beta\hat{H}}|\Psi_n\rangle\langle\Psi_n| \\
&= Ne^{-\beta\hat{H}}\sum_n |\Psi_n\rangle\langle\Psi_n| \\
&= Ne^{-\beta\hat{H}}
\end{aligned} \tag{5.43}$$

To determine N, we note that

$$\begin{aligned}
\mathrm{Tr}\,\hat{\rho}(\beta) &= N\,\mathrm{Tr}\,e^{-\beta\hat{H}} \\
&= 1
\end{aligned} \tag{5.44}$$

which implies

$$N = \frac{1}{\mathrm{Tr}\,e^{-\beta\hat{H}}} \tag{5.45}$$

Hence the density operator under thermal equilibrium is

$$\begin{aligned} \hat{\rho}(\beta) &= \frac{e^{-\beta\hat{H}}}{\mathrm{Tr}\, e^{-\beta\hat{H}}} \\ &= \frac{1}{Z}\, e^{-\beta\hat{H}} \end{aligned} \tag{5.46}$$

where

$$Z = \mathrm{Tr}\, e^{-\beta\hat{H}} \tag{5.47}$$

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

$$\begin{aligned} \langle \hat{A} \rangle &= \mathrm{Tr}\, \hat{\rho}\, \hat{A} \\ &= \frac{\mathrm{Tr}\left[e^{-\beta\hat{H}}\, \hat{A} \right]}{\mathrm{Tr}\left[e^{-\beta\hat{H}} \right]} \end{aligned} \tag{5.48}$$

The mean energy of the system (i.e. the internal energy) is given by U where

$$\begin{aligned} U &= \langle \hat{H} \rangle \\ &= \frac{\mathrm{Tr}\left[e^{-\beta\hat{H}}\, \hat{H} \right]}{\mathrm{Tr}\left[e^{-\beta\hat{H}} \right]} \\ &= -\frac{\partial}{\partial\beta} \ln \left(\mathrm{Tr}\, e^{-\beta\hat{H}} \right) \\ &= -\frac{\partial}{\partial\beta} \ln\, Z(T,V,N) \end{aligned} \tag{5.49}$$

From the partition function, we obtain all thermodynamic observables:

$$\begin{aligned} S &= -k\, \mathrm{Tr}\, (\hat{\rho} \ln \hat{\rho}) \qquad \text{(entropy)} \\ &= k\, \beta\, \langle \hat{H} \rangle + k \ln\, Z(T,V,N) \\ F &= U - TS \qquad \text{(Helmholtz free energy)} \end{aligned}$$

$$
\begin{aligned}
&= -kT \ln Z(T,V,N) \\
&= -kT \ln \left[\mathrm{Tr} \left\{ e^{-\beta \hat{H}} \right\} \right] \qquad (5.50)
\end{aligned}
$$

We now calculate the matrix elements of the density operator (in various representations) for some concrete cases.

5.4.1 Density matrix for a free particle in the momentum representation

We determine the density matrix in the momentum representation for a free particle in thermal equilibrium 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;

$$\hat{H}|\psi_{\mathbf{k}}\rangle = E|\psi_{\mathbf{k}}\rangle \qquad (5.51)$$

with

$$E = \frac{\hbar^2 \mathbf{k}^2}{2m} \qquad (5.52)$$

and $|\psi_{\mathbf{k}}\rangle$ defined by

$$
\begin{aligned}
\psi_{\mathbf{k}}(\mathbf{r}) &= \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}} \\
\mathbf{k} &= \frac{2\pi}{L}(n_x, n_y, n_z) \\
n_i &= 0, \pm 1, \pm 2, \ldots \qquad (5.53)
\end{aligned}
$$

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,

$$
\begin{aligned}
\langle \psi_{\mathbf{k}'} | \psi_{\mathbf{k}} \rangle &= \delta_{\mathbf{k},\mathbf{k}'} \\
&= \delta_{n_{x'},n_x}\, \delta_{n_{y'},n_y}\, \delta_{n_z,n_z} \qquad (5.54)
\end{aligned}
$$

and complete,

$$\sum_{\mathbf{k}} \psi^*_{\mathbf{k}}(\mathbf{r}')\psi_{\mathbf{k}}(\mathbf{r}) = \delta(\mathbf{r}' - \mathbf{r}) \qquad (5.55)$$

The canonical partition function is

$$\begin{aligned} Z(T,V,1) &= \mathrm{Tr}\, e^{-\beta\hat{H}} \\ &= \sum_{\mathbf{k}} \langle\psi_{\mathbf{k}}|e^{-\beta\hat{H}}|\psi_{\mathbf{k}}\rangle \\ &= \sum_{\mathbf{k}} e^{-\frac{\beta\hbar^2}{2m}\mathbf{k}^2} \end{aligned} \tag{5.56}$$

Since the eigenvalues $\mathbf{k}$ are very close together in a large volume, we can replace the sum in (5.56) by an integral.

$$\begin{aligned} Z(T,V,1) &= \frac{V}{(2\pi)^3}\int d\mathbf{k}\, e^{-\frac{\beta\hbar^2}{2m}\mathbf{k}^2} \\ &= \frac{V}{(2\pi)^3}\left(\frac{2m\pi}{\beta\hbar^2}\right)^{3/2} \\ &= \frac{V}{\lambda^3} \end{aligned} \tag{5.57}$$

where λ is called the *thermal wavelength*. The matrix elements of the density operator thus becomes

$$\langle\psi_{\mathbf{k}'}|\hat{\rho}|\psi_{\mathbf{k}}\rangle = \frac{\lambda^3}{V}\, e^{-\frac{\beta\hbar^2}{2m}\mathbf{k}^2}\, \delta_{\mathbf{k},\mathbf{k}'} \tag{5.58}$$

which is a diagonal matrix.

5.4.2 Density matrix for a free particle in the position representation

We can make a change of basis to find the density matrix for the same system in the position representation. We have

$$\begin{aligned} \langle\mathbf{r}'|\hat{\rho}|\mathbf{r}\rangle &= \sum_{\mathbf{k},\mathbf{k}'} \langle\mathbf{r}'|\mathbf{k}'\rangle\, \langle\mathbf{k}'|\hat{\rho}|\mathbf{k}\rangle\, \langle\mathbf{k}|\mathbf{r}\rangle \\ &= \sum_{\mathbf{k},\mathbf{k}'} \psi_{\mathbf{k}'}(\mathbf{r}')\left\{\frac{\lambda^3}{V}\, e^{-\frac{\beta\hbar^2}{2m}\mathbf{k}^2}\, \delta_{\mathbf{k},\mathbf{k}'}\right\}\psi^*_{\mathbf{k}}(\mathbf{r}) \\ &= \frac{\lambda^3}{V}\,\frac{1}{(2\pi)^3}\int d\mathbf{k}\, \exp\left\{-\frac{\beta\hbar^2}{2m}\mathbf{k}^2 + i\mathbf{k}\cdot(\mathbf{r}'-\mathbf{r})\right\} \end{aligned} \tag{5.59}$$

Problem 8: Show that Eq. (5.59) reduces to

$$\begin{aligned}\langle \mathbf{r}'|\hat{\rho}|\mathbf{r}\rangle &= \frac{\lambda^3}{V}\,\frac{1}{(2\pi)^3}\,\exp\left\{-\frac{m}{2\beta\hbar^2}(\mathbf{r}'-\mathbf{r})^2\right\}\left(\frac{2m\pi}{\beta\hbar^2}\right)^{3/2} \\ &= \frac{1}{V}\exp\left\{-\frac{\pi}{\lambda^2}(\mathbf{r}'-\mathbf{r})^2\right\} \qquad (5.60)\end{aligned}$$

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.

$$\langle \mathbf{r}'|\hat{\rho}|\mathbf{r}\rangle = \rho(\mathbf{r}) = \frac{1}{V} \qquad (5.61)$$

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.

5.4.3 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 in the quantum physics course:

$$\Psi_n(q) = \left(\frac{M\omega}{\pi\hbar}\right)^{1/4}\frac{H_n(x)}{\sqrt{2^n n!}}\,\exp\left\{-\frac{1}{2}x^2\right\} \qquad (5.62)$$

where q is the position and for brevity we have introduced the variable

$$x = \sqrt{\frac{M\omega}{\hbar}}\,q \qquad (5.63)$$

The energy eigenvalues are $E_n = \hbar\omega(n+\frac{1}{2})$ and H_n are the Hermite polynomials, defined by

$$\begin{aligned}H_n(x) &= (-1)^n e^{x^2}\left(\frac{d}{dx}\right)^n e^{-x^2} \\ &= \frac{e^{x^2}}{\sqrt{\pi}}\int_{-\infty}^{+\infty}(-2iu)^n\,\exp\{-u^2+2ixu\}\,du \qquad (5.64)\end{aligned}$$

The density matrix in the **energy representation** is simple:

$$\begin{aligned} \langle m|\hat{\rho}|n\rangle &= \rho_n \delta_{mn} \\ \rho_n &= \frac{1}{Z}\exp\Big\{-\beta\,\hbar\omega(n+\frac{1}{2})\Big\} \\ n &= 0,1,2,\ldots \end{aligned} \tag{5.65}$$

where

$$Z(T,V,1) = \Big[2\sinh\Big(\frac{1}{2}\beta\hbar\omega\Big)\Big]^{-1} \tag{5.66}$$

Problem 9: Verify Eqs. (5.65) and (5.66).

Problem 10: Show that in the **position representation** we have

$$\begin{aligned} \langle q'|\hat{\rho}|q\rangle &= \frac{1}{Z}\Big(\frac{M\omega}{\pi\hbar}\Big)^{1/2}\exp\Big\{-\frac{1}{2}(x^2+x'^2)\Big\} \\ &\times \sum_{n=0}^{\infty}\frac{1}{2^n n!}\exp\Big\{-\beta\,\hbar\omega(n+\frac{1}{2})\Big\}\,H_n(x)\,H_n(x') \end{aligned} \tag{5.67}$$

Hint: We have twice inserted the complete set of energy eigenfunctions.

Using the integral representation (5.64) for the Hermite polynomials in Eq. (5.67), we get:

$$\begin{aligned} \langle q'|\hat{\rho}|q\rangle &= \frac{1}{Z\pi}\Big(\frac{M\omega}{\pi\hbar}\Big)^{1/2}\exp\Big\{+\frac{1}{2}(x^2+x'^2)\Big\} \\ &\times \int_{-\infty}^{+\infty}du\int_{-\infty}^{+\infty}dv\sum_{n=0}^{\infty}\frac{(-2uv)^n}{n!}\exp\Big\{-\beta\,\hbar\omega(n+\frac{1}{2})\Big\}\exp\{-u^2+2ixu\} \\ &\times \exp\{-v^2+2ix'v\} \end{aligned} \tag{5.68}$$

The summation over n can be carried out as follows:

$$\begin{aligned} &\sum_{n=0}^{\infty}\frac{(-2uv)^n}{n!}\exp\Big\{-\beta\,\hbar\omega(n+\frac{1}{2})\Big\} \\ &= \exp\Big\{-\frac{1}{2}\beta\hbar\omega\Big\}\sum_{n=0}^{\infty}\frac{1}{n!}\Big[-2uv\exp(-\beta\hbar\omega)\Big]^n \\ &= \exp\Big\{-\frac{1}{2}\beta\hbar\omega\Big\}\exp\Big\{-2uv\,e^{-\beta\hbar\omega}\Big\} \end{aligned} \tag{5.69}$$

Then Eq. (5.68) becomes

$$\begin{aligned}\langle q'|\hat{\rho}|q\rangle &= \frac{1}{Z\pi}\left(\frac{M\omega}{\pi\hbar}\right)^{1/2}\exp\left\{+\frac{1}{2}(x^2+x'^2-\beta\hbar\omega)\right\} \\ &\times \int_{-\infty}^{+\infty}du\int_{-\infty}^{+\infty}dv\,\exp\left\{-u^2+2ixu-v^2+2ix'v-2uv\,e^{-\beta\hbar\omega}\right\} \qquad (5.70)\end{aligned}$$

The argument in the exponent is a quadratic form, which can be rewritten in the general form

$$-u^2+2ixu-v^2+2ix'v-2uv\,e^{-\beta\hbar\omega} = -\frac{1}{2}\mathbf{w}^T\cdot\mathbf{A}\cdot\mathbf{w}+i\mathbf{b}\cdot\mathbf{w} \qquad (5.71)$$

where

$$\begin{aligned}\mathbf{A} &= 2\begin{pmatrix}1 & e^{-\beta\hbar\omega}\\ e^{-\beta\hbar\omega} & 1\end{pmatrix} \\ \mathbf{b} &= 2\begin{pmatrix}x\\ x'\end{pmatrix} \\ \mathbf{w} &= \begin{pmatrix}u\\ v\end{pmatrix} \qquad (5.72)\end{aligned}$$

We now use the general formula

$$\int d^n\mathbf{w}\,\exp\left\{-\frac{1}{2}\mathbf{w}^T\cdot\mathbf{A}\cdot\mathbf{w}+i\mathbf{b}\cdot\mathbf{w}\right\} = \frac{(2\pi)^{n/2}}{[det\mathbf{A}]^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}\mathbf{b}^T\cdot\mathbf{A}^{-1}\cdot\mathbf{b}\right\} \qquad (5.73)$$

which holds if $\mathbf{A}$ is an invertible symmetric $n\times n$ matrix.

Problem 11: Verify (5.73).

Using (5.73) we get

$$\begin{aligned}\langle q'|\hat{\rho}|q\rangle &= \frac{1}{Z}\left(\frac{M\omega}{\pi\hbar}\right)^{1/2}\frac{e^{-\frac{1}{2}\beta\hbar\omega}}{[1-e^{-2\beta\hbar\omega}]^{\frac{1}{2}}} \\ &\times \exp\left\{\frac{1}{2}(x^2+x'^2)-[1-e^{-2\beta\hbar\omega}]^{-1}\,(x^2+x'^2-2xx'e^{-\beta\hbar\omega})\right\} \\ &= \frac{1}{Z}\left[\frac{M\omega}{2\pi\hbar\sinh(\beta\hbar\omega)}\right]^{\frac{1}{2}}\exp\left\{-\frac{1}{2}(x^2+x'^2)\coth(\beta\hbar\omega)+\frac{xx'}{\sinh(\beta\hbar\omega)}\right\} \qquad (5.74)\end{aligned}$$

Using the identity

$$\tanh\left(\frac{1}{2}\beta\hbar\omega\right) = \frac{\cosh(\beta\hbar\omega)-1}{\sinh(\beta\hbar\omega)} = \frac{\sinh(\beta\hbar\omega)}{1+\cosh(\beta\hbar\omega)} \qquad (5.75)$$

one finally gets

$$\begin{aligned}\langle q'|\hat{\rho}|q\rangle &= \frac{1}{Z}\left[\frac{M\omega}{2\pi\hbar\sinh(\beta\hbar\omega)}\right]^{\frac{1}{2}} \\ &\times \exp\left\{-\frac{M\omega}{4\hbar}\left[(q+q')^2\tanh\left(\frac{1}{2}\beta\hbar\omega\right)+(q-q')^2\coth\left(\frac{1}{2}\beta\hbar\omega\right)\right]\right\} \qquad (5.76)\end{aligned}$$

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:

$$\rho(q) = \left[\frac{M\omega}{\pi\hbar}\tanh\left(\frac{1}{2}\beta\hbar\omega\right)\right]^{\frac{1}{2}} \exp\left\{-\frac{M\omega}{\hbar}\tanh\left(\frac{1}{2}\beta\hbar\omega\right)q^2\right\} \qquad (5.77)$$

which is a Gaussian distribution with width

$$\sigma_q = \left[\frac{\hbar}{2M\omega\tanh\left(\frac{1}{2}\beta\hbar\omega\right)}\right]^{\frac{1}{2}} \qquad (5.78)$$

Problem 12: Show that in the limit of high temperatures, $\beta\hbar\omega \ll 1$,

$$\rho(q) \approx \left(\frac{m\omega^2}{2\pi kT}\right)^{\frac{1}{2}} \exp\left\{-\frac{M\omega^2 q^2}{2kT}\right\} \qquad (5.79)$$

and at low temperature, $\beta\hbar\omega \gg 1$,

$$\rho(q) \approx \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{2}} \exp\left\{-\frac{M\omega q^2}{\hbar}\right\} \qquad (5.80)$$

Therefore the density matrix thus contains, for high temperatures, the classical limit, and for very low temperatures, the quantum mechanical ground state density.

Chapter 6

Lie Groups

6.1 Introduction

We have already seen several examples of transformations of quantum mechanical state vectors, of the form

$$|\psi\rangle \to |\psi'\rangle = \hat{U}|\psi\rangle \tag{6.1}$$

To preserve the inner products of state vectors under this transformation, we require

$$\langle\phi'|\psi'\rangle = \langle\phi|\hat{U}^\dagger\hat{U}|\psi\rangle = \langle\phi|\psi\rangle \tag{6.2}$$

for all $|\phi\rangle$ and $|\psi\rangle$, or in other words that the transformation operator $\hat{U}$ is *unitary*, $\hat{U}^\dagger\hat{U} = \hat{I}$. Examples are the basis transformations discussed in Sect. 1.3.5, and the time development operator in Sect. 2.1. In the latter case, the transformation taking the state vector at time $t = 0$ to that at time t (in the Schrödinger picture) was found to be

$$\hat{U}_T(t) = \hat{T}(t, 0) = \exp\left[-\frac{i}{\hbar}\hat{H}t\right] \tag{6.3}$$

The time development operators for different values of t are an example of a *group* of transformation operators acting on the Hilbert space of state vectors. The general definition of a group G is a set of elements $\{a, b, c, \ldots\}$ with a rule of multiplication, $\{a, b\} \to ab \in G$ having the following properties:

- **Associativity**: for all $a, b, c \in G$ we have $a(bc) = (ab)c$
- **Unit element**: there exists $e \in G$ such that $ae = ea = a$ for all $a \in G$
- **Inverse**: for all $a \in G$ there exists an element, denoted by a^{-1}, such that $aa^{-1} = a^{-1}a = e$

If in addition the multiplication rule is commutative, $ab = ba$, then the group is said to be *Abelian*.

In the case of the time development operators (6.3) we clearly have an Abelian group, with

$$\begin{aligned} \hat{U}_T(t_1)\hat{U}_T(t_2) &= \hat{U}_T(t_2)\hat{U}_T(t_1) = \hat{U}_T(t_1+t_2) \\ \hat{U}_T(t)^{-1} &= \hat{U}_T(-t) \\ \text{Unit element} \qquad & \hat{U}_T(0) = \hat{I} \end{aligned} \tag{6.4}$$

This group has an infinite number of elements, corresponding to the different possible values of the continuous real parameter t. The elements are differentiable with respect to t: in fact

$$\frac{d}{dt}\hat{U}_T = -\frac{i}{\hbar}\hat{H}\hat{U}_T \tag{6.5}$$

so that

$$\hat{U}_T(t+dt) = \left(1 - \frac{i}{\hbar}\hat{H}\,dt\right)\hat{U}_T(t) \tag{6.6}$$

Thus the operator $-i\hat{H}\,dt/\hbar$ has the effect of producing the infinitesimal development in time dt. We say that the operator $\hat{H}/\hbar$ is the *generator* of the group of time development operators $\{\hat{U}_T\}$.

6.1.1 The translation group

A somewhat more complicated group is the group of *spatial translations* $\{\hat{U}_D\}$. In the position representation, these operate on the wave function $\psi(\mathbf{r}, t)$ to produce displacements:

$$\psi(\mathbf{r}, t) \to \psi'(\mathbf{r}, t) \quad \text{such that} \quad \psi'(\mathbf{r}+\mathbf{a}, t) = \psi(\mathbf{r}, t) \tag{6.7}$$

Thus we have

$$\psi'(\mathbf{r}, t) = \psi(\mathbf{r}-\mathbf{a}, t) = \hat{U}_D(\mathbf{a})\psi(\mathbf{r}, t) \tag{6.8}$$

We see that this is again an Abelian group, since

$$\begin{aligned} \hat{U}_D(\mathbf{a})\hat{U}_D(\mathbf{b})\psi(\mathbf{r}, t) &= \hat{U}_D(\mathbf{a})\psi(\mathbf{r}-\mathbf{b}, t) = \psi(\mathbf{r}-\mathbf{b}-\mathbf{a}, t) \\ &= \psi(\mathbf{r}-\mathbf{a}-\mathbf{b}, t) = \hat{U}_D(\mathbf{b})\hat{U}_D(\mathbf{a})\psi(\mathbf{r}, t) \end{aligned} \tag{6.9}$$

For an infinitesimal change $\mathbf{a} \to \mathbf{a} + \mathrm{d}\mathbf{a}$ we have

$$\psi(\mathbf{r}-\mathbf{a}-\mathrm{d}\mathbf{a}, t) = (1 - \mathrm{d}\mathbf{a}\cdot\nabla)\psi(\mathbf{r}-\mathbf{a}, t) \tag{6.10}$$

so that

$$\hat{U}_D(\mathbf{a}+\mathrm{d}\mathbf{a}) = (1-\mathrm{d}\mathbf{a}\cdot\nabla)\hat{U}_D(\mathbf{a}) \tag{6.11}$$

By analogy with the group of time development operators, we write this as

$$\hat{U}_D(\mathbf{a}+\mathrm{d}\mathbf{a}) = \left(1 - i\,\mathrm{d}\mathbf{a}\cdot\hat{\mathbf{D}}\right)\hat{U}_D(\mathbf{a}) \tag{6.12}$$

where the three operators $\hat{D}_j$ $(j = x, y, z)$ are the generators of the translation group,

$$\left(\hat{D}_x, \hat{D}_y, \hat{D}_z\right) = -i\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \tag{6.13}$$

In the case of a finite displacement a_x in the x-direction, we can Taylor expand the transformed wave function, to obtain

$$\begin{aligned}
\psi'(x,y,z,t) &= \psi(x-a_x,y,z,t) = \psi - a_x\frac{\partial\psi}{\partial x} + \frac{1}{2}a_x^2\frac{\partial^2\psi}{\partial x^2} - \ldots \\
&= \sum_{n=0}^{\infty}\frac{1}{n!}\left(-a_x\frac{\partial}{\partial x}\right)^n\psi \\
&= \sum_{n=0}^{\infty}\frac{1}{n!}\left(-ia_x\hat{D}_x\right)^n\psi \\
&= \exp\left(-ia_x\hat{D}_x\right)\psi
\end{aligned} \tag{6.14}$$

Thus when $\mathbf{a} = (a_x, 0, 0)$ we have $\hat{U}_D(\mathbf{a}) = \exp(-ia_x\hat{D}_x)$. Similarly, when $\mathbf{a} = (0, a_y, 0)$ we have $\hat{U}_D(\mathbf{a}) = \exp(-ia_y\hat{D}_y)$, and when $\mathbf{a} = (0, 0, a_z)$, $\hat{U}_D(\mathbf{a}) = \exp(-ia_z\hat{D}_z)$. Therefore, since the group is Abelian, we can write

$$\begin{aligned}
\hat{U}_D(\mathbf{a}) &= e^{-ia_x\hat{D}_x}e^{-ia_y\hat{D}_y}e^{-ia_z\hat{D}_z} \\
&= e^{-i(a_x\hat{D}_x+a_y\hat{D}_y+a_z\hat{D}_z)} = e^{-i\mathbf{a}\cdot\hat{\mathbf{D}}}
\end{aligned} \tag{6.15}$$

We recall from Chapter 1 that it is **not** true in general that

$$e^{\hat{A}}e^{\hat{B}} = e^{\hat{A}+\hat{B}} \tag{6.16}$$

This follows only if the operators $\hat{A}$ and $\hat{B}$ commute, $[\hat{A}, \hat{B}] = 0$. Thus the second line of Eq. (6.15) follows from the first only because $\hat{D}_x, \hat{D}_y$ and $\hat{D}_z$ commute with each other.

Notice that the spatial translation operators, like the time development operators, are unitary, since the generators are Hermitian:

$$\hat{U}_D^\dagger(\mathbf{a}) = e^{+i\mathbf{a}\cdot\hat{\mathbf{D}}^\dagger} = e^{+i\mathbf{a}\cdot\hat{\mathbf{D}}} = U_D(-\mathbf{a}) = [U_D(\mathbf{a})]^{-1} \tag{6.17}$$

We recall that this is because these operators preserve the inner products between state vectors, Eq. (6.2).

In general, a continuous group of operators $\hat{G}(a_1, a_2, \ldots, a_n)$ that can be written in the form

$$\hat{G}(a_1, a_2, \ldots, a_n) = \exp\left(-i \sum_{j=1}^{n} a_j \hat{X}_j\right) \tag{6.18}$$

where $a_1, a_2, \ldots, a_n$ are real parameters, is called a *Lie group* of dimension n, with generators $\{\hat{X}_j\}$. The number of generators that commute with each other, $[\hat{X}_j, \hat{X}_k] = 0$, is the *rank* of the group. Thus the group of time development operators is a Lie group with dimension = rank = 1, while the group of spatial displacements has dimension = rank = 3.

It should be clear that any Abelian Lie group has rank equal to its dimension. We shall see, however, that many of the important Lie groups in physics are *non-Abelian*, i.e. some or all of their generators do not commute with each other.

6.1.2 Symmetries and constants of the motion

We recognise the generators of the translation group, $\hat{D}_j$ in Eq. (6.13), as the quantum mechanical momentum operators, up to an overall factor of $\hbar$:

$$(\hat{p}_x, \hat{p}_y, \hat{p}_z) = -i\hbar \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) = \hbar \left(\hat{D}_x, \hat{D}_y, \hat{D}_z\right) \tag{6.19}$$

This is not a coincidence: the generators of Lie groups are often related to quantum mechanical observables. In particular, if the group that they generate is a *symmetry group* of the system, then the corresponding observables will be *constants of the motion*, i.e. conserved quantities whose values are time-independent.

By a symmetry group we mean in general a group of transformations $\{\hat{U}\}$ such that the transformed system behaves in exactly the same way as the untransformed one. In particular, if we transform any state of the system $|\psi\rangle$, allow it to evolve for a time, and then perform the inverse transformation, we shall find exactly the same state as if the transformations had not been performed at all. In mathematical terms, this means that if $\hat{U}$ is a symmetry transformation and $\hat{U}_T$ is the time development operator, then

$$\hat{U}^{-1}\hat{U}_T\hat{U}|\psi\rangle = \hat{U}_T|\psi\rangle \tag{6.20}$$

for any $|\psi\rangle$, which implies the operator equation

$$\hat{U}^{-1}\hat{U}_T\hat{U} = \hat{U}_T \tag{6.21}$$

or, acting on the left with $\hat{U}$,

$$\hat{U}_T\hat{U} = \hat{U}\hat{U}_T \tag{6.22}$$

In other words, the symmetry transformation operator and the time development operator commute. Now if $\hat{U}$ is an element of a Lie group,

$$\hat{U} = \exp\left(-i\sum a_j\hat{X}_j\right) \tag{6.23}$$

then, recalling Eq. (6.3) for $\hat{U}_T$, this can only be the case if the generators $\hat{X}_j$ commute with the Hamiltonian operator $\hat{H}$.

Thus we have shown that if a Lie group is a symmetry group then its generators must commute with the Hamiltonian. Furthermore, since the symmetry transformation operators must be unitary, the generators must be Hermitian operators. They can therefore be associated with quantum mechanical observables. And from Chapter 2 we know that observables that commute with the Hamiltonian represent constants of the motion.

Consider for example the case of a single particle moving freely in the absence of any forces. The position of the origin of coordinates is arbitrary, so if we make a translation $\mathbf{r} \to \mathbf{r} - \mathbf{a}$ at time $t = 0$, then allow the particle to move, then translate back, $\mathbf{r} \to \mathbf{r} + \mathbf{a}$ at time $t > 0$, the result will be equivalent to no translation. Thus the generators of displacements, $\hat{D}_j$, are associated with constants of the motion. And indeed the form of the Hamiltonian for this system

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2m} \tag{6.24}$$

is such that the operators $\hat{p}_j = \hbar\hat{D}_j$ do commute with it.

In the presence of a potential that depends on x, $V = V(x)$, a translation in the x-direction will move the particle into a region of different potential and its time development will be changed, so that p_x will no longer be a constant of the motion, while p_y and p_z will remain so. The symmetry group of the system has been reduced from the full translation group to the *subgroup* of translations in the yz-plane, generated by $\hat{D}_y$ and $\hat{D}_z$. Correspondingly, $\hat{D}_x$, and hence $\hat{p}_x$, does not commute with the Hamiltonian: in fact

$$[\hat{p}_x, \hat{H}] = \hbar[\hat{D}_x, V(x)] = -i\hbar\frac{dV}{dx} \tag{6.25}$$

In the case of a single particle moving in a central potential, $V = V(|\mathbf{r}|)$, a displacement of the centre of force will clearly alter the motion and therefore none of the components of the particle's momentum are conserved. However, if we consider the larger system of a pair of particles moving under the influence of a mutual attraction (or repulsion), $V = V(|\mathbf{r}_1 - \mathbf{r}_2|)$, we can see that a

simultaneous displacement of them both by the *same amount*, $\mathbf{r}_{1,2} \to \mathbf{r}_{1,2} - \mathbf{a}$, will not affect the motion. Denoting the generators of displacements of particles 1 and 2 by $\hat{D}_j^{(1)}$ and $\hat{D}_j^{(2)}$ respectively, the corresponding symmetry transformation will be

$$\begin{aligned} \hat{U}_D(\mathbf{a}) &= \hat{U}_D^{(1)}(\mathbf{a})\,\hat{U}_D^{(2)}(\mathbf{a}) \\ &= \exp\left[-i\mathbf{a}\cdot\hat{\mathbf{D}}^{(1)}\right]\exp\left[-i\mathbf{a}\cdot\hat{\mathbf{D}}^{(2)}\right] \\ &= \exp\left[-i\mathbf{a}\cdot(\hat{\mathbf{D}}^{(1)}+\hat{\mathbf{D}}^{(2)})\right] \end{aligned} \tag{6.26}$$

where the last line follows because displacements of particle 1 and particle 2 commute. Therefore the sum of the generators of displacements of the two particles is associated with a constant of the motion, which (multiplying by $\hbar$) we recognise as the *total momentum* of the system.

6.2 The rotation group, SO(3)

Consider next the group of rotations in three dimensions. We can characterise any rotation by an axis and an angle of rotation around that axis. The effect on the position vector $\mathbf{r}$ of a rotation by angle φ about the z-axis is as follows:

$$\begin{aligned} x &\to x' = x\cos\varphi - y\sin\varphi \\ y &\to y' = x\sin\varphi + y\cos\varphi \\ z &\to z' = z \end{aligned} \tag{6.27}$$

which we can represent by a 3×3 matrix

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \to \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{6.28}$$

That is,

$$\mathbf{r} \to \mathbf{r}' = \hat{R}_z(\varphi)\,\mathbf{r} \tag{6.29}$$

where

$$\hat{R}_z(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{6.30}$$

To find the generator of rotations around the z-axis, we consider an infinitesimal rotation $d\varphi$, for which

$$\hat{R}_z(d\varphi) = \begin{pmatrix} 1 & -d\varphi & 0 \\ d\varphi & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{6.31}$$

i.e.

$$\hat{R}_z(d\varphi) = 1 - id\varphi\, \hat{X}_z \tag{6.32}$$

where

$$\hat{X}_z = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{6.33}$$

Problem 1: Show that

$$\hat{R}_z(\varphi) = \exp\left(-i\varphi\, \hat{X}_z\right) \tag{6.34}$$

where $\hat{R}_z(\varphi)$ is given by Eq. (6.30) and $\hat{X}_z$ by Eq. (6.33).

Hint: Note that $\hat{X}_z^3 = \hat{X}_z$.

Similarly, rotations about the x- and y-axes are represented by the matrices

$$\hat{R}_x(\varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}, \quad \hat{R}_y(\varphi) = \begin{pmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{pmatrix} \tag{6.35}$$

which can be written as

$$\hat{R}_{x,y}(\varphi) = \exp\left(-i\varphi\, \hat{X}_{x,y}\right) \tag{6.36}$$

where

$$\hat{X}_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \hat{X}_y = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \tag{6.37}$$

Since our choice of coordinate axes is arbitrary, it must therefore be the case that a rotation through angle φ about a general axis along the unit vector $\mathbf{n} = (n_x, n_y, n_z)$ is represented by

$$\hat{R}_n(\varphi) = \exp\left[-i\varphi\left(n_x\hat{X}_x + n_y\hat{X}_y + n_z\hat{X}_z\right)\right] = \exp\left(-i\varphi\, \mathbf{n}\cdot\hat{\mathbf{X}}\right) \tag{6.38}$$

which shows that the rotation group is a Lie group of dimension 3. Furthermore, since the generator matrices $\hat{X}_{x,y,z}$ are Hermitian, the rotation matrices $\hat{R}_n(\varphi)$ are unitary. In fact, since $i\hat{X}_{x,y,z}$ and $n_{x,y,z}$ are real, the rotation matrices are real unitary, i.e. *orthogonal* matrices. Finally, since $|\exp \hat{A}| = \exp(\mathrm{Tr}\hat{A})$ for any square matrix (see Problem 6 of Chapter 1) and $\hat{X}_{x,y,z}$ are traceless matrices, $|\hat{R}_n(\varphi)| = 1$. Thus the rotation group is denoted by SO(3), the *special orthogonal group* of dimension 3, the S for special indicating that the determinant is 1.

The rotation group is *non-Abelian*: successive rotations about different axes do not commute. Correspondingly, the generators $\hat{X}_{x,y,z}$ do not commute amongst themselves. By explicit calculation using the above matrices, we find

$$[\hat{X}_x, \hat{X}_y] = i\hat{X}_z \ , \quad [\hat{X}_y, \hat{X}_z] = i\hat{X}_x \ , \quad [\hat{X}_z, \hat{X}_x] = i\hat{X}_y \tag{6.39}$$

that is,

$$[\hat{X}_j, \hat{X}_k] = i\varepsilon_{jkl}\hat{X}_l \tag{6.40}$$

Problem 2: Notice that the elements of the matrices representing the generators of the rotation group are given by

$$\left(\hat{X}_j\right)_{kl} = -i\varepsilon_{jkl} \tag{6.41}$$

Use this fact to derive the commutation relations (6.40).

The maximal subgroup of commuting rotations that we can make are those about a single axis, for example the z-axis:

$$\hat{R}_z(\varphi_1)\hat{R}_z(\varphi_2) = \hat{R}_z(\varphi_2)\hat{R}_z(\varphi_1) = \hat{R}_z(\varphi_1 + \varphi_2) \tag{6.42}$$

Thus the rotation group has rank = 1.

The generators of a Lie group form a *Lie algebra*. An algebra is a set of objects with a rule for multiplication and a rule for addition (whereas a group has only a rule for multiplication). The elements of the Lie algebra are the set of linear combinations of the generators,

$$\hat{A} = \sum_{j=1}^{n} a_j \hat{X}_j \tag{6.43}$$

with real coefficients a_j.

The rule for multiplication in the Lie algebra is to form the commutator and divide by i. If we simply multiplied the operators, or did not divide by i, the result would not in general be a real

linear combination of the generators, i.e. the algebra would not *close*. The properties of the group are determined by the outcome of this multiplication rule. In general we can have

$$[\hat{X}_j, \hat{X}_k] = i \sum_{l=1}^{n} f_{jkl} \hat{X}_j \tag{6.44}$$

where the real constants f_{jkl} are called the *structure constants* of the group. For an Abelian group the structure constants are all zero, while for the rotation group we have $f_{jkl} = \varepsilon_{jkl}$.

6.2.1 Angular momentum conservation

So far, we have discussed the rotation group in terms of its action on the position vector $\mathbf{r}$. Consider now its action on a quantum mechanical wave function $\psi(\mathbf{r}, t)$. We have to find the unitary transformation operator $\hat{U}_R$ such that

$$\psi'(\hat{R}\,\mathbf{r}, t) = \psi(\mathbf{r}, t) \tag{6.45}$$

and therefore

$$\psi'(\mathbf{r}, t) = \hat{U}_R\,\psi(\mathbf{r}, t) = \psi(\hat{R}^{-1}\mathbf{r}, t) \tag{6.46}$$

for a general rotation $\hat{R}$. As usual, we do this by studying the effects of infinitesimal transformations. From Eq. (6.32), an infinitesimal rotation about the z-axis gives

$$\psi'(\mathbf{r}, t) = \psi\left((1 + i\,d\varphi\,\hat{X}_z)\mathbf{r}, t\right) \tag{6.47}$$

and therefore, using Cartesian coordinates,

$$\begin{aligned}
\psi'(x, y, z, t) &= \psi(x + y\,d\varphi, y - x\,d\varphi, z, t) \\
&= \psi(x, y, z, t) - d\varphi \left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\psi(x, y, z, t) \\
&= (1 - \frac{i}{\hbar}\,d\varphi\,\hat{L}_z)\psi
\end{aligned} \tag{6.48}$$

where we recognise $\hat{L}_z$ as the z-component of the orbital angular momentum operator:

$$\hat{L}_z = x\hat{p}_y - y\hat{p}_x = \frac{\hbar}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) \tag{6.49}$$

Thus $\hat{L}_z/\hbar$ is the generator of rotations about the z-axis. Similarly, $\hat{L}_x/\hbar$ and $\hat{L}_y/\hbar$ are the generators of rotations about the x and y axes. Notice that these operators generate the same Lie algebra (6.40) as the matrices $\hat{X}_j$ introduced earlier:

$$[\hat{L}_j, \hat{L}_k] = i\hbar\varepsilon_{jkl}\hat{L}_l \quad \Rightarrow \quad \left[\frac{\hat{L}_j}{\hbar}, \frac{\hat{L}_k}{\hbar}\right] = i\varepsilon_{jkl}\frac{\hat{L}_l}{\hbar} \tag{6.50}$$

We therefore obtain the following fundamental relation between the SO(3) generators and the orbital angular momentum operators:

$$\hat{\mathbf{L}} = \hbar\hat{\mathbf{X}} \tag{6.51}$$

Correspondingly, the transformation operator for the general rotation (6.38), i.e. a finite rotation φ about the unit vector $\mathbf{n}$, is given by

$$\hat{U}_R = \exp\left(-\frac{i}{\hbar}\varphi\,\mathbf{n}\cdot\hat{\mathbf{L}}\right) \tag{6.52}$$

Now we can invoke the general discussion in the previous section to relate the symmetry of a system under rotations to the conservation of the generators of the rotation group, i.e. angular momentum conservation. For example, as we discussed earlier, the motion of a particle in a potential that depends only on x is symmetric with respect to displacements in the yz-plane, and it is also symmetric under rotations in that plane, i.e. rotations about the x-axis. We therefore expect L_x, but not L_y or L_z, to be conserved. If the potential is central, $V = V(|\mathbf{r}|)$, the system is symmetric under all rotations and therefore all components of the angular momentum are conserved.

Problem 3: Consider a pair of particles moving under the influence of a mutual attraction $V = V(|\mathbf{r}_1 - \mathbf{r}_2|)$. Show that the transformation operator for the rotation of the whole system through angle φ about the unit vector $\mathbf{n}$ is given by

$$\hat{U}_R = \exp\left[-\frac{i}{\hbar}\varphi\,\mathbf{n}\cdot(\hat{\mathbf{L}}^{(1)} + \hat{\mathbf{L}}^{(2)})\right] \tag{6.53}$$

where $\hat{\mathbf{L}}^{(1)}$ and $\hat{\mathbf{L}}^{(2)}$ are the angular momentum operators for the two particles separately. Show further that the total angular momentum $\hat{\mathbf{L}} = \hat{\mathbf{L}}^{(1)} + \hat{\mathbf{L}}^{(2)}$ is conserved, whereas $\hat{\mathbf{L}}^{(1)}$ and $\hat{\mathbf{L}}^{(2)}$ separately are not.

6.2.2 Representations of SO(3)

We have seen that the 3×3 matrices $\hat{X}_j$ in Eqs. (6.33) and (6.37) satisfy the same commutation relations as the more abstract operators $\hat{L}_j/\hbar$. We say that these matrices form a *representation* of the Lie algebra generated by the operators. Correspondingly, the rotation matrices $\hat{R}_n$ form a representation of the Lie group SO(3). In fact, they are the smallest matrices that can form a *faithful* representation, i.e. one in which each element is uniquely represented by a different matrix. The smallest faithful representation is called the *fundamental* representation of a group.

However, there are infinitely many other sets of matrices that satisfy the same Lie algebra. A trivial example is to represent every generator by $\hat{0}$, the null matrix, which certainly satisfies the algebra (6.40):

$$[\hat{0}, \hat{0}] = i\varepsilon_{jkl}\hat{0} \tag{6.54}$$

Then every rotation is represented by the identity operator $\hat{I}$. This smallest, but extremely unfaithful, representation is called the *singlet* representation, $\underline{1}$.

Non-trivial representations can be constructed by appealing to the theory of angular momentum presented in the quantum physics course. We note first that the *quadratic Casimir operator*

$$\hat{\mathbf{X}}^2 = \sum_j \hat{X}_j^2 \tag{6.55}$$

commutes with all the generators:

$$[\hat{\mathbf{X}}^2, \hat{X}_k] = 0 \tag{6.56}$$

Problem 4: Verify Eq. (6.56).

Now $\hat{\mathbf{X}}^2 = \hat{\mathbf{L}}^2/\hbar^2$, and we know that the magnitude of the orbital angular momentum is quantised according to $\mathbf{L}^2 = \hbar^2 l(l+1)$ where $l = 0, 2, 3, \ldots$. The states with a given orbital angular momentum are therefore eigenstates of the quadratic Casimir operator $\hat{\mathbf{X}}^2$ with eigenvalue $l(l+1)$. The corresponding wave functions are linear combinations of the spherical harmonics $Y_{lm}(\theta, \varphi)$, where the m signifies that the eigenvalue of $\hat{L}_z$ is $m\hbar$, i.e. the eigenvalue of $\hat{X}_z$ is m. The quantum number m can take $2l+1$ values $m = -l, -l+1, \ldots, l-1, l$. Under rotations, the spherical harmonics with a given value of l and different values of m transform amongst themselves, and therefore the action of rotations on these states can be represented by $(2l+1) \times (2l+1)$ matrices. For example, when $l = 1$ we can represent

$$Y_{11}(\theta, \varphi) = -\sqrt{\frac{3}{8\pi}}\,\sin\theta\, e^{i\varphi} \quad \text{by} \quad \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$Y_{10}(\theta, \varphi) = \sqrt{\frac{3}{4\pi}}\,\cos\theta \quad \text{by} \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$Y_{1-1}(\theta,\varphi) = \sqrt{\frac{3}{8\pi}}\,\sin\theta\, e^{-i\varphi} \quad \text{by} \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{6.57}$$

Then, defining the angular momentum ladder operators $\hat{L}_\pm = \hat{L}_x \pm i\hat{L}_y$, we have

$$\begin{array}{lll} \hat{L}_+ Y_{11} = 0\,, & \hat{L}_+ Y_{10} = \sqrt{2}\hbar Y_{11}\,, & \hat{L}_+ Y_{1-1} = \sqrt{2}\hbar Y_{10} \\ \hat{L}_- Y_{11} = \sqrt{2}\hbar Y_{10}\,, & \hat{L}_- Y_{10} = \sqrt{2}\hbar Y_{1-1}\,, & \hat{L}_- Y_{1-1} = 0 \\ \hat{L}_z Y_{11} = \hbar Y_{11}\,, & \hat{L}_z Y_{10} = 0\,, & \hat{L}_z Y_{1-1} = -\hbar Y_{1-1} \end{array} \tag{6.58}$$

Therefore in this representation the generator matrices $\hat{L}_j/\hbar$ are represented by the matrices $\hat{X}'_j$ where

$$\begin{aligned} \hat{X}'_x &= \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\ \hat{X}'_y &= \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \\ \hat{X}'_z &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \end{aligned} \tag{6.59}$$

This so-called *spherical representation* of the Lie algebra is equivalent to the *Cartesian representation* generated by the matrices $\hat{X}_j$: they are related by a change of basis. They generate the *triplet* representation, $\underline{3}$, of the rotation group.

Problem 5: Suppose $\hat{G}$ is an element of a Lie group with generators $\{\hat{X}_j\}$,

$$\hat{G} = \exp\left(-i\,a_j \hat{X}_j\right)\,, \tag{6.60}$$

and $\hat{U}$ is a unitary operator corresponding to a change of basis,

$$\hat{G} \to \hat{G}' = \hat{U}\hat{G}\hat{U}^\dagger\,. \tag{6.61}$$

Show that

$$\hat{G}' = \exp\left(-i\,a_j \hat{X}'_j\right) \tag{6.62}$$

where

$$\hat{X}'_j = \hat{U}\hat{X}_j\hat{U}^\dagger \ . \tag{6.63}$$

Find the unitary matrix $\hat{U}$ which relates the Cartesian and spherical forms of the triplet representation of SO(3).

Answer:

$$\hat{U} = \frac{e^{i\phi}}{\sqrt{2}} \begin{pmatrix} -1 & i & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & i & 0 \end{pmatrix} \tag{6.64}$$

where ϕ is an arbitrary phase angle.

In general, the $(2l+1) \times (2l+1)$ matrices that represent the action of the angular momentum operators on the states of a system with orbital angular momentum quantum number l generate a representation of SO(3) called the $\underline{2l+1}$ representation. These are the so-called *irreducible representations* of SO(3). The term 'irreducible' refers to the fact that the representation matrices cannot be brought into *block-diagonal* form by any change of basis, whereas a *reducible* representation can be written entirely in that form.

Notice that in the spherical representation we label the $2l+1$ basis states of an irreducible representation of SO(3) according to the eigenvalues m of $\hat{X}_z$. We cannot specify any other quantum numbers because the other generators do not commute with $\hat{X}_z$, i.e. because SO(3) has rank 1. In general, if a group has rank r, the r commuting generators are compatible observables and we can label the states by r quantum numbers which represent their eigenvalues.

Reducible representations of SO(3) are produced when we combine orbital angular momenta. For example, when we combine particles 1 and 2 with $l_1 = l_2 = 1$ we can represent the states as linear combinations of the products of spherical harmonics $Y_{1m_1}(\theta_1, \varphi_1)\, Y_{1m_2}(\theta_1, \varphi_2)$. This makes 9 states $|m_1 m_2\rangle$ with $m_{1,2} = 1$, 0 or –1. The 9×9 matrices representing the action of the angular momentum operators on these states form the $\underline{3}\times\underline{3}$ representation of SO(3). However, we know that by combining $l_1 = 1$ and $l_2 = 1$ we can make total angular momentum $L = 0, 1$ or 2. Each of these separately forms an irreducible representation $\underline{2L+1}$. The angular momentum operators cannot transform any state with a given value of L into one with a different value of L. Therefore the $\underline{3}\times\underline{3}$ representation must be expressible in a block-diagonal form with 1×1, 3×3 and 5×5 matrices along the diagonal, representing the action of rotations on the $L = 0$, 1 and 2 multiplets of states respectively. We express this reduction by writing

$$\underline{3} \times \underline{3} = \underline{1} + \underline{3} + \underline{5} \tag{6.65}$$

Problem 6: Write down the SO(3) reduction formula, corresponding to Eq. (6.65), for the general case of combining angular momenta l_1 and l_2. Verify that the total numbers on the left- and right-hand sides agree.

6.3 The group SU(2)

As its name implies, SU(2) is the group of 2×2 unitary matrices with unit determinant. Since any unitary operator can be written as $\hat{U} = \exp(-i\hat{H})$ where $\hat{H}$ is Hermitian, and since $|\hat{U}| = \exp(-i\,\mathrm{Tr}\,\hat{H})$, we see that the generators of SU(2) must be 2×2 Hermitian, traceless matrices. There are 3 linearly independent 2×2 Hermitian, traceless matrices, which can be taken to be the Pauli matrices

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} , \tag{6.66}$$

Then any element of SU(2) can be written as

$$\hat{U} = \exp(-i\,\alpha_j\,\hat{\sigma}_j) \tag{6.67}$$

where $\{\alpha_j\}$ are real parameters and, as usual, a sum over $j = x, y, z$ is understood.

It is clear from Eq. (6.67) that SU(2) is a Lie group, but at this stage it is a purely mathematical construct. We notice, however that the Lie algebra of the generators $\hat{\sigma}_j$ is

$$[\hat{\sigma}_j, \hat{\sigma}_k] = 2i\varepsilon_{jkl}\,\hat{\sigma}_l \tag{6.68}$$

which is the same as that in Eq. (6.40) if we just make the rescaling $\hat{X}_j = \frac{1}{2}\hat{\sigma}_j$. Therefore the Lie algebras of SU(2) and SO(3) are identical. The two groups are then said to be *homomorphic*.

However, the homomorphism of SU(2) and SO(3) does **not** mean that the two groups are identical. For a start, SU(2) has a representation in terms of 2×2 matrices, which SO(3) does not. The difference lies not in the algebra of the generators of the groups but in the ranges of their parameters.

We can regard SU(2) as the group of rotations of a spin-$\frac{1}{2}$ system such as a single electron. Convenient basis states are those with spin up and down relative to the z-axis:

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \quad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{6.69}$$

Now consider the effect of a rotation around the y-axis:

$$\hat{U} = \exp(-i\,\alpha\,\hat{\sigma}_y) \tag{6.70}$$

Since $\hat{\sigma}_y^2 = \hat{I}$, the unit matrix, we can easily Taylor expand this to obtain

$$\hat{U} = \cos\alpha\,\hat{I} - i\sin\alpha\,\hat{\sigma}_y = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \tag{6.71}$$

Thus the action of this operator on the spin-up state is

$$\hat{U}|\uparrow\rangle = \cos\alpha|\uparrow\rangle + \sin\alpha|\downarrow\rangle \tag{6.72}$$

Now suppose we wish to transform the spin-up state into spin-down. This clearly corresponds to a rotation through angle π around the y-axis. However, we see from Eq. (6.72) that to transform $|\uparrow\rangle$ into $|\downarrow\rangle$ we need $\alpha = \pi/2$, not π. In general, to perform a rotation through angle θ about the y-axis we need $\alpha = \theta/2$. Notice, however, that to get back to $\hat{U}|\uparrow\rangle = |\uparrow\rangle$ we need $\alpha = 2\pi$, i.e. $\theta = 4\pi$. We need to rotate *twice* around the axis of rotation to get back to the state we started from! A rotation through only 2π changes $|\uparrow\rangle$ to $-|\uparrow\rangle$, which has observable consequences, for example in interference experiments involving polarised neutron beams.

More generally, a rotation of the spin through angle φ about an axis along the unit vector $\mathbf{n}$ is represented by the SU(2) matrix

$$\hat{U}_n(\varphi) = \exp\left(-i\frac{\varphi}{2}\,\mathbf{n}\cdot\hat{\sigma}\right) \tag{6.73}$$

where, to cover all the elements of the group, the range of φ must be $0 \le \varphi < 4\pi$. This is equivalent to the expression (6.38) for the group SO(3), with the substitution $\hat{X}_j = \frac{1}{2}\hat{\sigma}_j$ discussed above, except that the range of φ in SO(3) is only $0 \le \varphi < 2\pi$. Thus the group SU(2) has the same Lie algebra as SO(3) but twice as many elements.

Problem 7: Show that

$$\hat{U}_n(\varphi) = \cos(\varphi/2)\hat{I} - i\sin(\varphi/2)\,\mathbf{n}\cdot\hat{\sigma} \tag{6.74}$$

and hence that $\hat{U}_n(2\pi) = -\hat{I}$, independent of the direction of the unit vector $\mathbf{n}$.

6.3.1 Representations of SU(2)

The representations of SU(2) are those that can be formed by combining spin one-half objects. The representations of the spin operators and the SU(2) generators will be related in the same way as

the orbital angular momentum and the SO(3) generators,

$$\hat{S}_j = \hbar \hat{X}_j \tag{6.75}$$

where $\hat{X}_j = \frac{1}{2}\hat{\sigma}_j$ in the fundamental representation. The irreducible representations can then be labelled by the total spin quantum number S, which gives the eigenvalue of the magnitude-squared of the total spin $\mathbf{S}^2 = \hbar^2 S(S+1)$. Equivalently, S gives the eigenvalue $S(S+1)$ of the quadratic Casimir operator $\hat{\mathbf{X}}^2$, which, as in SO(3), commutes with all the generators. The representation will consist of $(2S+1) \times (2S+1)$ matrices which transform the $2S+1$ basis states with a given value of S amongst themselves. The basis states can be labelled by the eigenvalues of $\hat{X}_z$, which range from $-S$ to $+S$ at unit intervals.

Since the total spin quantum number S can be an integer (from combining an even number of spins) or a half-integer (from combining an odd number), there is an irreducible representation $\underline{2S+1}$ corresponding to every positive integer. As for SO(3), the reduction formulae for reducible representations follow from the rules for combining angular momenta. Fore example, combining three spin one-half objects gives the reducible representation $\underline{2} \times \underline{2} \times \underline{2}$. Now

$$\begin{aligned} \underline{2} \times \underline{2} &= \underline{1} + \underline{3} \\ \underline{2} \times \underline{1} &= \underline{2} \\ \underline{2} \times \underline{3} &= \underline{2} + \underline{4} \end{aligned} \tag{6.76}$$

and so

$$\underline{2} \times \underline{2} \times \underline{2} = \underline{2} + \underline{2} + \underline{4} \tag{6.77}$$

corresponding to two distinct spin one-half irreducible representations and a spin three-halves one.

6.4 The group SO(4)

As its name implies, SO(4) is the group of 4×4 real orthogonal matrices with unit determinant. As usual we write the group elements as

$$\hat{G} = \exp\left(-i \sum_{j=1}^{n} a_j \hat{X}_j\right) \tag{6.78}$$

The generators $\{\hat{X}_j\}$ must then be linearly independent 4×4 purely imaginary Hermitian matrices with trace zero. There are six of these, so the dimension of SO(4) is 6. The first three generators

can be chosen to be those of SO(3) or SU(2), filled out with zeros in the fourth row and column:

$$\hat{X}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \hat{X}_2 = \begin{pmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \hat{X}_3 = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{6.79}$$

The remaining generators can then be defined as $\hat{X}_{j+3} = \hat{W}_j$ where

$$\hat{W}_1 = \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, \quad \hat{W}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix}, \quad \hat{W}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{pmatrix} \tag{6.80}$$

We then find that the generators satisfy the following commutation relations (j, k = 1,2 or 3)

$$\begin{aligned} \left[\hat{X}_j, \hat{X}_k\right] &= i\varepsilon_{jkl}\hat{X}_l \\ \left[\hat{X}_j, \hat{W}_k\right] &= i\varepsilon_{jkl}\hat{W}_l \\ \left[\hat{W}_j, \hat{W}_k\right] &= i\varepsilon_{jkl}\hat{X}_l \end{aligned} \tag{6.81}$$

If we define the symmetric and antisymmetric combinations,

$$\hat{X}_j^{\pm} = \frac{1}{2}(\hat{X}_j \pm \hat{W}_j) \tag{6.82}$$

then the commutation relations become simply

$$\begin{aligned} \left[\hat{X}_j^+, \hat{X}_k^+\right] &= i\varepsilon_{jkl}\hat{X}_l^+ \\ \left[\hat{X}_j^-, \hat{X}_k^-\right] &= i\varepsilon_{jkl}\hat{X}_l^- \\ \left[\hat{X}_j^+, \hat{X}_k^-\right] &= 0 \end{aligned} \tag{6.83}$$

Thus the Lie algebra of SO(4) is equivalent to two mutually commuting SU(2) subalgebras. (We shall see shortly that they correspond to SU(2) and not SO(3) subgroups.) Since the maximal commuting subgroup is generated by a single generator from each SU(2), the rank of SO(4) is 2.

6.4.1 Representations of SO(4)

Each of the SU(2) subalgebras of SO(4) can be regarded as applying to a hypothetical 'spin', which we denote by **K** rather than **S** to avoid confusion with the real spin. They are related to the

generators as in Eq. (6.75):

$$\hat{K}_j^{\pm} = \hbar \hat{X}_j^{\pm} \tag{6.84}$$

The eigenvalues of $(\hat{\mathbf{K}}^{\pm})^2$ will then be of the form $\hbar^2 k_{\pm}(k_{\pm}+1)$. The irreducible representations of SO(4) can therefore be labelled by the two quantum numbers k_+ and k_- which are the 'total spin' quantum numbers of the irreducible representations of the two commuting SU(2) subgroups. The corresponding multiplets of states will have $(2k_+ + 1)(2k_- + 1)$ members. For example, the fundamental representation $\underline{4}$ has $k_+ = k_- = \frac{1}{2}$. Notice that if the subgroups were SO(3) instead of SU(2), then k_+ and k_- would have to be integers and the $\underline{4}$ representation could not be formed.

6.4.2 SO(4) symmetry of the hydrogen atom

All of the above may seem purely mathematical, but it turns out that SO(4) symmetry is in fact manifest in the dynamics of the hydrogen atom. The atomic shells of degenerate energy levels in hydrogen correspond to irreducible representations of SO(4). The principal quantum number n is equal to $2k+1$, where k labels the irreducible representation with $k_+ = k_- = k$.

The SO(4) symmetry of the hydrogen atom arises from the existence of a second conserved vector quantity $\mathbf{M}$, with components proportional to the SO(4) generators $\hat{W}_j$, in addition to the conserved angular momentum vector $\mathbf{L}$ with components proportional to the generators $\hat{X}_j$. This quantity is the *Runge-Lenz* vector

$$\mathbf{M} = \frac{1}{m}\mathbf{p} \times \mathbf{L} - \frac{\kappa}{r}\mathbf{r} \tag{6.85}$$

where $\kappa = e^2/4\pi\varepsilon_0$ is the force constant.

In classical mechanics, $\mathbf{M}$ is a vector directed along the major axis of the orbital ellipse, with a magnitude proportional to the eccentricity of the ellipse. Its constancy reflects the fact that the classical orbit is a closed curve, which is a special property of the inverse square force law.

Problem 8: Show that $d\mathbf{M}/dt = 0$ in classical mechanics.

Problem 9: Show further that $\mathbf{L} \cdot \mathbf{M} = 0$ and that

$$\mathbf{M}^2 = \frac{2}{m} H \mathbf{L}^2 + \kappa^2 \tag{6.86}$$

where $H = \mathbf{p}^2/2m - \kappa/r$ is the Hamiltonian.

In quantum mechanics we should represent $\mathbf{M}$ by the Hermitian operator

$$\hat{\mathbf{M}} = \frac{1}{2m}(\hat{\mathbf{p}} \times \hat{\mathbf{L}} - \hat{\mathbf{L}} \times \hat{\mathbf{p}}) - \frac{\kappa}{r}\mathbf{r} \tag{6.87}$$

because $\hat{\mathbf{p}}$ and $\hat{\mathbf{L}}$ do not commute and therefore $\hat{\mathbf{p}} \times \hat{\mathbf{L}}$ is not equal to $-\hat{\mathbf{L}} \times \hat{\mathbf{p}}$. We then find instead of (6.86) that

$$\hat{\mathbf{M}}^2 = \frac{2}{m}\hat{H}\left(\hat{\mathbf{L}}^2 + \hbar^2\right) + \kappa^2 \tag{6.88}$$

After some fairly arduous operator algebra we find the following commutation relations

$$\begin{aligned}
\left[\hat{L}_j, \hat{H}\right] &= \left[\hat{M}_j, \hat{H}\right] = 0 \\
\left[\hat{L}_j, \hat{L}_k\right] &= i\hbar\varepsilon_{jkl}\hat{L}_l \\
\left[\hat{L}_j, \hat{M}_k\right] &= i\hbar\varepsilon_{jkl}\hat{M}_l \\
\left[\hat{M}_j, \hat{M}_k\right] &= -2i\hbar\frac{\hat{H}}{m}\varepsilon_{jkl}\hat{L}_l
\end{aligned} \tag{6.89}$$

Thus $\hat{L}_j$ and $\hat{M}_j$ are constants of the motion and, comparing with Eq. (6.81), we can relate them to the generators of an SO(4) Lie algebra as follows:

$$\hat{L}_j = \hbar\hat{X}_j\ , \quad \hat{M}_j = \hbar\sqrt{\frac{-2\hat{H}}{m}}\hat{W}_j \tag{6.90}$$

or, in terms of the symmetric and antisymmetric combinations (6.82):

$$\hat{L}_j = \hbar\left(\hat{X}_j^+ + \hat{X}_j^-\right)\ , \quad \hat{M}_j = \hbar\sqrt{\frac{-2\hat{H}}{m}}\left(\hat{X}_j^+ - \hat{X}_j^-\right) \tag{6.91}$$

which we may express in terms of the hypothetical spin operators (6.84) as

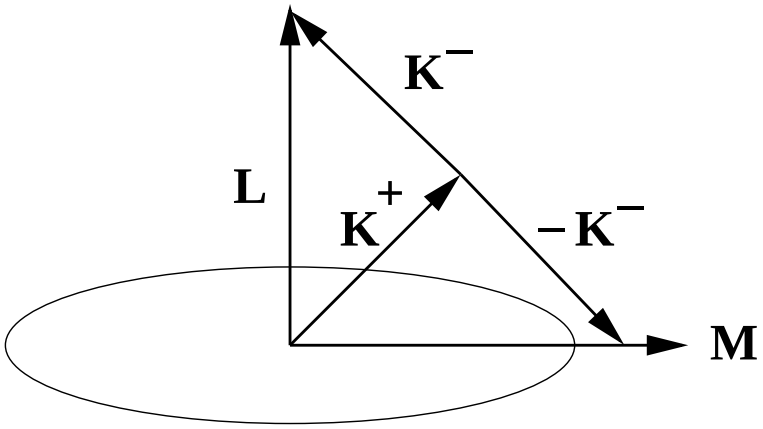
$$\hat{\mathbf{L}} = \hat{\mathbf{K}}^+ + \hat{\mathbf{K}}^-\ , \quad \hat{\mathbf{M}} = \sqrt{\frac{-2\hat{H}}{m}}\left(\hat{\mathbf{K}}^+ - \hat{\mathbf{K}}^-\right) \tag{6.92}$$

Therefore we find

$$\hat{\mathbf{L}} \cdot \hat{\mathbf{M}} = \sqrt{\frac{-2\hat{H}}{m}}\left[\left(\hat{\mathbf{K}}^+\right)^2 - \left(\hat{\mathbf{K}}^-\right)^2\right] \tag{6.93}$$

Now we recall that the irreducible representations of SO(4) are eigenstates of $\left(\hat{\mathbf{K}}^\pm\right)^2$ with eigenvalues $\hbar^2 k_\pm(k_\pm + 1)$. Therefore the eigenvalues of $\hat{\mathbf{L}} \cdot \hat{\mathbf{M}}$ can only be zero, in accordance with Problem 9, if we have $k_+ = k_-$ for the states of the hydrogen atom.

We can picture the composition of **L** and **M** according to Eq. (6.92) as shown below.



To compute the energy levels of the hydrogen atom, we note from Eqs. (6.88) and (6.92) that

$$\frac{2\hat{H}}{m}\hat{\mathbf{L}}^2 - \hat{\mathbf{M}}^2 = -\frac{2\hbar^2}{m}\hat{H} - \kappa^2 = \frac{4\hat{H}}{m}\left[\left(\hat{\mathbf{K}}^+\right)^2 + \left(\hat{\mathbf{K}}^-\right)^2\right] \tag{6.94}$$

which implies that the eigenvalues of $\hat{H}$ are E where

$$\hbar^2\frac{4E}{m}\left[k_+(k_+ + 1) + k_-(k_- + 1) + \frac{1}{2}\right] = -\kappa^2 \tag{6.95}$$

Hence, setting $k_+ = k_- = k$,

$$E = -\frac{m\kappa^2}{2\hbar^2(2k+1)^2} \tag{6.96}$$

where $k = 0, \frac{1}{2}, 1, \ldots$. Setting $k = (n-1)/2$ with $n = 1, 2, \ldots$, this does indeed give the familiar formula for the energy levels of hydrogen,

$$E_n = -\frac{m\kappa^2}{2\hbar^2 n^2} \tag{6.97}$$

where n is the principal quantum number.

Problem 10: Show that number of states in the shell with $k_+ = k_- = (n-1)/2$ agrees with the usual counting of angular momentum states with $l < n$.
