# Theoretical Physics 2 <br> Lecture Notes and Examples 

B R Webber and N R Cooper

Lent Term 2004

## Preface

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

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

## Course Outline

- Operator Methods in Quantum Mechanics (2 lectures): Mathematical foundations of non-relativistic quantum mechanics; vector spaces; operator methods for discrete and continuous eigenspectra; generalized form of the uncertainty principle; simple harmonic oscillator; delta-function potential; introduction to second quantization.
- Angular Momentum (2 lectures): Eigenvalues/eigenvectors of the angular momentum operators (orbital/spin); spherical harmonics and their applications; Pauli matrices and spinors; addition of angular momenta.
- Approximation Methods for Bound States (2 lectures): Variational methods and their application to problems of interest; perturbation theory (time-independent and time depen-
dent) including degenerate and non-degenerate cases; the JWKB method and its application to barrier penetration and radioactive decay.
- Scattering Theory (2 lectures): Scattering amplitudes and differential cross-section; partial wave analysis; the optical theorem; Green functions; weak scattering and the Born approximation; *relation between Born approximation and partial wave expansions; *beyond the Born approximation.
- Identical Particles in Quantum Mechanics (2 lectures): Wave functions for noninteracting systems; symmetry of many-particle wave functions; the Pauli exclusion principle; fermions and bosons; exchange forces; the hydrogen molecule; scattering of identical particles; *second quantization method for many-particle systems; *pair correlation functions for bosons and fermions;
- Density Matrices (2 lectures): Pure and mixed states; the density operator and its properties; position and momentum representation of the density operator; applications in statistical mechanics.


## Problem Sets

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

## Books

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

At about the level of the course: Liboff, Quantum Mechanics, $3^{\text {rd }}$ Ed., Addison-Wesley.
At a more advanced level: Sakurai, Quantum Mechanics, $2^{\text {nd }}$ Ed., Addison-Wesley; Reed and Simon, Methods of Modern Mathematical Physics, Academic Press.

## Contents

1 Operator Methods In Quantum Mechanics ..... 1
1.1 Introduction ..... 1
1.1.1 Mathematical foundations ..... 2
1.1.2 Hilbert space ..... 3
1.1.3 The Schwartz inequality ..... 4
1.1.4 Some properties of vectors in a Hilbert space ..... 5
1.1.5 Orthonormal systems ..... 5
1.1.6 Operators on Hilbert space ..... 6
1.1.7 Eigenvectors and eigenvalues ..... 10
1.1.8 Observables ..... 15
1.1.9 Generalised uncertainty principle ..... 16
1.1.10 Basis transformations ..... 18
1.1.11 Matrix representation of operators ..... 19
1.1.12 Mathematical interlude: Dirac delta function ..... 20
1.1.13 Operators with continuous or mixed (discrete-continuous) spectra ..... 21
1.2 Applications ..... 25
1.2.1 Harmonic oscillator ..... 25
1.2.2 Delta-function potential well ..... 31
1.3 Introduction to second quantisation ..... 34
1.3.1 Vibrating string ..... 34
1.3.2 Quantisation of vibrating string ..... 37
1.3.3 General second quantisation procedure ..... 38
2 Angular Momentum ..... 41
2.1 Introduction ..... 41
2.2 Orbital angular momentum ..... 41
2.2.1 Eigenvalues of orbital angular momentum ..... 47
2.2.2 Eigenfunctions of orbital angular momentum ..... 50
2.2.3 Mathematical interlude: Legendre polynomials and spherical harmonics ..... 53
2.2.4 Angular momentum and rotational invariance ..... 56
2.3 Spin angular momentum ..... 58
2.3.1 Spinors ..... 62
2.4 Addition of angular momenta ..... 64
2.4.1 Addition of spin- $\frac{1}{2}$ operators ..... 65
2.4.2 Addition of spin- $\frac{1}{2}$ and orbital angular momentum ..... 67
2.4.3 General case ..... 69
3 Approximation Methods For Bound States ..... 71
3.1 Introduction ..... 71
3.2 Variational methods ..... 71
3.2.1 Variational theorem ..... 72
3.2.2 Interlude : atomic units ..... 74
3.2.3 Hydrogen molecular ion, $\mathrm{H}_{2}^{+}$ ..... 75
3.2.4 Generalisation: Ritz theorem ..... 78
3.2.5 Linear variation functions ..... 80
3.3 Perturbation methods ..... 83
3.3.1 Time-independent perturbation theory ..... 83
3.3.2 Time-dependent perturbation theory ..... 90
3.4 JWKB method ..... 96
3.4.1 Derivation ..... 97
3.4.2 Connection formulae ..... 99
3.4.3 *JWKB treatment of the bound state problem ..... 101
3.4.4 Barrier penetration ..... 103
3.4.5 Alpha decay of nuclei ..... 105
4 Scattering Theory ..... 109
4.1 Introduction ..... 109
4.2 Spherically symmetric square well ..... 109
4.3 Mathematical interlude ..... 111
4.3.1 Brief review of complex analysis ..... 111
4.3.2 Properties of spherical Bessel/Neumann functions ..... 113
4.3.3 Expansion of plane waves in spherical harmonics ..... 115
4.4 The quantum mechanical scattering problem ..... 116
4.4.1 Born approximation ..... 120
4.5 *Formal time-independent scattering theory ..... 126
4.5.1 *Lippmann-Schwinger equation in the position representation ..... 127
4.5.2 *Born again! ..... 128
5 Identical Particles in Quantum Mechanics ..... 131
5.1 Introduction ..... 131
5.2 Multi-particle systems ..... 131
5.2.1 Pauli exclusion principle ..... 133
5.2.2 Representation of $\Psi(1,2, \ldots, N)$ ..... 134
5.2.3 Neglecting the symmetry of the many-body wave function ..... 134
5.3 Fermions ..... 135
5.4 Bosons ..... 135
5.5 Exchange forces ..... 136
5.6 Helium atom ..... 138
5.6.1 Ground state ..... 139
5.7 Hydrogen molecule ..... 142
5.8 Scattering of identical particles ..... 148
5.8.1 Scattering of identical spin zero bosons ..... 149
5.8.2 Scattering of fermions ..... 149
5.9 *Modern electronic structure theory ..... 150
5.9.1 *The many-electron problem ..... 150
5.9.2 *One-electron methods ..... 151
5.9.3 *Hartree approximation ..... 151
5.9.4 *Hartree-Fock approximation ..... 152
5.9.5 *Density functional methods ..... 154
5.9.6 *Shortcomings of the mean-field approach ..... 156
5.9.7 *Quantum Monte Carlo methods ..... 157
6 Density Operators ..... 159
6.1 Introduction ..... 159
6.2 Pure and mixed states ..... 160
6.3 Properties of the Density Operator ..... 161
6.3.1 Density operator for spin states ..... 164
6.3.2 Density operator in the position representation ..... 166
6.4 Density operator in statistical mechanics ..... 168
6.4.1 Density operator for a free particle in the momentum representation ..... 170
6.4.2 Density operator for a free particle in the position representation ..... 171
6.4.3 *Density matrix for the harmonic oscillator ..... 172

## Chapter 1

## Operator Methods In Quantum Mechanics

### 1.1 Introduction

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

- Postulate 1: The state of a quantum-mechanical system is completely specified by a function $\Psi(\mathbf{r}, t)$ (which in general can be complex) that depends on the coordinates of the particles (collectively denoted by $\mathbf{r}$ ) and on the time. This function, called the wave function or the state function, has the important property that $\Psi^{*}(\mathbf{r}, t) \Psi(\mathbf{r}, t) d \mathbf{r}$ is the probability that the system will be found in the volume element $d \mathbf{r}$, located at $\mathbf{r}$, at the time $t$.
- Postulate 2: To every observable $A$ in classical mechanics, there corresponds a linear Hermitian operator $\hat{A}$ in quantum mechanics.
- Postulate 3: In any measurement of the observable $A$, the only values that can be obtained are the eigenvalues $\{a\}$ of the associated operator $\hat{A}$, which satisfy the eigenvalue equation

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

where $\Psi_{a}$ 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 $\Psi$, and the eigenfunctions $\left\{\Psi_{a}\right\}$ of $\hat{A}$ are also normalised, then the probability of obtaining the value $a$ in a measurement of the observable $A$ is given by

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

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

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

is said to be normalised.)

- Postulate 5: As a result of a measurement of the observable $A$ in which the value $a$ is obtained, the wave function of the system becomes the corresponding eigenfunction $\Psi_{a}$. (This is sometimes called the collapse of the wave function.)
- Postulate 6: Between measurements, the wave function evolves in time according to the time-dependent Schrödinger equation

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

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

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

We now explore the mathematical structure underpinning quantum mechanics.

### 1.1.1 Mathematical foundations

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

### 1.1.2 Hilbert space

A Hilbert space $H$,

$$
\begin{equation*}
H=\{|a\rangle,|b\rangle,|c\rangle, \ldots\}, \tag{1.1}
\end{equation*}
$$

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, $\mid$ null $\rangle \in H$ with the property

$$
\begin{equation*}
|a\rangle+\mid \text { null }\rangle=|a\rangle \tag{1.2}
\end{equation*}
$$

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

$$
\begin{equation*}
|a\rangle+|-a\rangle=\mid \text { null }\rangle \tag{1.3}
\end{equation*}
$$

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

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

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

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

The last equation can also be written as

$$
\begin{equation*}
\langle a \mid b\rangle=\langle b \mid a\rangle^{*} \tag{1.11}
\end{equation*}
$$

From the above, we can deduce that

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

and

$$
\begin{align*}
\left(\left|a_{1}\right\rangle+\left|a_{2}\right\rangle,|b\rangle\right) & =\left(\left|a_{1}\right\rangle,|b\rangle\right)+\left(\left|a_{2}\right\rangle,|b\rangle\right)  \tag{1.14}\\
& =\left\langle a_{1} \mid b\right\rangle+\left\langle a_{2} \mid b\right\rangle \tag{1.15}
\end{align*}
$$

The norm of a vector is defined by

$$
\begin{equation*}
\|a\|=\sqrt{\langle a \mid a\rangle} \tag{1.16}
\end{equation*}
$$

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

### 1.1.3 The Schwartz inequality

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

$$
\begin{equation*}
\|a\| \quad\|b\| \geq|\langle a \mid b\rangle| \tag{1.17}
\end{equation*}
$$

with the equality only being valid for the case

$$
\begin{equation*}
|a\rangle=\lambda|b\rangle \tag{1.18}
\end{equation*}
$$

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

## Proof

Define a $|c\rangle$ such that

$$
\begin{equation*}
|c\rangle=|a\rangle+\lambda|b\rangle \tag{1.19}
\end{equation*}
$$

where $\lambda$ is an arbitrary complex number. Whatever $\lambda$ may be:

$$
\begin{align*}
\langle c \mid c\rangle & =\langle a \mid a\rangle+\lambda\langle a \mid b\rangle+\lambda^{*}\langle b \mid a\rangle+\lambda \lambda^{*}\langle b \mid b\rangle  \tag{1.20}\\
& \geq 0 \tag{1.21}
\end{align*}
$$

Choose for $\lambda$ the value

$$
\begin{equation*}
\lambda=-\frac{\langle b \mid a\rangle}{\langle b \mid b\rangle} \tag{1.22}
\end{equation*}
$$

and substitute into the above equation, which reduces to

$$
\begin{equation*}
\langle a \mid a\rangle-\frac{\langle a \mid b\rangle\langle b \mid a\rangle}{\langle b \mid b\rangle} \geq 0 \tag{1.23}
\end{equation*}
$$

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

$$
\begin{align*}
\langle a \mid a\rangle\langle b \mid b\rangle & \geq\langle a \mid b\rangle\langle b \mid a\rangle  \tag{1.24}\\
& \geq|\langle a \mid b\rangle|^{2} \tag{1.25}
\end{align*}
$$

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

### 1.1.4 Some properties of vectors in a Hilbert space

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

$$
\begin{equation*}
\||a\rangle-\left|a_{n}\right\rangle \| \leq \epsilon \tag{1.26}
\end{equation*}
$$

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 $\left\{\left|a_{n}\right\rangle\right\}$, in the sense that

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \||a\rangle-\left|a_{n}\right\rangle \|=0 \tag{1.27}
\end{equation*}
$$

Then the sequence $\left\{\left|a_{n}\right\rangle\right\}$ has a unique limiting value $|a\rangle$.
The above properties are necessary for vector spaces of infinite dimension that occur in QM.

### 1.1.5 Orthonormal systems

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

$$
\begin{equation*}
\langle a \mid b\rangle=0 \tag{1.28}
\end{equation*}
$$

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

$$
\begin{equation*}
\left\langle a_{n} \mid a_{m}\right\rangle=\delta_{n, m} \tag{1.29}
\end{equation*}
$$

where

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

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

$$
\begin{equation*}
|a\rangle=\sum_{n} \alpha_{n}\left|a_{n}\right\rangle \tag{1.30}
\end{equation*}
$$

where in general $\alpha_{n}$ are complex numbers whose values are

$$
\begin{equation*}
\alpha_{m}=\left\langle a_{m} \mid a\right\rangle \tag{1.31}
\end{equation*}
$$

Proof

$$
\begin{align*}
\left\langle a_{m} \mid a\right\rangle & =\left\langle a_{m}\right|\left(\sum_{n} \alpha_{n}\left|a_{n}\right\rangle\right) \\
& =\sum_{n} \alpha_{n}\left\langle a_{m} \mid a_{n}\right\rangle \\
& =\sum_{n} \alpha_{n} \delta_{m, n} \\
& =\alpha_{m} \tag{1.32}
\end{align*}
$$

Thus we can write

$$
\begin{equation*}
|a\rangle=\sum_{n}\left|a_{n}\right\rangle\left\langle a_{n} \mid a\right\rangle \tag{1.33}
\end{equation*}
$$

Note that this implies

$$
\begin{equation*}
\hat{I}=\sum_{n}\left|a_{n}\right\rangle\left\langle a_{n}\right| \tag{1.34}
\end{equation*}
$$

called the "resolution of the identity operator" or the closure relation. The complex numbers $\alpha_{n}$ are called the $a_{n}$-representation of $|a\rangle$, i.e. they are the components of the vector $|a\rangle$ in the basis $\left\{\left|a_{n}\right\rangle\right\}$.

### 1.1.6 Operators on Hilbert space

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

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

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

$$
\begin{equation*}
\| \hat{A}|a\rangle\|\leq C\||a\rangle \| \tag{1.36}
\end{equation*}
$$

$\forall|a\rangle \in H$, and $C$ is a real positive constant $(<\infty)$.
Bounded linear operators are continuous, i.e. if

$$
\begin{equation*}
\left|a_{n}\right\rangle \rightarrow|a\rangle \tag{1.37}
\end{equation*}
$$

then it follows that

$$
\begin{equation*}
\hat{A}\left|a_{n}\right\rangle \rightarrow \hat{A}|a\rangle \tag{1.38}
\end{equation*}
$$

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

$$
\begin{equation*}
\hat{A}|a\rangle=\hat{B}|a\rangle \tag{1.39}
\end{equation*}
$$

The following definitions are valid $\forall|a\rangle \in H$ :
Unit operator, $\hat{I}$

$$
\begin{equation*}
\hat{I}|a\rangle=|a\rangle \tag{1.40}
\end{equation*}
$$

Zero operator, $\hat{0}$

$$
\begin{equation*}
\hat{0}|a\rangle=\mid \text { null }\rangle \tag{1.41}
\end{equation*}
$$

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

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

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

$$
\begin{equation*}
(\hat{A} \hat{B})|a\rangle=\hat{A}(\hat{B}|a\rangle) \tag{1.43}
\end{equation*}
$$

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

$$
\begin{equation*}
(|b\rangle, \hat{A}|a\rangle)=\left(\hat{A}^{\dagger}|b\rangle,|a\rangle\right) \tag{1.44}
\end{equation*}
$$

or

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

The adjoint of an operator has the following properties:

$$
\begin{align*}
(\alpha \hat{A})^{\dagger} & =\alpha^{*} \hat{A}^{\dagger}  \tag{1.46}\\
(\hat{A}+\hat{B})^{\dagger} & =\hat{A}^{\dagger}+\hat{B}^{\dagger}  \tag{1.47}\\
(\hat{A} \hat{B})^{\dagger} & =\hat{B}^{\dagger} \hat{A}^{\dagger}  \tag{1.48}\\
\left(\hat{A}^{\dagger}\right)^{\dagger} & =\hat{A} \tag{1.49}
\end{align*}
$$

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

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

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

$$
\begin{equation*}
\hat{U} \hat{U}^{\dagger}=\hat{U}^{\dagger} \hat{U}=\hat{I} \tag{1.51}
\end{equation*}
$$

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 $|a\rangle$

$$
\begin{equation*}
\hat{P}|b\rangle=\langle a \mid b\rangle|a\rangle=|a\rangle\langle a \mid b\rangle \tag{1.52}
\end{equation*}
$$

We write this symbolically as

$$
\begin{equation*}
\hat{P}=|a\rangle\langle a| \tag{1.53}
\end{equation*}
$$

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

$$
\begin{equation*}
\hat{P}^{2}=|a\rangle\langle a \mid a\rangle\langle a|=|a\rangle\langle a| \tag{1.54}
\end{equation*}
$$

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

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

$$
\begin{equation*}
[\hat{A}, \hat{B}]=\hat{A} \hat{B}-\hat{B} \hat{A} \tag{1.55}
\end{equation*}
$$

Note that in general

$$
\begin{equation*}
\hat{A} \hat{B} \neq \hat{B} \hat{A} \tag{1.56}
\end{equation*}
$$

Properties of commutators:

$$
\begin{gather*}
{[\hat{A}, \hat{B}]=-[\hat{B}, \hat{A}]}  \tag{1.57}\\
{[\hat{A},(\hat{B}+\hat{C})]=[\hat{A}, \hat{B}]+[\hat{A}, \hat{C}]}  \tag{1.58}\\
{[\hat{A}, \hat{B} \hat{C}]=[\hat{A}, \hat{B}] \hat{C}+\hat{B}[\hat{A}, \hat{C}]}  \tag{1.59}\\
{[\hat{A},[\hat{B}, \hat{C}]]+[\hat{B},[\hat{C}, \hat{A}]]+[\hat{C},[\hat{A}, \hat{B}]]=\hat{0}}  \tag{1.60}\\
{[\hat{A}, \hat{B}]^{\dagger}=\left[\hat{B}^{\dagger}, \hat{A}^{\dagger}\right]} \tag{1.61}
\end{gather*}
$$

## 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 $\left[\hat{P}, \hat{Q}^{n}\right]$ to its simplest possible form.

Answer: Let

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

Then $\hat{R}_{1}=[\hat{P}, \hat{Q}]=a \hat{I}$ and

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

(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 $\hat{R}_{2}=2 a \hat{Q}, \hat{R}_{3}=3 a \hat{Q}^{2}$ etc. This implies that

$$
\hat{R}_{n}=\left[\hat{P}, \hat{Q}^{n}\right]=n a \hat{Q}^{n-1}
$$

Note that in general,

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

- Reduce the commutator

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

to its simplest form.
Answer: Use results above to get

$$
\left[\hat{P}, e^{i \hat{Q}}\right]=i a e^{i \hat{Q}}
$$

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

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

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

Answer: $-2 \hat{I}$

- Find the commutator $[\hat{A}, \hat{B}]$

Answer: $-2 \hat{B}$

### 1.1.7 Eigenvectors and eigenvalues

If

$$
\begin{equation*}
\hat{A}|a\rangle=a|a\rangle \tag{1.63}
\end{equation*}
$$

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

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

$$
\begin{align*}
\hat{A}|a\rangle & =a|a\rangle  \tag{1.64}\\
\hat{A}\left|a^{\prime}\right\rangle & =a^{\prime}\left|a^{\prime}\right\rangle \tag{1.65}
\end{align*}
$$

and if $a \neq a^{\prime}$, then

$$
\begin{equation*}
\left\langle a \mid a^{\prime}\right\rangle=0 \tag{1.66}
\end{equation*}
$$

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

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

Answer: To be discussed in class.

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

$$
\begin{equation*}
|\psi\rangle=\sum_{a}|a\rangle\langle a \mid \psi\rangle \tag{1.67}
\end{equation*}
$$

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

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

$$
\begin{array}{r}
\hat{Q}^{\dagger} \hat{Q}^{\dagger}=0 \\
\hat{Q} \hat{Q}^{\dagger}+\hat{Q}^{\dagger} \hat{Q}=\hat{I} \tag{1.68}
\end{array}
$$

The Hamiltonian for the system is given by

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

where $\alpha$ is a real constant.

- Show that $\hat{H}$ is self-adjoint
- Find an expression for $\hat{H}^{2}$ in terms of $\hat{H}$ Answer: Use the anti-commutator property of $\hat{Q}$ to get $\hat{H}^{2}=\alpha \hat{H}$.
- Deduce the eigenvalues of $\hat{H}$ using the results obtained above. Answer:The eigenvalues are 0 and $\alpha$.


## Problem 4 : Manipulating Operators

- Show that if $|a\rangle$ is an eigenvector of $\hat{A}$ with eigenvalue $a$, then it is an eigenvector of $f(\hat{A})$ with eigenvalue $f(a)$.
- Show that

$$
\begin{equation*}
(\hat{A} \hat{B})^{\dagger}=\hat{B}^{\dagger} \hat{A}^{\dagger} \tag{1.69}
\end{equation*}
$$

and in general

$$
\begin{equation*}
(\hat{A} \hat{B} \hat{C} \ldots)^{\dagger}=\ldots \hat{C}^{\dagger} \hat{B}^{\dagger} \hat{A}^{\dagger} \tag{1.70}
\end{equation*}
$$

- Show that $\hat{A} \hat{A}^{\dagger}$ is Hermitian even if $\hat{A}$ is not.
- Show that if $\hat{A}$ is Hermitian, then the expectation value of $\hat{A}^{2}$ are non-negative, and the eigenvalues of $\hat{A}^{2}$ are non-negative.
- Suppose there exists a linear operator $\hat{A}$ that has an eigenvector $|\psi\rangle$ with eigenvalue $a$. If there also exists an operator $\hat{B}$ such that

$$
\begin{equation*}
[\hat{A}, \hat{B}]=\hat{B}+2 \hat{B} \hat{A}^{2} \tag{1.71}
\end{equation*}
$$

then show that $\hat{B}|\psi\rangle$ is an eigenvector of $\hat{A}$ and find the eigenvalue.
Answer: Eigenvalue is $1+a+2 a^{2}$.

## 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 $\left[\hat{A}, \hat{B}^{n}\right]=n \hat{B}^{n-1}[\hat{A}, \hat{B}]$ and $\left[\hat{A}^{n}, \hat{B}\right]=n \hat{A}^{n-1}[\hat{A}, \hat{B}]$.

Answer: To show this, consider the following steps:

$$
\begin{align*}
{\left[\hat{A}, \hat{B}^{n}\right] } & =\hat{A} \hat{B}^{n}-\hat{B}^{n} \hat{A}  \tag{1.72}\\
& =\hat{A} \hat{B} \hat{B}^{n-1}-\hat{B} \hat{A} \hat{B}^{n-1}+\hat{B}(\hat{A} \hat{B}) \hat{B}^{n-2}-\hat{B}(\hat{B} \hat{A}) \hat{B}^{n-3}+\cdots \hat{B}^{n-1} \hat{A} \hat{B}-\hat{B}^{n-1} \hat{B} \hat{A} \\
& =[\hat{A}, \hat{B}] \hat{B}^{n-1}+\hat{B}[\hat{A}, \hat{B}] \hat{B}^{n-2}+\cdots+\hat{B}^{n-1}[\hat{A}, \hat{B}]
\end{align*}
$$

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

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

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

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

as required.

- (b) Just as in (a), show that for any analytic function, $f(x)$, we have $[\hat{A}, f(\hat{B})]=[\hat{A}, \hat{B}] f^{\prime}(\hat{B})$, where $f^{\prime}(x)$ 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

$$
\begin{align*}
{[\hat{A}, f(\hat{B})] } & =\left[\hat{A}, \sum_{n} a_{n} \hat{B}^{n}\right]  \tag{1.73}\\
& =\sum_{n} a_{n}\left[\hat{A}, \hat{B}^{n}\right] \\
& =[\hat{A}, \hat{B}] \sum_{n} n a_{n} \hat{B}^{n-1} \\
& =[\hat{A}, \hat{B}] f^{\prime}(\hat{B})
\end{align*}
$$

- (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:

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

Using part (a), we can write

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

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

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

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 $\hat{B}$ :

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

Answer: To show this, define

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

where $\lambda$ is a real parameter. Then,

$$
\begin{align*}
f(0) & =\hat{B}  \tag{1.76}\\
f(1) & =e^{\hat{A}} \hat{B} e^{-\hat{A}} \\
f^{\prime}(\lambda) & =e^{\lambda \hat{A}}[\hat{A}, \hat{B}] e^{-\lambda \hat{A}} \\
f^{\prime}(0) & =[\hat{A}, \hat{B}] \\
f^{\prime \prime}(\lambda) & =e^{\lambda \hat{A}}[\hat{A},[\hat{A}, \hat{B}]] e^{-\lambda \hat{A}} \\
f^{\prime \prime}(0) & =[\hat{A},[\hat{A}, \hat{B}]]
\end{align*}
$$

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

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

This implies

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

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

### 1.1.8 Observables

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

$$
\begin{equation*}
|\psi\rangle=\sum_{n=1}^{D}\left|\psi_{n}\right\rangle\left\langle\psi_{n} \mid \psi\right\rangle \tag{1.77}
\end{equation*}
$$

(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 $\left|\psi_{n}\right\rangle$. 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

$$
\begin{equation*}
P\left(a_{n}\right)=\left|\left\langle\psi_{n} \mid \psi\right\rangle\right|^{2} \tag{1.78}
\end{equation*}
$$

(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

$$
\begin{equation*}
\sum_{n=1}^{D} P\left(a_{n}\right)=1 \tag{1.79}
\end{equation*}
$$

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

$$
\begin{equation*}
\hat{A}=\sum_{n} a_{n}\left|\psi_{n}\right\rangle\left\langle\psi_{n}\right| \tag{1.80}
\end{equation*}
$$

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

$$
\begin{align*}
\langle\hat{A}\rangle & =\sum_{n} a_{n} P\left(a_{n}\right)=\sum_{n} a_{n}\left|\left\langle\psi_{n} \mid \psi\right\rangle\right|^{2} \\
& =\sum_{n} a_{n}\left\langle\psi \mid \psi_{n}\right\rangle\left\langle\psi_{n} \mid \psi\right\rangle=\langle\psi| \hat{A}|\psi\rangle \tag{1.81}
\end{align*}
$$

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

Problem 5: A system described by the Hamiltonian $\hat{H}_{0}$ has just two orthogonal energy eigenstates, $|1\rangle$ and $|2\rangle$ with

$$
\begin{align*}
& \langle 1 \mid 1\rangle=1 \\
& \langle 1 \mid 2\rangle=0 \\
& \langle 2 \mid 2\rangle=1 \tag{1.82}
\end{align*}
$$

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

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

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

Answer: This will be done in class.

### 1.1.9 Generalised uncertainty principle

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

$$
\begin{equation*}
[\hat{A}, \hat{B}]=i \hat{C} \tag{1.84}
\end{equation*}
$$

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

$$
\begin{equation*}
\Delta A \Delta B \geq \frac{1}{2}|\langle\hat{C}\rangle| \tag{1.85}
\end{equation*}
$$

where

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

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

Problem 6: Prove (1.85).

If the eigenvalues of $\hat{A}$ are non-degenerate, the normalised eigenvectors $\left|\psi_{n}\right\rangle$ are unique to within a phase factor i.e. the kets $\left|\psi_{n}\right\rangle$ and $e^{i \theta}\left|\psi_{n}\right\rangle$, where $\theta$ 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 $\left|a_{n}, b_{p}, c_{q}, \ldots\right\rangle$ are unique to within a phase factor; they are a basis for which the orthonormality relations are

$$
\begin{equation*}
\left\langle a_{n^{\prime}}, b_{p^{\prime}}, c_{q^{\prime}}, \ldots \mid a_{n}, b_{p}, c_{q}, \ldots\right\rangle=\delta_{n^{\prime} n} \delta_{p^{\prime} p} \delta_{q^{\prime} q} \cdots \tag{1.87}
\end{equation*}
$$

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

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

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


### 1.1.10 Basis transformations

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

$$
\begin{equation*}
\left|\phi_{n}\right\rangle=\hat{U}\left|\psi_{n}\right\rangle \tag{1.88}
\end{equation*}
$$

where

$$
\begin{equation*}
\hat{U}=\sum_{i}\left|\phi_{i}\right\rangle\left\langle\psi_{i}\right| \tag{1.89}
\end{equation*}
$$

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

## Problem 8:

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

$$
\begin{equation*}
\hat{A}\left|\psi_{n}\right\rangle=a_{n}\left|\psi_{n}\right\rangle \tag{1.90}
\end{equation*}
$$

show that the operator

$$
\begin{equation*}
\hat{A}^{\prime}=\hat{U} \hat{A} \hat{U}^{\dagger} \tag{1.91}
\end{equation*}
$$

has $\hat{U}\left|\psi_{n}\right\rangle$ as its eigenvector with eigenvalue $a_{n}$.

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

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

### 1.1.11 Matrix representation of operators

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

$$
\begin{equation*}
\hat{A}=\hat{I} \hat{A} \hat{I}=\sum_{n} \sum_{n^{\prime}}\left|n^{\prime}\right\rangle\left\langle n^{\prime}\right| \hat{A}|n\rangle\langle n| \tag{1.92}
\end{equation*}
$$

where the set $\{|n\rangle\}$ are a set of basis vectors in the Hilbert space and the complex numbers $\left\langle n^{\prime}\right| \hat{A}|n\rangle$ 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 | $A=A^{T}$ | $A_{p q}=A_{q p}$ |
| Antisymmetric | $A=-A^{T}$ | $A_{p q}=-A_{q p}$ |
| Orthogonal | $A=\left(A^{T}\right)^{-1}$ | $\left(A^{T} A\right)_{p q}=\delta_{p q}$ |
| Real | $A=A^{*}$ | $A_{p q}=A_{p q}^{*}$ |
| Pure Imaginary | $A=-A^{*}$ | $A_{p q}=-A_{p q}^{*}$ |
| Hermitian | $A=A^{\dagger}$ | $A_{p q}=A_{q p}^{*}$ |
| Anti-Hermitian | $A=-A^{\dagger}$ | $A_{p q}=-A_{q p}^{*}$ |
| Unitary | $A=\left(A^{\dagger}\right)^{-1}$ | $\left(A^{\dagger} A\right)_{p q}=\delta_{p q}$ |
| Singular | $\|A\|=0$ |  |

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

## Problem 9:

- If $A, B, C$ are $3 n \times n$ square matrices, show that $\operatorname{Tr}(A B C)=\operatorname{Tr}(C A B)=\operatorname{Tr}(B C A)$, where $\operatorname{Tr}$ 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 $U=(A+i I)(A-i I)^{-1}$ is unitary. ( $I$ here is the identity matrix.)
- Show that $|I+\epsilon A|=I+\epsilon \operatorname{Tr} A+O\left(\epsilon^{2}\right)$ where $A$ is an $n \times n$ square matrix.
- Show that $\left|e^{A}\right|=e^{\operatorname{Tr} A}$ where $A$ is a $n \times n$ square matrix.


### 1.1.12 Mathematical interlude: Dirac delta function

## Definition

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

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

Its integral properties are

$$
\begin{align*}
\int_{-\infty}^{\infty} f(x) \delta(x) d x & =f(0) \\
\int_{-\infty}^{\infty} \delta(x) d x & =1 \\
\int_{-\infty}^{\infty} f\left(x^{\prime}\right) \delta\left(x-x^{\prime}\right) d x^{\prime} & =f(x) \\
\int_{-\infty}^{\infty} \delta\left(x-x^{\prime}\right) d x^{\prime} & =1 \tag{1.93}
\end{align*}
$$

Note that

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

In three dimensions, the above definition is generalised as follows

$$
\begin{equation*}
\int_{\text {all space }} f(\mathbf{r}) \delta(\mathbf{r}-\mathbf{a}) d \mathbf{r}=f(\mathbf{a}) \tag{1.94}
\end{equation*}
$$

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

Some useful properties

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

$$
\begin{align*}
\delta^{\prime}(x) & =-\delta^{\prime}(-x) \\
x \delta(x) & =0 \\
\delta(a x) & =\frac{1}{|a|} \delta(x) \\
\delta\left(x^{2}-a^{2}\right) & =\frac{1}{|2 a|}[\delta(x-a)-\delta(x+a)] \\
\int_{-\infty}^{\infty} \delta(a-x) \delta(x-b) d x & =\delta(a-b) \\
f(x) \delta(x-a) & =f(a) \delta(x-a) \\
x \delta^{\prime}(x) & =-\delta(x) \\
\int g(x) \delta[f(x)-a] d x & =\left.\frac{g(x)}{|d f / d x|}\right|_{x=x_{0}, f\left(x_{0}\right)=a} \tag{1.95}
\end{align*}
$$

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

$$
x \delta^{\prime}(x)=-\delta(x)
$$

we proceed as follows

$$
\begin{align*}
\int_{-\infty}^{\infty} f(x) x \delta^{\prime}(x) d x & =\int_{-\infty}^{\infty} \frac{d}{d x}(f x \delta) d x-\int_{-\infty}^{\infty} \delta \frac{d}{d x}(f x) d x \\
& =-\int_{-\infty}^{\infty} \delta(x)\left(x \frac{d f}{d x}+f\right) d x \\
& =-\int_{-\infty}^{\infty} \delta(x) f(x) d x \tag{1.96}
\end{align*}
$$

where we have used integration by parts.

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

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

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

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

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

$$
\begin{equation*}
\left\langle a \mid a^{\prime}\right\rangle=\delta\left(a-a^{\prime}\right) \tag{1.97}
\end{equation*}
$$

The resolution of the identity (or closure relation) becomes

$$
\begin{equation*}
\int d a|a\rangle\langle a|=\hat{I} \tag{1.98}
\end{equation*}
$$

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

$$
\begin{equation*}
|\psi\rangle=\int d a^{\prime}\left|a^{\prime}\right\rangle\left\langle a^{\prime} \mid \psi\right\rangle \tag{1.99}
\end{equation*}
$$

with $\left\langle a^{\prime} \mid \psi\right\rangle$ denoting $|\psi\rangle$ in the $A$ representation. The inner product for two state vectors $|\psi\rangle$ and $|\phi\rangle$ is defined as

$$
\begin{align*}
\langle\psi \mid \phi\rangle & =\int d a^{\prime}\left\langle\psi \mid a^{\prime}\right\rangle\left\langle a^{\prime} \mid \phi\right\rangle \\
& =\int \psi^{*}\left(a^{\prime}\right) \phi\left(a^{\prime}\right) d a^{\prime} \tag{1.100}
\end{align*}
$$

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

$$
\begin{equation*}
|\psi\rangle=\sum_{a^{\prime}}\left|a^{\prime}\right\rangle\left\langle a^{\prime} \mid \psi\right\rangle+\int\left|a^{\prime}\right\rangle\left\langle a^{\prime} \mid \psi\right\rangle d a^{\prime} \tag{1.101}
\end{equation*}
$$

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

## Position and momentum representations for free particles

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

$$
\begin{align*}
\hat{x}\left|x^{\prime}\right\rangle & =x^{\prime}\left|x^{\prime}\right\rangle \\
\hat{p}\left|p^{\prime}\right\rangle & =p^{\prime}\left|p^{\prime}\right\rangle \\
\left\langle x \mid x^{\prime}\right\rangle & =\delta\left(x-x^{\prime}\right) \\
\left\langle p \mid p^{\prime}\right\rangle & =\delta\left(p-p^{\prime}\right) \tag{1.102}
\end{align*}
$$

These definitions, the fundamental commutator

$$
\begin{equation*}
[\hat{x}, \hat{p}]=i \hbar \tag{1.103}
\end{equation*}
$$

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

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

Problem $\mathbf{1 0}^{\dagger 1}$ : Verify the formulae (1.104)

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

$$
\hat{p}\left|p^{\prime}\right\rangle=p^{\prime}\left|p^{\prime}\right\rangle
$$

then we have

$$
\begin{align*}
\left\langle x^{\prime}\right| \hat{p}\left|p^{\prime}\right\rangle & =\int d x^{\prime \prime}\left\langle x^{\prime}\right| \hat{p}\left|x^{\prime \prime}\right\rangle\left\langle x^{\prime \prime} \mid p^{\prime}\right\rangle \\
& =\int d x^{\prime \prime}\left(-i \hbar \frac{\partial}{\partial x^{\prime}} \delta\left(x^{\prime}-x^{\prime \prime}\right)\right)\left\langle x^{\prime \prime} \mid p^{\prime}\right\rangle \\
& =-i \hbar \frac{\partial}{\partial x^{\prime}} \int d x^{\prime \prime} \delta\left(x^{\prime}-x^{\prime \prime}\right)\left\langle x^{\prime \prime} \mid p^{\prime}\right\rangle \\
& =-i \hbar \frac{\partial}{\partial x^{\prime}}\left\langle x^{\prime} \mid p^{\prime}\right\rangle \tag{1.105}
\end{align*}
$$

On the other hand, we also have

$$
\left\langle x^{\prime}\right| \hat{p}\left|p^{\prime}\right\rangle=p^{\prime}\left\langle x^{\prime} \mid p^{\prime}\right\rangle
$$

[^0]Therefore

$$
\begin{equation*}
-i \hbar \frac{\partial}{\partial x^{\prime}}\left\langle x^{\prime} \mid p^{\prime}\right\rangle=p^{\prime}\left\langle x^{\prime} \mid p^{\prime}\right\rangle \tag{1.106}
\end{equation*}
$$

which implies

$$
\begin{equation*}
\left\langle x^{\prime} \mid p^{\prime}\right\rangle=\frac{1}{\sqrt{2 \pi \hbar}} \exp \left(\frac{i p^{\prime} x^{\prime}}{\hbar}\right) \tag{1.107}
\end{equation*}
$$

where we have chosen the normalisation such that

$$
\begin{align*}
\left\langle p^{\prime \prime} \mid p^{\prime}\right\rangle & =\int d x^{\prime}\left\langle p^{\prime \prime} \mid x^{\prime}\right\rangle\left\langle x^{\prime} \mid p^{\prime}\right\rangle \\
& =\int d x^{\prime}\left\langle x^{\prime} \mid p^{\prime \prime}\right\rangle^{*}\left\langle x^{\prime} \mid p^{\prime}\right\rangle \\
& =\frac{1}{(2 \pi \hbar)} \int d x^{\prime} \exp \left(\frac{i\left(p^{\prime}-p^{\prime \prime}\right) x^{\prime}}{\hbar}\right) \\
& =\delta\left(p^{\prime \prime}-p^{\prime}\right) \tag{1.108}
\end{align*}
$$

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

$$
\begin{align*}
|\mathbf{r}\rangle & =|x, y, z\rangle \\
\hat{\mathbf{r}}|\mathbf{r}\rangle & =\mathbf{r}|\mathbf{r}\rangle \\
\left\langle\mathbf{r}^{\prime} \mid \mathbf{r}^{\prime \prime}\right\rangle & =\delta\left(\mathbf{r}^{\prime}-\mathbf{r}^{\prime \prime}\right) \\
|\mathbf{p}\rangle & =\left|p_{x}, p_{y}, p_{z}\right\rangle \\
\hat{\mathbf{p}}|\mathbf{p}\rangle & =\mathbf{p}|\mathbf{p}\rangle \\
\left\langle\mathbf{p}^{\prime} \mid \mathbf{p}^{\prime \prime}\right\rangle & =\delta\left(\mathbf{p}^{\prime}-\mathbf{p}^{\prime \prime}\right) \\
\left\langle\mathbf{r}^{\prime}\right| \hat{\mathbf{p}}\left|\mathbf{r}^{\prime \prime}\right\rangle & =-i \hbar \nabla_{\mathbf{r}^{\prime}} \delta\left(\mathbf{r}^{\prime}-\mathbf{r}^{\prime \prime}\right) \\
\left\langle\mathbf{p}^{\prime}\right| \hat{\mathbf{r}}\left|\mathbf{p}^{\prime \prime}\right\rangle & =i \hbar \nabla_{\mathbf{p}^{\prime}} \delta\left(\mathbf{p}^{\prime}-\mathbf{p}^{\prime \prime}\right) \\
\langle\mathbf{r} \mid \mathbf{p}\rangle & =\frac{1}{(2 \pi \hbar)^{3 / 2}} \exp (i \mathbf{r} \cdot \mathbf{p} / \hbar) \tag{1.109}
\end{align*}
$$


[^0]:    ${ }^{1}$ Problems that you may choose to skip on a first reading are indicated by $\dagger$.

