

Notes on Quantum Mechanics
Subramanian Ramakrishnan

1 Contents

- 1. Contents
- 2. Introduction
- 3. Mathematical Formalism
- 4. Review of Classical Mechanics
- 5. Problems with Classical Physics
- 6. Framework for Quantum Mechanics - The Postulates
- 7. Measurements and Ensembles
- 8. The Schrodinger Equation
- 9. One Dimensional Problems
- 10. The Harmonic Oscillator

2 Introduction

As engineering continues to expand into domains hitherto considered outside the circumference of its definition, a remarkable feature has been the giant strides it has taken towards the realm of the extremely small. In this context, the recent flurry of activity that finds a common umbrella under the name of ‘nano-technologies’ inevitably leads us to the regime of quantum phenomena. The present set of notes is the result of an attempt to understand quantum mechanics as it relates to modelling nano-scale dynamics. The principal text followed in this study is [1]. However the notes tend to go further than the text and draw parallels with the mathematical theory of elasticity wherever relevant, in order to enhance

the understanding of a reader familiar with the latter. The arena for both theories is the action of linear operators on vector spaces. In contrast to the theory of elasticity which is cast in the formalism of finite dimensional real vector spaces, quantum mechanics is set in infinite dimensional complex vector spaces. Indeed, significant issues arise in this leap from the real and finite to the complex and infinite that demand careful attention. However, it is felt that the ‘path of least resistance’ to the mathematical formalism of quantum mechanics begins from the mathematical structure of the theory of elasticity and it is in this spirit that these notes have been written.

The object of primary interest in the analysis of a dynamical system is a differential equation. Indeed, in modern mathematics, the differential equation(s) of motion itself is called a dynamical system (or a vector field). Solving the system yields functions of time (and space), which, along with initial conditions, completely determine the state of the system at any given time. All physical dynamical variables of interest such as energy and momentum are known functions of the state of the system. In other words, knowledge of the state of the system at any given point in time categorically solves the dynamical problem.

Classical mechanics offers three different (yet equivalent) methods of deriving the equations of motion viz. the Newtonian, Lagrangian and Hamiltonian approaches. An important point to note is that while ordinary differential equations represent the dynamics of discrete particles, partial differential equations are used to model continuous systems. The archetypes for both systems are the point particle and the wave familiar from elementary physics. In what might seem to be stating the obvious, we note that classical mechanics treats the dynamics of particles and waves as distinct physical phenomena. In other words, particles, by definition, are not considered to exhibit wave like properties and vice versa. Mathematically speaking, this translates to dynamical systems being represented by either ordinary or partial differential equations but never as a combination of both! Apart from the feature that classical mechanics considers deterministic evolution of systems, we also note that, a classical dynamical system spans a continuous spectrum of possible values for any dynamical variable. In other words, nothing prevents a dynamical system, in principle, from assuming any value of energy, momentum and so on, during its evolution as long as these values do not violate relativistic considerations.

The history of how the above principles were challenged by a sequence of experiments in atomic physics beginning in the twilight of the 19th century is well chronicled. The laws of classical physics failed to mesh with irrefutable experimental evidence in the atomic and sub-atomic regimes and had to give way to the principles of quantum mechanics, in those regimes. At a fundamental level, quantum mechanics established that, natural laws, in the smallest scales that we understand, are essentially probabilistic. The other remarkable consequence is that the very act of observation (measurement, to be precise) interferes with the dynamical evolution of a quantum system. The study of this aspect, after having been mostly dormant for a good 75 years, has acquired great importance in the current time when technology seeks to understand and control quantum systems.

These notes begin with an axiomatic development of the mathematical formalism of quantum mechanics. Most of the initial definitions and results are from basic linear algebra. However, a new notation useful in quantum mechanics, called the Dirac notation, is introduced.

3 Mathematical Formalism

3.1 Linear Vector Spaces

Definition 1: A linear vector space \mathbf{V} over a field of scalars S is a collection of objects, called vectors, with two binary operations $(\mathbf{V}, +, *)$ defined such that:

- (1) $(\mathbf{V}, +)$ is an abelian (commutative) group. This implies that, under the operation '+', (a) \mathbf{V} is closed, (b) \exists a unique inverse and (c) \exists a unique identity.
- (2) The second operation '*' (called scalar multiplication) is distributive and associative.

Definition 2: Denote $V \in \mathbf{V}$ by $|V\rangle$, $\forall V$ and call it 'ket' V .

Definition 3: Consider $|1\rangle, |2\rangle, \dots, |n\rangle$. The set is linearly independent if and only if $\sum_i a_i |i\rangle = 0 \Rightarrow a_i = 0 \forall i$. That is, no element of the vector space can be expressed as a linear combination of the others.

Theorem 1: Any vector $|V\rangle$ in an n -dimensional vector space may be written as a linear combination of n linearly independent vectors $|1\rangle, |2\rangle, \dots, |n\rangle$.

Definition 4: A set of n linearly independent vectors in an n dimensional space is called a basis.

Let $|1\rangle, |2\rangle, \dots, |n\rangle$ be a basis for \mathbf{V} . This implies $|V\rangle = \sum_i a_i |i\rangle \quad \forall |V\rangle \in \mathbf{V}$.

Theorem 2: Given a basis, the above expansion of any vector is unique.

3.2 Inner Product Spaces

As such, the definition of a vector space and the idea of a set of basis vectors that span the entire space provides us only with a skeletal structure. In order for the structure to be useful in modelling physical situations (and also for further mathematical development), we need a consistent method of associating scalar quantities with the elements of a vector space. The most basic association is, of course, the ‘norm’ or the ‘magnitude’ of a vector. At this point, it is useful to begin to view this association as a ‘function’ or ‘mapping’ of the vector space to the set of scalars. It is immediate that such a function should satisfy the two basic conditions for any relation defined between two sets to be a function viz. (1): Every vector in the space should map to some scalar (i.e., every vector should have a norm) and (2): No vector should map to more than one scalar (i.e., the norm should be unique). Furthermore, the norm should be a non-negative real number.

The definition of an inner product, essentially, ‘generates’ scalars out of pairs of vectors. Formally, this means a mapping from the product space of the vector space with itself (this is how we create pairs of vectors) to the set of scalars. That is, $\mathbf{V} \times \mathbf{V} \longrightarrow S$. It follows at once that the norm of a vector is a result of considering the inner product of the vector with itself.

Definition 5: An Inner Product Space is a vector space with an inner product defined on it as follows:

$$(1) \langle V|W\rangle = \langle W|V\rangle^* \quad \forall |V\rangle, |W\rangle \in \mathbf{V} \text{ (skew-symmetry)}$$

(2) $\langle V|V\rangle \geq 0 \quad \forall |V\rangle \in \mathbf{V}$ with 0 if and only if $|V\rangle = |0\rangle$ (positive semi-definite)

$$(3) \langle V|(a|W\rangle + b|Z\rangle) = \langle V|aW + bZ\rangle = a \langle V|W\rangle + b \langle V|Z\rangle \quad \forall |V\rangle, |W\rangle \in \mathbf{V}, \quad \forall a, b \in S \text{ (Linearity)}.$$

Some comments are in order here.

Firstly, the set of scalars considered here is the set of complex numbers. It is important to note that a complex vector space does NOT imply that the vectors themselves are complex in any way, just that the set of scalars associated with such a vector space is complex. Hence, the $*$ in (1) of the definition denotes complex conjugation. This is the first point of departure from a real vector space, where the inner product is defined to be symmetric. However, we note that the set of real numbers is a subset of the set of complex numbers. Hence there is no ambiguity in the norm of every vector being a non-negative real number under the definition of the inner product. Needless to add, in general, the inner product of two vectors in this space yields a complex number.

Secondly, the object $\langle V|$ which, by definition (2), generates the norm for the ket $|V \rangle$ (by mapping the ket to the set of real(trivially complex)numbers, to be precise) is called the bra $\langle V|$. Formally, $\langle V|$ is the dual of $|V \rangle$ and the space of the bra vectors (call it \mathbf{V}^*) is called the dual space of \mathbf{V} . It is clear that \mathbf{V}^* is distinct from the vector space \mathbf{V} . They may be viewed loosely as ‘complex mirror-images’ of each other. The idea of a dual space arises from the question of mapping a vector space to the set of scalars. The key point is that, in the case of real vector spaces, this distinction between the space and its dual is not apparent because of the set of scalars being the set of real numbers and the concomitant symmetry in the definition of the inner product. In order to clarify this statement, we wish to touch on a subtle point that seldom enters a usual discussion of inner products in a real vector space. Consider a vector \mathbf{a} in a real vector space. In a basis, the components of the vector may be expressed as a column (or row) vector. Conventionally, the components are collected

in a column vector: $\begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ a_n \end{pmatrix}$. Now, in order to generate the norm of this vector, we need to

multiply it on the left with its transpose given by: $(a_1 \ a_2 \ \cdot \ \cdot \ a_n)$. The transpose is, indeed, the dual of \mathbf{a} and vice versa.

Thirdly, the definition of the inner product presents our first exposure to the ingenuity of the Dirac notation. It is clear that the terms ‘bra’ and ‘ket’ are a result of a bifurcation of the

word bracket. In this notation, a full bracket $\langle V|V \rangle$, where V is any vector, represents a scalar. Furthermore, this notation brings out clearly the duality inherent in the definition of an inner product in a complex vector space. It keeps reminding us consistently that dual vectors (spaces) map vectors (spaces) to scalars. In other words, in order to generate a scalar out of a ket, we need to ‘hit it on the left’ with a bra. The plot thickens as we progress to the scene of linear operators acting on kets in a vector space transforming them to other kets (vectors are interesting and useful only because they transform in a certain way, anyway!) and in that situation the full power of the Dirac notation becomes apparent.

As an exercise in the algebraic manipulation of bras and kets, let us examine the linearity axiom (3 in Definition 5) in terms of a linear combination of bra vectors. The result is:

$$\begin{aligned}
 \langle aW + bZ|V \rangle &= \langle V|aW + bZ \rangle^* \quad (\text{from property 1}) \\
 &= (a \langle V|W \rangle + b \langle V|Z \rangle)^* \\
 &= a^* \langle V|W \rangle^* + b^* \langle V|Z \rangle^* \\
 &= a^* \langle W|V \rangle + b^* \langle Z|V \rangle.
 \end{aligned} \tag{1}$$

We conclude that **‘hitting’ a ket with a bra on the left invokes complex conjugation of the scalars.**

For the sake of completeness we gather the following two definitions that are familiar from real vector spaces.

Definition 6: Two vectors are orthogonal if and only if their inner product vanishes.

Definition 7: The norm of a vector $|V \rangle$ is defined as $|V| = \sqrt{\langle V|V \rangle}$.

3.2.1 Inner Product in a Basis

Before computing the inner product in a basis we note that, as in the case of any normalized set of basis vectors,

$$\langle i|j \rangle = \delta_{ij} \tag{2}$$

where δ_{ij} is the kronecker delta. Let

$$\begin{aligned}
 |V\rangle &= \sum_i v_i |i\rangle \\
 |W\rangle &= \sum_j w_j |j\rangle \\
 \Rightarrow \langle V| &= \sum_i v_i^* \langle i| \\
 \Rightarrow \langle V|W\rangle &= \sum_i v_i^* \langle i| \sum_j w_j |j\rangle \\
 &= \sum_i \sum_j v_i^* w_j \langle i|j\rangle \\
 &= \sum_i \sum_j v_i^* w_j \delta_{ij}.
 \end{aligned} \tag{3}$$

However,

$$\begin{aligned}
 \sum_j w_j \delta_{ij} &= w_i \text{ from property of the kronecker delta} \\
 \Rightarrow \langle V|W\rangle &= \sum_i v_i^* w_i.
 \end{aligned} \tag{4}$$

As a corollary, we have,

$$\begin{aligned}
 \langle V|V\rangle &= \sum_i v_i^* v_i \\
 &= \sum_i |v_i|^2 \geq 0.
 \end{aligned} \tag{5}$$

Notes: (1) In the column vector format, the ket $|V\rangle$ may be written as

$$\begin{pmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ v_n \end{pmatrix}.$$

Then, the corresponding bra $\langle V|$ is $(v_1^* \ v_2^* \ \cdot \ \cdot \ v_n^*)$.

$$(2) \langle V|W\rangle = (v_1^* \ v_2^* \ \cdot \ \cdot \ v_n^*) \begin{pmatrix} w_1 \\ w_2 \\ \cdot \\ \cdot \\ w_n \end{pmatrix}.$$

(3) We may also denote $\langle V|$ as the adjoint of $|V\rangle$. $Adj(|V\rangle) = \langle V|$. The corresponding bra vector is called the Hermitian conjugate or the ‘complex-conjugate transpose’ of the ket vector.

3.2.2 Transformations between Orthonormal (ON) bases

Let

$$|V\rangle = \sum_i v_i |i\rangle .$$

Let $|j\rangle$ be another ON basis.

$$\begin{aligned} \langle j|V\rangle &= \sum_i v_i \langle j|i\rangle \\ &= \sum_i v_i \delta_{ji} \\ &= \sum_i v_i \delta_{ij} \\ &= v_j . \end{aligned} \tag{6}$$

This leads to the following familiar and important result, in the Dirac notation,

$$\begin{aligned} |V\rangle &= \sum_i v_i |i\rangle \\ &= \sum_i \langle i|V\rangle |i\rangle \\ &= \sum_i |i\rangle \langle i|V\rangle . \end{aligned} \tag{7}$$

The next step is to examine how the above result transforms into the dual space. We know:

$$\langle aV| = \langle V|a^* . \tag{8}$$

Now consider:

$$a|V\rangle = b|W\rangle + c|Z\rangle + \dots \tag{9}$$

Using Eqn 8, the adjoint of the above equation 9 is:

$$\langle V|a^* = \langle W|b^* + \langle Z|c^* + \dots \tag{10}$$

Applying 10 to the ket $|V\rangle = \sum_i v_i |i\rangle$ the adjoint is

$$\langle V| = \sum_i \langle i|v_i^*. \quad (11)$$

Recalling that $v_i = \langle i|V\rangle$ and $v_i^* = \langle V|i\rangle$ we then get the important result that

$$|V\rangle = \sum_i |i\rangle \langle i|V\rangle \text{ maps to} \quad (12)$$

$$\langle V| = \sum_i \langle i| \langle V|i\rangle. \quad (13)$$

3.2.3 Gram-Schmidt Theorem

The Gram-Schmidt process converts a linearly independent (LI) basis to an orthonormal (ON) one. The idea behind the technique can be best explained in the two dimensional case. Consider a LI set of two vectors $\mathbf{1}$ and $\mathbf{2}$ in the plane. From the triangle law of vector addition,

$$\mathbf{p}_{21}(\text{projection of } \mathbf{2} \text{ on } \mathbf{1}) + \mathbf{2}' = \mathbf{2}.$$

$$\Rightarrow \mathbf{2}' = \mathbf{2} - \mathbf{p}_{21} \text{ (with } \mathbf{2}' \text{ perpendicular to } \mathbf{1})$$

$$\Rightarrow \mathbf{1}' = \frac{\mathbf{1}}{|\mathbf{1}|} \text{ and}$$

$$\mathbf{2}' = \frac{\mathbf{2}}{|\mathbf{2}'|}.$$

Thus the above exercise produces an ON basis.

In Dirac notation let $|\hat{1}\rangle, |\hat{2}\rangle, \dots$ be a LI basis. Let $|1\rangle, |2\rangle, \dots$ be the ON basis that we seek. Then the Gram-Schmidt procedure gives

$$|1\rangle = \frac{|\hat{1}\rangle}{\| |\hat{1}\rangle \|}.$$

To proceed further, consider

$$|2'\rangle = |\hat{2}\rangle - |1\rangle \langle 1|\hat{2}\rangle. \text{Hence}$$

$$|2\rangle = \frac{|2'\rangle}{\| |2'\rangle \|} \quad (14)$$

and so on.

3.2.4 The Schwartz and Triangle Inequalities

These important theorems may be stated as follows:

Schwarz Inequality: $|\langle V|W \rangle| \leq |V| |W| \quad \forall |V \rangle, |W \rangle \in \mathbf{V}$.

Triangle Inequality: $|V + W| \leq |V| + |W| \quad \forall |V \rangle, |W \rangle \in \mathbf{V}$.

3.3 Linear Operators on Vector Spaces

In order to proceed further, we need to transform vectors to other vectors. An object that effects this transformation in a linear manner is called a linear operator. The action of a linear operator Ω (henceforth, only linear operators are considered) on a vector $|V \rangle$ is represented as:

$$\Omega|V \rangle = |V' \rangle. \quad (15)$$

We say Ω transforms $|V \rangle$ to $|V' \rangle$. We also note that only operators that transform vectors to other vectors within the same vector space are considered in the discussion. The action of an operator on a bra is represented as:

$$\langle W|\Omega = \langle W'|. \quad (16)$$

The linearity of the operator is made explicit in the following expressions, where α and β are scalars.

$$\Omega(\alpha|V \rangle) = \alpha\Omega|V \rangle \quad (17)$$

$$\Omega(\alpha|V \rangle + \beta|W \rangle) = \alpha\Omega|V \rangle + \beta\Omega|W \rangle \quad (18)$$

$$(\langle V|\alpha)\Omega = \langle V|\Omega\alpha \quad (19)$$

$$(\langle V|\alpha + \langle W|\beta)\Omega = \langle V|\Omega\alpha + \langle W|\Omega\beta \quad (20)$$

The product of two operators represents the instruction that each one of them act on a vector in the same sequence that they appear in the product. Symbolically,

$$\begin{aligned} \Lambda\Omega|V \rangle &= \Lambda(\Omega|V \rangle) \\ &= \Lambda|\Omega V \rangle. \end{aligned} \quad (21)$$

Definition 8: The commutator of two operators Ω and Λ is given by:

$$[\Omega, \Lambda] = \Omega\Lambda - \Lambda\Omega \quad (22)$$

The most important thing about an operator product is the order in which the operators appear in the product. Two identities involving the commutator are:

$$[\Omega, \Lambda\Theta] = \Lambda [\Omega, \Theta] + [\Omega, \Lambda] \Theta \quad (23)$$

$$[\Lambda\Omega, \Theta] = \Lambda [\Omega, \Theta] + [\Lambda, \Theta] \Omega \quad (24)$$

We note that the above identities resemble the product rule for differentiation with care exercised in the ordering.

3.3.1 Matrix Representation of Linear Operators

Just as the concept of an abstract vector becomes useful only when the vector is represented by components in some basis, the utility of an abstract operator is realized in terms of components. Given a basis, the natural representation of a linear operator on a vector space is in terms of components collected in a matrix. Formally, we say the operator acts on the vector space via its matrix representation in some basis. Indeed, the matrix components are basis dependent but they “facilitate the computation of basis-independent quantities, by rendering the abstract operator more tangible.” [1]

A point to note here is that since a basis spans the entire space, the action of an operator on any vector is fully defined once its action on the basis is known. Hence, to compute the matrix representation of an operator Ω acting on an arbitrary vector, let us first consider its action on the basis. Let the basis $|i\rangle$ be transformed as

$$\Omega|i\rangle = |i'\rangle \quad (25)$$

Then, any vector in the space undergoes a change that can be computed as:

$$\begin{aligned} \Omega|V\rangle &= \Omega \sum_i v_i |i\rangle \\ &= \sum_i v_i \Omega|i\rangle \\ &= \sum_i v_i |i'\rangle. \end{aligned} \quad (26)$$

We use the following important property to find the components of $|i'\rangle$ in the original basis.

To find the components of a vector in any basis, hit it with the bra of that basis.

Hence,

$$\langle j|i' \rangle = \langle j|\Omega|i \rangle = \Omega_{ji} \quad (27)$$

Now, going back to the case of the action of Ω on an arbitrary vector, let

$$\Omega|V \rangle = |V' \rangle$$

We need to express the components of $|V' \rangle$, v' , in terms of Ω_{ij} and the components of $|V \rangle$.

We have:

$$\begin{aligned} v_i' &= \langle i|V' \rangle \\ &= \langle i|\Omega|V \rangle \\ &= \langle i|\Omega(\sum_j v_j|j \rangle) \\ &= \sum_j v_j \langle i|\Omega|j \rangle \\ &= \sum_j v_j \Omega_{ij} \end{aligned} \quad (28)$$

Hence, we have the important result:

$$v_i' = \sum_j \Omega_{ij} v_j. \quad (29)$$

In matrix form the above equation may be expressed as:

$$\begin{pmatrix} v_1' \\ v_2' \\ \cdot \\ v_n' \end{pmatrix} = \begin{pmatrix} \langle 1|\Omega|1 \rangle & \langle 1|\Omega|2 \rangle & \cdot & \cdot & \langle 1|\Omega|n \rangle \\ \langle 2|\Omega|1 \rangle & \langle 2|\Omega|2 \rangle & \cdot & \cdot & \langle 2|\Omega|n \rangle \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \langle n|\Omega|1 \rangle & \langle n|\Omega|2 \rangle & \cdot & \cdot & \langle n|\Omega|n \rangle \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \cdot \\ v_n \end{pmatrix} \quad (30)$$

Note: In the above equation, the elements of the first column are components of the first transformed vector $|1' \rangle = \Omega|1 \rangle$ and so on.

3.3.2 The Identity and Projection Operators

The Identity Operator: Based on the matrix representation $\Omega_{ij} = \langle i|\Omega|j \rangle$ for an arbitrary operator Ω that was explained in the previous sub-section, we have the special

case of the identity operator, given by,

$$\begin{aligned}
 I_{ij} &= \langle i|I|j \rangle \\
 &= \langle i|j \rangle \\
 &= \delta_{ij}.
 \end{aligned}
 \tag{31}$$

The Projection Operator: In a real vector space, consider $\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \dots + v_n\mathbf{e}_n$. Let us consider a set of operators \mathbf{P}_i such that

$$\mathbf{P}_i(\mathbf{v}) = v_i\mathbf{e}_i. \tag{32}$$

We note that **no summation** is implied over the repeated index in the above equation. Thus, the projection operators project the given vector onto the basis vectors. Now, if we have an orthonormal basis,

$$\langle \mathbf{e}_i, \mathbf{v} \rangle = v_i \tag{33}$$

$$\Rightarrow \mathbf{P}_i(\mathbf{v}) = \langle \mathbf{e}_i, \mathbf{v} \rangle \mathbf{e}_i. \tag{34}$$

Let us examine the analog of the above for complex vector spaces in the Dirac notation. Consider the expansion of a ket $|V \rangle$ in a basis:

$$\begin{aligned}
 |V \rangle &= \sum_i v_i |i \rangle \\
 &= \sum_i \langle i|V \rangle |i \rangle \\
 &= \sum_i |i \rangle \langle i|V \rangle.
 \end{aligned}
 \tag{35}$$

Let us write this as:

$$|V \rangle = \left(\sum_i |i \rangle \langle i| \right) |V \rangle \tag{36}$$

Comparing the above equation with the identity operator given by:

$$|V \rangle = I|V \rangle \tag{37}$$

we obtain the following identity:

$$\begin{aligned}
 \sum_i |i \rangle \langle i| &= I \\
 &= \sum_i \mathbf{P}_i.
 \end{aligned}
 \tag{38}$$

where

$$\mathbf{P}_i = |i\rangle\langle i|, \text{ no summation.} \quad (39)$$

\mathbf{P}_i is called the projection operator for the basis ket $|i\rangle$. The identity $\mathbf{P}_i = |i\rangle\langle i|$ is called the completeness relation which states that the sum of all the projection operators for a given basis equals the identity operator. In other words, the completeness relation simply implies that the sum of projections of a vector along all the n directions equals the vector itself.

Projections operate on the dual space as follows:

$$\begin{aligned} \langle V|\mathbf{P}_i &= \langle V|i\rangle\langle i| \\ &= v_i^* \langle i|. \end{aligned} \quad (40)$$

Intuitively, one can project a vector onto an element of a basis “only once”. Further projections onto the same basis element will act as the identity operator while further projections onto any other basis vector should yield the null vector since the basis is ON. We see this as:

$$\begin{aligned} \mathbf{P}_i\mathbf{P}_j &= |i\rangle\langle i|j\rangle\langle j| \\ &= |i\rangle\delta_{ij}\langle j| \\ &= |i\rangle\langle i| \text{ if and only if } i = j \\ &= \mathbf{P}_i \\ &= \mathbf{P}_j. \end{aligned} \quad (41)$$

In general,

$$\mathbf{P}_i\mathbf{P}_j = \delta_{ij}\mathbf{P}_j. \quad (42)$$

The matrix representations for \mathbf{P}_i is given as follows.

Consider $\mathbf{P}_i = |i\rangle\langle i|$. Following our rule from before, to generate a matrix representation $(\mathbf{P}_i)_{kl}$ we hit the above expression with a basis bra and a basis ket on the left and the right. We have:

$$\begin{aligned} (\mathbf{P}_i)_{kl} &= \langle k|i\rangle\langle i|l\rangle \\ &= \delta_{ki}\delta_{il} \\ &= \delta_{kl}\delta_{li}. \end{aligned} \quad (43)$$

For the case of operator products,

$$\begin{aligned}
(\Omega\Lambda)_{ij} &= \langle i|\Omega\Lambda|j\rangle \\
&= \langle i|\Omega I \Lambda|j\rangle \\
&= \sum_k \langle i|\Omega|k\rangle \langle k|\Lambda|j\rangle \\
&= \sum_k \Omega_{ik}\Lambda_{kj}.
\end{aligned} \tag{44}$$

3.3.3 Adjoint of an Operator

Given a ket $\alpha|V\rangle$, the corresponding bra is $\langle V|\alpha^*$. Hence, it is natural to raise the question of the counterpart in the dual space of an operator acting on a ket. In the case of the scalar α , complex conjugation was the only operation required to obtain the corresponding bra. In the case of operators, since they have a matrix representation, we need the complex-conjugate-transpose of the matrix representation to describe the picture in the dual space. We can make the statement that the action of an operator on a ket given by:

$$\Omega|V\rangle = |\Omega V\rangle \tag{45}$$

has the corresponding bra given by:

$$\langle \Omega V| = \langle V|\Omega^\dagger. \tag{46}$$

Here Ω^\dagger is called the adjoint of Ω . In a basis,

$$\begin{aligned}
(\Omega^\dagger)_{ij} &= \langle i|\Omega^\dagger|j\rangle \\
&= \langle \Omega i|j\rangle \\
&= \langle j|\Omega i\rangle^* \text{ (from inner product definition)} \\
&= \langle j|\Omega|i\rangle^*.
\end{aligned} \tag{47}$$

Thus we have the useful identity in the matrix representation:

$$\Omega^\dagger_{ij} = \Omega^*_{ji}. \tag{48}$$

It is important to remind ourselves, again, that expressions such as the above one are valid **only in a given a basis**. Or, as a mathematician would say, all statements are modulo a given basis.

A useful property of the adjoint is:

$$(\Omega\Lambda)^\dagger = \Lambda^\dagger \Omega^\dagger. \quad (49)$$

In order to demonstrate the properties developed so far, we work out an example involving scalars, kets and operators.

Example: Compute the adjoint of :

$$\alpha_1|V_1\rangle = \alpha_2|V_2\rangle + \alpha_3|V_3\rangle\langle V_4|V_5\rangle + \alpha_4\Omega\Lambda|V_6\rangle. \quad (50)$$

Solution: The adjoint is:

$$\begin{aligned} \langle V_1|\alpha_1^* &= \langle V_2|\alpha_2^* + \langle V_3|\alpha_3^* \langle V_5|V_4\rangle + \langle \Omega\Lambda V_6|\alpha_4^* \\ &= \langle V_2|\alpha_2^* + \langle V_3|\alpha_3^* \langle V_5|V_4\rangle + \langle V_6|(\Omega\Lambda)^\dagger\alpha_4^* \\ &= \langle V_2|\alpha_2^* + \langle V_3|\alpha_3^* \langle V_5|V_4\rangle + \langle V_6|\Lambda^\dagger\Omega^\dagger\alpha_4^*. \end{aligned} \quad (51)$$

RULE FOR ADJOINTS: When a product is encountered in the process of taking the adjoint, reverse the order of all factors and make the substitutions $\Omega \rightleftharpoons \Omega^\dagger$, $| \rangle \rightleftharpoons \langle |$, and $\alpha \rightleftharpoons \alpha^*$.

3.3.4 Hermitian and Unitary Operators

Definition 9: An operator Ω is Hermitian if $\Omega^\dagger = \Omega$ and anti-Hermitian if $\Omega^\dagger = -\Omega$.

Note: It is clear from the above definition that Hermitian operators play a role in (**finite**) complex vector spaces similar to that of symmetric matrices in real vector spaces. The emphasis on ‘finite’ is due to the complications that arise in the infinite dimensional case that we shall discuss in the sequel. For now, the analogy with symmetric matrices illuminates the parallel with the theory of elasticity wherein stress matrices, by definition, are symmetric. A point of view is that transformation of stresses is what propels the theory in the case of elasticity and a similar role in terms of transforming kets in a useful manner is taken up by Hermitian operators here.

Definition 10: An operator U is unitary iff. $UU^\dagger = I$, where I is the identity operator.

Theorem: Unitary operators preserve the inner product.

Proof: Let two arbitrary kets $|V_1\rangle$ and $|V_2\rangle$ undergo a unitary transformation under the operator U . Let

$$\begin{aligned} |V_1'\rangle &= U|V_1\rangle \\ |V_2'\rangle &= U|V_2\rangle. \end{aligned} \tag{52}$$

$$\begin{aligned} \langle V_2'|V_1'\rangle &= \langle UV_2|UV_1\rangle \\ &= \langle V_2|U^\dagger U|V_1\rangle \\ &= \langle V_2|V_1\rangle. \end{aligned} \tag{53}$$

Note: Unitary operators are analogous to rotations in \mathbf{R}^n . Rotations preserve dot products in \mathbf{R}^n and unitary operators preserve the inner product in the complex case.

3.3.5 Active and Passive Transformations

Consider a unitary operator U acting on a vector space \mathbf{V} according to $|V\rangle \mapsto U|V\rangle \forall |V\rangle \in \mathbf{V}$. We raise the question of how does an arbitrary operator Ω transform under U ? The answer is obtained by examining explicitly the action of Ω on vectors that have undergone the unitary transform effected by U .

$$\langle V'|\Omega|V\rangle \mapsto \langle UV'| \Omega |UV\rangle = \langle V'|U^\dagger \Omega U|V\rangle$$

Hence, under the unitary transformation U , the operator Ω transforms as:

$$\Omega' = U^\dagger \Omega U \tag{54}$$

Note: We note the analogy again with \mathbf{R}^n . Under the action of a rotation matrix \mathbf{R} , the matrix representation of any operator \mathbf{A} (or, in less formal language, any matrix \mathbf{A}) transforms as $\mathbf{A} \mapsto (\mathbf{R})^T \mathbf{A} \mathbf{R}$.

Note: A consequence of the above discussion is of importance in quantum mechanics. It is that, in generating a dynamical picture out of the formalism of vector spaces, the theory may be formulated in terms of transformations on vectors or, alternatively, on operators. In quantum mechanics this dualism translates into the equivalence of the Schrodinger picture (based on the dynamical evolution of states of a system represented by kets) and the Heisenberg picture (based on the dynamical evolution of Hermitian operators that act on the kets).

3.4 The Eigenvalue Problem

The eigenvalue problem is of great importance in the study of linear algebra and its applications. Why is this so special?

Consider a linear operator Ω acting on an arbitrary, non-vanishing, ket $|V\rangle$ given by:

$$\Omega |V\rangle = |V'\rangle .$$

Now, unless Ω happens to be the identity operator, we cannot expect a simple relationship between $|V\rangle$ and $|V'\rangle$. However, associated with every operator, there exists a special class of kets (known as eigen kets) such that the action of the operator on any of those kets results in a scalar multiple of the same ket. The scalar multiples and the special kets are known as the eigenvalues and eigenkets **of that operator**. Appealing to geometric intuition, we may also informally view the eigen kets as those vectors whose “direction” is unchanged under the action of the operator. The problem of obtaining the eigenvalues and eigenkets of an operator is termed the eigenvalue problem. For a given operator Ω the eigenvalue problem may be stated in terms of the following equation:

$$\Omega |V\rangle = \omega |V\rangle .$$

We collect the main results from considering the eigenvalue problem in the following theorems.

Theorem: The eigenvalues of a Hermitian operator are real.

The above theorem is a very powerful result from a physical point of view. If we were to associate physical quantities with the eigenvalues of a Hermitian operator, we are guaranteed that they will be real numbers. The parallel with the situation in the theory of elasticity follows at once. There, stress matrices (again, matrix representations of stress tensors to be precise!) are symmetric and hence their eigenvalues are guaranteed to be real. In other words, principal stresses are always real numbers. Even though a detailed discussion of the role of Hermitian operators in quantum mechanics will be undertaken later in these notes, we mention here that all measurable quantities (known as **observables**) in quantum mechanics are represented by Hermitian operators. Moreover, the actual measured values for an observable (say, energy) will be one of the eigenvalues of the corresponding Hermitian

operator (say, the energy operator). Therefore, solving the eigenvalue problem is often the central issue in solving for the dynamics of a quantum system.

Theorem: There exists, for every Hermitian operator, a basis comprising its ON eigenvectors. The operator is diagonal in this basis and the diagonal entries are precisely the eigenvalues of the operator.

Indeed, the above theorem reminds us of the structure of the stress matrix in the principal frame. The shear stresses vanish in this case.

Naturally, the next question is: if there exists a basis in which the operator is diagonal, how does one represent the operator in this basis, in a canonical way? The answer is provided by the following theorem.

Theorem: If Ω is a Hermitian operator, there exists a **unitary** operator U such that $U^\dagger \Omega U$ is diagonal. The columns of U are the eigenvectors of Ω .

Apart from establishing the intimate relationship between Hermitian and Unitary operators, the above theorem also implies that the diagonalization of a Hermitian operator is tantamount to solving its eigenvalue problem. The next result is a statement on the simultaneous diagonalization of two Hermitian operators.

Theorem: If Ω and Λ are two commuting Hermitian operators, i.e., $[\Omega, \Lambda] = 0$, then there exists at least a basis of common eigenvectors that diagonalizes them both.

The following two theorems on unitary operators are mentioned for the sake of completeness.

Theorem: The eigenvalues of a unitary operator are complex numbers of unit modulus.

Theorem: If a unitary operator has no multiple eigenvalues (no degeneracy), then its eigenvectors are mutually orthogonal.

3.4.1 The Propagator

The idea of the propagator is important in developing the dynamical picture of a quantum system. To introduce the concept, we consider a simple example from classical dynamics. Consider a system comprising two masses of m units each and two springs (stiffness k)

connected in series. The coupled equations of motion may be written as:

$$\ddot{x}_1 = -\frac{2k}{m}x_1 + \frac{k}{m}x_2 \quad (55)$$

$$\ddot{x}_2 = \frac{k}{m}x_1 - \frac{2k}{m}x_2 \quad (56)$$

The problem is to obtain $x_1(t)$ and $x_2(t)$, given the initial conditions. The abstract form of the above equations may be written as:

$$|\ddot{x}\rangle = \Omega |x(t)\rangle \quad (57)$$

In the matrix representation the above equation takes the following form:

$$\begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{pmatrix} = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (58)$$

where the elements of the Hermitian matrix Ω_{ij} are given by:

$$\begin{aligned} \Omega_{11} = \Omega_{22} &= -\frac{2k}{m} \\ \Omega_{12} = \Omega_{21} &= \frac{k}{m}. \end{aligned} \quad (59)$$

Now, the key point here (as has been mentioned repeatedly before) is that the operator Ω has the matrix representation above only in a certain basis. In this basis, the state $|x\rangle$ can be represented as:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} x_1 + \begin{pmatrix} 0 \\ 1 \end{pmatrix} x_2 \quad (60)$$

Also, the abstract counterpart of the above equation may be written as:

$$|x\rangle = |1\rangle x_1 + |2\rangle x_2. \quad (61)$$

It is in this $|1\rangle, |2\rangle$ basis that Ω has the matrix representation described above. We note that this basis has a convenient physical interpretation since $|1\rangle, |2\rangle$ represent unit displacements of the two masses. However, the price that is paid for this is the mathematical complexity that the equations of motion become coupled between the two variables. It is then natural to ask for a basis wherein the equations of motion are decoupled thereby making the mathematical solution easier to obtain. That such a basis is provided by the eigenvectors

of the operator is the central idea of the rest of this discussion. On obtaining the solution, one can transform to the $|1\rangle$, $|2\rangle$ basis for clarity in the physics.

From the theorems on Hermitian operators stated in the previous sub-section it follows that the basis in which the equations of motion are decoupled (in other words, the basis in which the matrix representation of the operator is diagonal) is the eigenbasis of the operator. Assuming the absence of degeneracy, i.e., that the eigen values are distinct, let the eigenvectors of Ω be denoted by $|1\rangle$ and $|2\rangle$ and given by the following equations:

$$\begin{aligned}\Omega|1\rangle &= -\omega_1^2 |1\rangle \\ \Omega|2\rangle &= -\omega_2^2 |2\rangle.\end{aligned}\tag{62}$$

Here, the eigenvalues are denoted by ω^2 for convenience. For the system under consideration, the eigenvalues and eigenvectors may be evaluated in the $|1\rangle$, $|2\rangle$ basis as:

$$\omega_1 = \left(\frac{k}{m}\right)^{\frac{1}{2}},\tag{63}$$

$$|1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.\tag{64}$$

$$\omega_2 = \left(\frac{3k}{m}\right)^{\frac{1}{2}},\tag{65}$$

$$|2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.\tag{66}$$

In the new basis, the abstract representation of $|x(t)\rangle$ is:

$$|x\rangle = |1\rangle x_1 + |2\rangle x_2.\tag{67}$$

The explicit matrix form of the above equation reads:

$$\begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{pmatrix} = \begin{pmatrix} -\omega_1^2 & 0 \\ 0 & -\omega_2^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.\tag{68}$$

Now, it is straightforward to obtain the solutions to the de-coupled equations

$$\ddot{x}_i + \omega_i^2 x_i = 0; \quad i = 1, 2\tag{69}$$

subject to vanishing initial velocities, as:

$$x_i(t) = x_i(0) \cos \omega_i t; \quad i = \mathbf{1}, \mathbf{2} \quad (70)$$

Using this result, the general solution can be written as:

$$|x(t)\rangle = |\mathbf{1}\rangle x_1(0) \cos \omega_1 t + |\mathbf{2}\rangle x_2(0) \cos \omega_2 t \quad (71)$$

where the scalars $x_1(0)$ and $x_2(0)$ are obtained from the inner products

$$\begin{aligned} x_1(0) &= \langle \mathbf{1} | x(0) \rangle \\ x_2(0) &= \langle \mathbf{2} | x(0) \rangle . \end{aligned} \quad (72)$$

The above exercise can be summarized in an algorithmic manner to compute the state of the system $|x\rangle$ given the initial condition $|x(0)\rangle$. The steps are:

- (1): Solve the eigenvalue problem for the operator Ω .
- (2): Find the coefficients $x_1(0) = \langle \mathbf{1} | x(0) \rangle$ and $x_2(0) = \langle \mathbf{2} | x(0) \rangle$ in the expansion:

$$|x\rangle = |\mathbf{1}\rangle x_1 + |\mathbf{2}\rangle x_2.$$

(2): Multiply each coefficient with a time dependent function $\cos \omega_i t$ where $i = \mathbf{1}, \mathbf{2}$ to obtain the full coefficients in the expansion of $|x(t)\rangle$.

Using the above algorithm, the final state vector for any given time is obtained by multiplying the initial state vector by a matrix. This matrix, which is independent of the initial state, is called the **propagator** for the system. In abstract notation,

$$|x(t)\rangle = U(t)|x(0)\rangle . \quad (73)$$

It is clear from the above equation that the propagator $U(t)$ is an operator acting on the vector representing the initial state of the system. A representation for $U(t)$ in the Dirac notation can be written as:

$$U(t) = \sum_i |i\rangle \langle i| \cos \omega_i t \quad i = \mathbf{1}, \mathbf{2} \quad (74)$$

Finally, we can frame the algorithm described earlier in terms of the propagator. To solve the equation

$$|\ddot{x}\rangle = \Omega|x\rangle \quad (75)$$

(1): Solve the eigenvalue problem for the operator Ω .

(2): Construct the propagator U in terms of the eigenvalues and eigenvectors.

(3): $|x(t)\rangle = U(t)|x(0)\rangle$.

We can now explore the implications of the concept of the propagator based on the results that have been derived.

3.4.2 The Normal Modes

The eigenstates of Ω undergo uniquely simple evolution under the action of the propagator. To see this, let $|x(0)\rangle = |\mathbf{1}\rangle$. Using the representation of $U(t)$ in the Dirac notation, we can write

$$\begin{aligned} |\mathbf{1}(t)\rangle &= U(t)|\mathbf{1}\rangle \\ \Rightarrow |\mathbf{1}(t)\rangle &= \left(\sum_i |i\rangle\langle i| \cos\omega_i t\right)|\mathbf{1}\rangle \\ \Rightarrow |\mathbf{1}(t)\rangle &= |\mathbf{1}\rangle \cos\omega_1 t. \end{aligned} \tag{76}$$

Evidently, the eigenvector $|\mathbf{1}\rangle$ is only multiplied by the phase factor $\cos\omega_1 t$ as it evolves in time. In other words, it does not undergo a change of “direction”. This is true of all the eigenvectors of Ω and they are also known as the **normal modes** of oscillation of the system. As a useful corollary we also note that a system that starts off in a linear combination of $|\mathbf{1}\rangle$ and $|\mathbf{2}\rangle$ evolves into the corresponding linear combination of the normal modes.

Having developed the formalism of vector spaces thus far, for the first time, we now apply the ideas to quantum dynamics. It is important to note that the following are postulates to begin with. However the conclusions that emerge from the theory based on the postulates have withstood the test of experiments over the last 75 years.

(1): The **state** of a quantum system, denoted by a **ket** $|\psi\rangle$ evolves according to Schrödinger’s equation $i\hbar \dot{|\psi\rangle} = \mathbf{H}|\psi\rangle$, where $i = \sqrt{-1}$, $\hbar = \frac{h}{2\pi}$ h – *Planck’s constant* and \mathbf{H} is the Hamiltonian (total energy) of the system. It is important to note that the Hamiltonian is a Hermitian operator.

(2): To solve for the dynamics of the system, in other words to obtain $|\psi(t)\rangle$ in terms of $|\psi(0)\rangle$, solve the eigenvalue problem of \mathbf{H} .

(3): Construct the propagator $U(t)$ in terms of the eigenvalues and eigenvectors of \mathbf{H} .

$$(4): |\psi(t)\rangle = U(t)|\psi(0)\rangle.$$

3.5 Generalization to Infinite Dimensions

3.5.1 The Inner Product and the Dirac Delta Function

The development so far tacitly assumed that the vector spaces involved are finite dimensional. However, vector spaces associated with quantum mechanics (as well as classical continuous systems) are infinite dimensional. Naturally, this has implications for the definitions of the vectors themselves as well as for the properties of the operators acting on them.

The primary characteristic of an infinite dimensional vector space is that the vectors (kets, in our case) are continuous functions. Hence, infinite dimensional vector spaces are also known as function spaces. In other words, the state of a system is represented by a continuous function. The first important consequence of this situation is that the definition of the inner product is extended as follows. Let us assume that the vectors are functions of a single independent variable x and let the domain of the functions be the entire real line (\mathbf{R}^1).

Definition: Let $f(x), g(x) \in \mathbf{V}$. The inner product between these vectors is defined as

$$\langle f|g \rangle = \int_{-\infty}^{\infty} f^*(x)g(x)dx \quad (77)$$

where $f^*(x)$ is the conjugate of $f(x)$. The most important feature of this generalization of replacing the sum in the inner product with an integral is the concomitant change in the normalization conditions of the basis vectors that is required for mathematical consistency. We recall that in the finite dimensional case, we had the normalization conditions for the basis vectors as

$$\langle i|j \rangle = \delta_{ij}. \quad (78)$$

In the present case, for basis vectors at two points in the domain x and x' , the condition involves the Dirac Delta function and is defined as

$$\langle x|x' \rangle = \delta(x - x') \quad (79)$$

where $\delta(x - x')$, the Dirac Delta function has the following properties:

$$\delta(x - x') = 0 \quad \forall x \neq x' \quad (80)$$

$$\int_{-\infty}^{\infty} \delta(x - x') dx = 1, \quad -\infty < x < \infty \quad (81)$$

$$\int_{-\infty}^{\infty} \delta(x - x') f(x') dx = f(x). \quad (82)$$

Two subtle points are to be noted here. Firstly, in our definition of orthogonality, we are constrained to consider basis vectors at any two points in the domain due to the infinite dimensional nature of the vector space. Secondly, the “definition” of the Dirac Delta function makes sense only under the integral sign. The Delta function in the neighbourhood of a point acts on any function in the domain in such a way as to “pick out” the value of the function at that point. It is possible to exploit the properties of integral transforms in conjunction with the above property in order to generate useful alternative definitions of the Dirac Delta function. As an example of this, we work out the definition of the Delta function using the classical Fourier transform.

The Fourier transform of a “well behaved” function $f(x)$ is defined by

$$f(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-ikx) f(x) dx \quad (83)$$

with the inverse

$$f(x') = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(ikx') f(k) dk. \quad (84)$$

Substituting the forward transform into the inverse, we obtain

$$f(x') = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(ikx') \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-ikx) f(x) dx \right] dk \quad (85)$$

$$\Rightarrow f(x') = \int_{-\infty}^{\infty} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} dk \exp(ik(x - x')) \right] f(x) dx. \quad (86)$$

Comparison with the property of the Delta function given by

$$\int_{-\infty}^{\infty} \delta(x - x') f(x') dx = f(x)$$

immediately gives

$$\delta(x - x') = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \exp(ik(x - x')). \quad (87)$$

In conclusion of the brief comments on the Delta function, we note that when they were introduced in 1930 by Dirac, they were mathematically unacceptable but “worked” for the physicist. Delta functions are but one example of a class of functions known as generalized functions or distributions, the rigorous theory of which was enunciated by the french mathematician Laurent Schwartz circa 1950.

3.5.2 Operators in Infinite Dimensions

The essential feature of an operator in an infinite dimensional vector space is that it maps functions to functions. The two key questions here are: (1) Given a basis for the function space, what are the matrix elements of an operator in the basis? and (2) Is the definition of Hermiticity the same as it is in the finite dimensional case?

Since the differential operator D , whose action is defined by $D|f\rangle = \left|\frac{df}{dx}\right\rangle$ is a fundamental one, we address the two questions above in the context of this operator.

(1) What are the matrix elements of D in the $|x\rangle$ basis?: Let D act on the ket $|f\rangle$. Since the discussion is confined to the $|x\rangle$ basis, following the anstaz of taking the inner product of the resultant ket $D|f\rangle$ with $\langle x|$, we obtain

$$\begin{aligned} \langle x|D|f\rangle &= \langle x|\frac{df}{dx}\rangle \\ &= \frac{df(x)}{dx}. \end{aligned} \quad (88)$$

The above expression, simple as it appears, masks the role of the Delta function in its derivation. To see this, let us consider the action of D on the basis (functions). From the normalization conditions we have,

$$\begin{aligned} \langle x|D|x'\rangle &= D_{xx'} \\ &= \delta'(x - x') \\ &= \delta(x - x')\frac{d}{dx'}. \end{aligned} \quad (89)$$

Hence the matrix elements of the action of D on the basis may be written as

$$D_{xx'} = \delta(x - x')\frac{d}{dx'} \quad (90)$$

It is due to the presence of the Delta function in the above expression that the action of D on an arbitrary $f(x)$ reduces to $\frac{df(x)}{dx}$.

(2) Is the definition of Hermiticity of an operator the same as in the finite dimensional case?

To answer this important question, let us consider the matrix elements of D as defined using the Delta function. If D is Hermitian, we would have $D_{xx'} = D_{x'x}^*$. But this is not the case since

$$\begin{aligned} D_{xx'} &= \delta'(x - x') \\ D_{x'x}^* &= \delta'(x' - x)^* \\ &= \delta'(x' - x) \\ &= -\delta'(x - x'). \end{aligned}$$

Now, if we consider $K = -iD$, where i is a pure imaginary number of unit modulus, $K_{xx'} = K_{x'x}^*$. The point is that **satisfying the above identity is not a sufficient condition for K to be Hermitian.**

Consider two kets in a function space $|f\rangle, |g\rangle$ which are represented in the X basis by two functions $f(x)$ and $g(x)$ in the interval (a, b) . If K is a Hermitian operator, it must satisfy

$$\begin{aligned} \langle g|K|f\rangle &= \langle g|Kf\rangle \\ &= \langle Kf|g\rangle^* \\ &= \langle f|K^\dagger|g\rangle^* \\ &= \langle f|K|g\rangle^*. \end{aligned} \tag{91}$$

Keeping in mind the definition of the inner product in a function space, one can show that the above identity is satisfied if and only if the surface term emerging out of the integral vanishes. Mathematically, this condition is

$$-ig^*(x)f(x)|_a^b = 0. \tag{92}$$

Thus, the conclusion here is that the conditions for Hermiticity carry over to the infinite dimensional case with the important rider that the functions themselves should satisfy the

boundary conditions prescribed by Eqn 92. We note that this is automatically satisfied in certain physical situations, for example, the case of a string held fixed at the end points. Another instance is when one considers a space of periodic functions ($f(\theta) = f(\theta + 2\pi)$). Here, $K = -i\frac{d}{d\theta}$ is a Hermitian operator.

4 Review of Classical Mechanics

The central problem of classical dynamics is solving for the motion (trajectories) of bodies moving under the influence of specified forces. In this section we briefly review the three formulations of classical mechanics viz. the Newtonian, Lagrangian and the Hamiltonian. The passage to quantum mechanics is most straightforward from the Lagrangian and Hamiltonian formulations. Quantum mechanics emerged first in the Hamiltonian formalism while the connection with the Lagrangian approach was made explicit by the path integral approach of Richard Feynman.

4.1 The Newtonian and Lagrangian Formalisms

Consider a point particle of mass m moving along the x axis under the influence of a potential $V(x)$. In the Newtonian scheme one formulates the following differential equation (DE) of motion:

$$m\ddot{x} = -\frac{dV}{dx} \quad (93)$$

The trajectory of the particle is a function of time $x(t)$ that satisfies the DE of motion at all times. Here, if one is given the initial state of the particle characterized by its initial position and initial velocity, one can precisely solve for the trajectory in time. From a mathematical point of view the equation of motion (EOM) being a second order DE, the general solution essentially involves two arbitrary constants of integration which identically correspond to the initial position and the initial velocity. Specific values assigned to these quantities pick out unique particular solutions (trajectories) from the family of general solutions to the DE.

The point of view here is that the trajectory evolves in every instant of time so as to satisfy the EOM. Indeed, in this local description, the actual trajectory between two spatial

points (say, $x_{initial}$ and x_{final}) is viewed as a piecing together of infinitesimal displacements such that the EOM is satisfied at all times.

In the Lagrangian formalism the same dynamical problem is posed in a different way: given that the particle is at x_i and x_f at times t_i and t_f respectively, what distinguishes the actual trajectory from all the other possible trajectories that connect the initial and final points?

The Lagrangian approach is thus global, seeking to determine the entire trajectory in one stroke, in contrast to the Newtonian approach that is constantly concerned with the displacement in the imminent, infinitesimal time step.

The Lagrangian method of solving for the dynamics comprises three parts:

(1) Define a **scalar function** L called the Lagrangian, given by $L = T - V$ where T and V are the kinetic energy and potential energy of the particle. Thus, in general, $L = L(x, \dot{x}, t)$.

(2) Associate with **every** path $x(t)$ connecting the points (x_i, t_i) and (x_f, t_f) a quantity called the action $S[x(t)]$ defined by:

$$S[x(t)] = \int_{t_i}^{t_f} L(x, \dot{x}) dt. \quad (94)$$

Note: S depends on the entire path (or function) $x(t)$. Hence it is a ‘function of a function’ and also called a functional.

(3) The actual trajectory (ie. the solution to the dynamical problem) is the one for which S is a minimum. The question that then arises naturally is: What is the analytical condition for S to be a minimum? The answer, provided by Hamilton’s principle from the calculus of variations is the celebrated Euler-Lagrange equation

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) - \frac{\partial L}{\partial x} = 0. \quad (95)$$

Indeed, the above condition for S to be a minimum is identically the equation of motion of the particle.

The salient features of the Lagrangian formalism are the following:

(1) The scheme is pivoted on constructing a **scalar** function, the Lagrangian, in order to obtain the EOM. In contrast, the Newtonian approach is vectorial and hence can swiftly get complicated.

(2) The Euler-Lagrange equations are form invariant (covariant) with respect to a change of co-ordinates. It is also to emphasize this aspect that, traditionally, the Euler-Lagrange equations are written in terms of generalized co-ordinates q_i . The EOM in the Newtonian approach, in general, remain simple only in cartesian co-ordinates.

(3) Conservation laws are easily obtained in the Lagrangian formalism. Suppose that the Lagrangian depends on a certain (generalized) velocity \dot{q}_i but not on the corresponding displacement q_i . It follows from the Euler-Lagrange equations that the corresponding generalized momentum p_i (defined by $p_i = \frac{\partial L}{\partial \dot{q}_i}$) is conserved (ie. a constant of the motion). Hence q_i is called a cyclic co-ordinate. Thus, a simple examination of the Lagrangian yields the conserved quantities of the motion.

A Note on $L = T - V$: The form of the Lagrangian $L = T - V$ is valid only when the forces under consideration are conservative. However, it is to be noted that the formalism holds even for non-conservative forces. In other words, even in the presence of non-conservative forces, the Euler-Lagrange equations are the correct EOM provided the correct Lagrangian for the problem can be constructed. It is entirely another matter that constructing the appropriate Lagrangian in such cases is an exercise fraught with difficulties. In fact, this was among the most daunting and exciting problems that faced particle physics in its march towards creating the standard model of fundamental forces, during the last century.

4.2 The Hamiltonian Formalism

In the Lagrangian formalism the independent variables are the co-ordinates q_i and the velocities \dot{q}_i . The momenta are derived quantities defined by

$$p_i = \frac{\partial L}{\partial \dot{q}_i}. \quad (96)$$

In the Hamiltonian formalism, the roles of q_i and p_i are interchanged. The velocities \dot{q}_i become derived quantities, derived from a scalar function called the Hamiltonian $H(p, q)$. The velocities are given by $\dot{q}_i = \frac{\partial H}{\partial p_i}$. The Hamiltonian $H(p, q)$ replaces the Lagrangian $L(q, \dot{q})$ in this formalism. The relation between the two is established by the Legendre transform given by:

$$H(q, p) = \sum_i p_i \dot{q}_i - L(q_i, \dot{q}_i) \quad (97)$$

The EOM are Hamilton's equations given by:

$$\begin{aligned}\frac{\partial H}{\partial p_i} &= \dot{q}_i; \\ -\frac{\partial H}{\partial q_i} &= \dot{p}_i.\end{aligned}\tag{98}$$

A key point here is that the EOM are first order equations. For a system with n degrees of freedom we have $2n$ first order equations. We also note that for a problem involving conservative forces, $H = T + V$, where T and V are the kinetic and potential energies of the system. Thus, in this situation, H represents the total energy of the system.

Let us illustrate the Hamiltonian formalism using the classic example of the simple harmonic oscillator.

Example: Consider the simple harmonic oscillator with a single degree of freedom in the co-ordinate x . The Lagrangian is :

$$L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2.\tag{99}$$

Hence, the canonical momentum may be written as:

$$p = \frac{\partial L}{\partial \dot{x}} = m\dot{x}.\tag{100}$$

Inverting this relation we obtain:

$$\dot{x} = \frac{p}{m}.$$

Now, the Hamiltonian can be written down as:

$$\begin{aligned}H &= T + V \\ &= \frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2 \\ &= \frac{p^2}{2m} + \frac{1}{2}kx^2.\end{aligned}\tag{101}$$

The EOM are :

$$\begin{aligned}\frac{\partial H}{\partial p} = \dot{q} &\Rightarrow \frac{p}{m} = \dot{x} \\ -\frac{\partial H}{\partial q} = \dot{p} &\Rightarrow -kx = \dot{p}.\end{aligned}\tag{102}$$

In order to generate the second order equation, let us differentiate with respect to time the first of the above equations to obtain

$$\dot{p} = m\ddot{x}. \quad (103)$$

Now use the second of the EOM to obtain the familiar equation:

$$m\ddot{x} = -kx. \quad (104)$$

4.3 Cyclic Co-ordinates, Poisson Brackets and Canonical Transformations

4.3.1 Cyclic Co-ordinates

The definition of a cyclic co-ordinate remains unchanged in the Hamiltonian scheme. If a co-ordinate is absent in the Hamiltonian, the corresponding conjugate momentum is conserved. This is a direct consequence of Hamilton's equations as can be readily seen. Suppose the expression for the Hamiltonian does not involve a certain co-ordinate q_k . Then

$$\begin{aligned} \dot{p}_k &= -\frac{\partial H}{\partial q_k} \\ &= 0. \end{aligned} \quad (105)$$

4.3.2 Poisson Brackets

The Poisson Bracket(PB) provides an elegant way of characterizing all conserved quantities in the Hamiltonian formalism. Let $\omega(p, q)$ be a dynamical variable, ie. some function of the state variables p, q . Consider :

$$\begin{aligned} d\omega &= \frac{\partial \omega}{\partial p} dp + \frac{\partial \omega}{\partial q} dq \\ \Rightarrow \frac{d\omega}{dt} &= \sum_i \frac{\partial \omega}{\partial p_i} \frac{dp_i}{dt} + \frac{\partial \omega}{\partial q_i} \frac{dq_i}{dt}. \end{aligned} \quad (106)$$

Using Hamilton's equations, we can now write

$$\begin{aligned}
\frac{d\omega}{dt} &= \sum_i \frac{\partial\omega}{\partial p_i} \left(\frac{-\partial H}{\partial q_i} \right) + \frac{\partial\omega}{\partial q_i} \left(\frac{\partial H}{\partial p_i} \right) \\
&= \sum_i \left(\frac{\partial\omega}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial\omega}{\partial p_i} \frac{\partial H}{\partial q_i} \right) \\
&= \{\omega, H\}.
\end{aligned} \tag{107}$$

Here, we have **defined** the Poisson Bracket between any two dynamical variables ω and λ as:

$$\{\omega, \lambda\} = \sum_i \left(\frac{\partial\omega}{\partial q_i} \frac{\partial\lambda}{\partial p_i} - \frac{\partial\omega}{\partial p_i} \frac{\partial\lambda}{\partial q_i} \right). \tag{108}$$

Hence we are led to the characterization that if the PB of a time-independent dynamical variable ω with the Hamiltonian vanishes, then ω is conserved. The immediate corollary is that since the PB of a Hamiltonian with itself $\{H, H\}$ vanishes, H is a constant of the motion. The above conclusions are exclusive to time independent Hamiltonians since, of course, it is meaningless to ask for conservation laws for a time dependent function.

Of great importance in the sequel are the PB relations between the fundamental dynamical variables of the Hamiltonian formalism, the qs and the ps . We have the following identities:

$$\begin{aligned}
\{q_i, q_j\} &= \{p_i, p_j\} = 0 \\
\{q_i, p_j\} &= \delta_{ij}.
\end{aligned} \tag{109}$$

It is also interesting to note that Hamilton's equations themselves may be expressed in terms of the PB as:

$$\begin{aligned}
\dot{q}_i &= \{q_i, H\} \\
\dot{p}_i &= \{p_i, H\}.
\end{aligned} \tag{110}$$

4.3.3 Canonical Transformations

It was noted earlier that an advantage of the Lagrangian formalism is that the Euler-Lagrange equations are form invariant under a smooth(non-singular forward and inverse), arbitrary, change of co-ordinates. This idea of symmetry, or invariance under co-ordinate transforms has profound dynamical implications.

A case in point is the powerful result known as Noether's theorem which establishes a one-to-one correspondence between continuous symmetries and conservation laws. The result may be roughly stated as follows: Corresponding to every continuous symmetry of the Lagrangian (Hamiltonian), there exists a conserved physical quantity.

As an application, it can be shown that the three cardinal conservation laws of classical dynamics viz. the conservation of energy, linear momentum and angular momentum emerge as a consequence of the invariance of a standard Lagrangian under time translations, space translations and rotations respectively. It is interesting to note that Noether's theorem is the precise mathematical formulation of the simple physical idea that the laws of dynamics (indeed, the laws of physics in general) cannot change with respect to time, space or orientation. Certainly, universal physical laws should hold for a certain experiment that was conducted a century ago, repeated now, or a century hence!

Having touched upon the importance of symmetry in dynamics let us now look at canonical transformations. Consider a transformation of co-ordinates in the Hamiltonian formalism. Let $(q, p) \rightarrow (\bar{q}, \bar{p})$. If Hamilton's equations are satisfied in the new co-ordinate system, the transformation is called a canonical transformation.

The next question is, given a new set of co-ordinates, how can we tell whether they are the result of a canonical transformation?

The answer is provided by the following elegant set of necessary and sufficient conditions in the new co-ordinates. They are the commutation relations given by:

$$\begin{aligned} \{\bar{q}_i, \bar{q}_j\} &= \{\bar{p}_i, \bar{p}_j\} = 0 \\ \{\bar{q}_i, \bar{p}_j\} &= \delta_{ij} \end{aligned} \tag{111}$$

It is noteworthy that the above set of conditions make no reference either to the functional form of the Hamiltonian in the new co-ordinates or the form of the co-ordinate transformation itself.

An indication of the amount of structure that a canonical transformation preserves may be obtained from the observation that the Poisson Bracket between any two dynamical variables remains invariant under such a transformation. We have for two dynamical variables σ and ω ,

$$\{\omega, \sigma\}_{q,p} = \{\omega, \sigma\}_{\bar{q},\bar{p}} \tag{112}$$

With this, we conclude this brief review of some aspects of classical dynamics.

5 Problems with Classical Physics

An oft touted unique feature of quantum mechanics that captures the imagination is wave-particle duality. In this section we focus on the inadequacies of classical physics as exposed by certain experiments in the twentieth century. Attempts to explain the experimental results led to the radical revision of certain classical concepts and ultimately led to the framework of quantum mechanics. The wave-particle duality of radiation and matter emerged as an essential idea in this development.

5.1 Particles and Waves

Particles and waves are treated as distinct entities in classical mechanics. Let us examine some elementary facts about both from the classical perspective.

A particle is a localized bundle of energy and momentum that is described at any instant by the state variables (q, \dot{q}) in a given co-ordinate system. The state variables evolve in time in a manner dictated by the equations of motion. Given the existence and uniqueness of solutions to the EOM, knowledge of (q, \dot{q}) at any given instant in time guarantees that we can deduce (q, \dot{q}) for any future time.

A wave is a disturbance spread over space that is described by a wave function $\psi(\vec{r}, t)$ which characterizes the disturbance at the point \vec{r} at time t . The EOM of the wave is the classical PDE

$$\nabla^2 \psi = \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} \quad (113)$$

where c is the speed of the wave.

A class of waves of interest are plane waves that are periodic in space and time. A plane wave in one dimension may be written as

$$\begin{aligned} \psi(x, t) &= A \exp \left[i \left(\frac{2\pi}{\lambda} x - \frac{2\pi}{T} t \right) \right] \\ &= A \exp [i\phi]. \end{aligned} \quad (114)$$

At a given time t , the wave is periodic in space with a period λ called its *wavelength*. At a given point x , it is periodic in time, repeating itself every T seconds where T is called the *timeperiod*. Useful quantities associated with λ and T are $k = \frac{2\pi}{\lambda}$ called the *wavenumber* and $\omega = \frac{2\pi}{T}$ called the *frequency*. Moreover, ϕ is called the phase and k the *wavenumber* measures the phase change per unit length at any fixed time t . Also, ω the *frequency* measures phase change per unit time at a point x . The wave travels at a speed given by $v = \frac{\omega}{k}$. The scale A in front of the phase factor is called the *amplitude*. The *intensity* of a wave is defined as $I = |\psi|^2$. Indeed, for a plane wave, $I = |A|^2$. In physical terms, if ψ is an electromagnetic wave, the intensity is a measure of the energy and momentum carried by the wave.

In three dimensions plane waves are written as:

$$\psi(\vec{r}, t) = A \exp[i(\vec{k} \cdot \vec{r} - \omega t)], \quad \omega = |\vec{k}|v.$$

Here, \vec{k} is called the wave vector and each component k_i of \vec{k} gives the phase change per unit length along the i th axis.

5.2 A Classical Experiment with Waves and Particles

We now discuss the essential version of an experiment that highlights the inadequacy of classical mechanics and provides the rationale for the introduction of quantum ideas.

Let us first consider a stream of particles being produced from a source S as shown in the figure. The particles, in succession, pass through the slits $S1$ **or** $S2$ to reach a detector which records the intensity of arrival $I(x)$. Here the intensity $I(x)$ is defined as the number of particles arriving in unit time at a given point which is at a distance x as measured consistently from one end of the detector screen. The distributions $I1(x)$ and $I2(x)$ with **only** $S1$ or $S2$ open respectively, is shown in part (a) of the figure. Clearly, the intensity of arrival peaks at the respective slit (when the other slit is closed) since we expect most particles not to scatter in their journey from the slit to the detector. The interesting situation is when both slits are open simultaneously. Classical mechanics predicts categorically that the intensity in this case I is given by $I1 + I2$ and follows the pattern shown roughly in part (b) of the figure.

Now if we consider a stream of waves (say, light waves) under the same conditions and repeat the experiment with both slits open, we will observe the well known phenomenon of *interference*. In the framework of classical mechanics, the key fact that distinguishes the wave experiment from the particle experiment and characterizes the interference is that in the case of waves, the intensity I is **not** the sum of $I1$ and $I2$.

A series of experiments with light at the turn of the 20th century contradicted the above classical prediction in the following manner. Firstly, light beams of very low intensity produced interference patterns befitting a stream of particles. In other words, if we sufficiently control the light beam intensity and keep it low, we observe that the detector behaves as if it is being hit with a stream of particles. This led to the celebrated idea of light being a stream of **photons** or quanta of energy. However, we cannot begin to deal with photons using the classical mechanics of particles due to the following experimentally observed phenomenon. Consider the situation when we allow only one photon (a discrete bundle of energy, in other words) to emerge out of the source. Now, if both slits are open, the singleton photon produces an interference pattern in the detector! This leads us to the bizarre conclusion that

if we manage to control the source in such a way that it ejects only a single photon, that photon begins to behave like a wave.

Indeed, the square peg of classical mechanics was being confronted with the round hole of quantum phenomena.

5.3 The First Quantum Principles

The experimental facts mentioned in the previous sub-section led Max Born to postulate as follows: “with *each* photon is associated a wave ψ , called the *probability amplitude* whose modulus squared $|\psi(x)|^2$ gives the *probability* of finding the particle at x ”. In the light of this hypothesis, the results of the double slit experiment may be interpreted as follows: **every** incoming photon of energy E and momentum p has a wave function ψ associated with it. This wave function is a plane wave with $\omega = \frac{E}{\hbar}$ and $k = \frac{p}{\hbar}$ where $\hbar = \frac{h}{2\pi}$, h being the Planck’s constant. This wave interferes with itself and forms the pattern $|\psi(x)|^2$ along the detector, which gives the probability that the given photon will arrive at a point x . Here the consequences of a probabilistic description kick in. A given photon does not reveal the nature of the distribution nor does it even indicate the probabilistic nature of the process. However, if we prepare an ensemble of photons all with wavefunctions identically equal to ψ and wait for all of them to arrive at the detector, the number of arrived photons at any given x will be proportional to $|\psi(x)|^2$. On the other hand, if the intensity of the beam is increased then a number of photons with identical wave functions arrive simultaneously and are spread over the detector. The intensity distribution in this case right away assumes the shape of the probability distribution and the energy flow appears as continuous with the results in agreement with the classical predictions. In other words, the classical(macroscopic) limit is obtained from what is essentially a quantum phenomenon.

It bears repetition that associated with **every** photon is a wavefunction ψ . If the beam is monochromatic (of high intensity), then every photon is described by the same ψ . A large ensemble of such photons leads to the classically predicted results in experiments.

The next big step was taken by de Broglie who conjectured that just as in the case of photons, all matter has associated wave functions. Of course this idea was confirmed by a gamut of experiments in the sub-atomic domain and led to quantum mechanics being established

as an integral part of modern physics. The mesh with classical mechanics emerges from the fact that even though all of matter has wave aspects, the associated wave characteristics are extremely small for matter in the macroscopic realm. For example, a mass of 1 gm moving at 1 cm/sec has an associated wave length of 10^{-26} cm. Clearly, the double slit experiment conducted with a stream of such masses will be unable to detect interference effects arising from such miniscule wavelengths and classical results will obtain.

The prime conclusion from all the above statements is that all information about a quantum system is contained in ψ . The dynamics of the system is the dynamics of ψ and indeed ψ is treated as a ket. Drawing from the mathematical explorations of the first section and the experimental results mentioned in the present one, the next section deals with the set of fundamental postulates of quantum mechanics.

6 Framework for Quantum Mechanics - A Set of Postulates

The first part of this section provides a set of four fundamental postulates of quantum mechanics. In order for us to see the contrast and obtain a better perspective, the corresponding classical postulate (in the Hamiltonian formalism) is mentioned first.

C1: The state of a particle at any given time is fully specified by the position and momentum variables, $x(t)$ and $p(t)$. Hence, in geometric terms, the arena for the dynamics is a two dimensional phase space.

Q1: The state of a quantum particle is represented by a ket vector $|\psi(t)\rangle$ in a Hilbert space.

C2: Every dynamical variable ω is a function of x and p : $\omega = \omega(x, p)$.

Q2: The independent **variables** x and p of classical mechanics are replaced by Hermitian **operators** X and P with matrix elements in the eigenbasis of X given by:

$$\begin{aligned}\langle x|X|x'\rangle &= x\delta(x - x') \\ \langle x|P|x'\rangle &= -i\hbar\delta'(x - x').\end{aligned}$$

It follows that this anzatz of replacing the fundamental dynamical variables with operators

extends to all the classical variables. Thus the operator corresponding to a general ω is obtained by:

$$\Omega(X, P) = \omega(x \mapsto X, p \mapsto P).$$

For the sake of complete clarity, we elucidate the above equation as follows: The operator Ω is the same function of the operators X and P as the variable ω is of the variables x and p .

C3: If the particle is in a state given by x and p , the measurement of the variable ω yields a value $\omega(x, p)$. The accuracy of the measurement is limited only by the accuracy of the measuring apparatus. **The state of the system remains unaffected by the measurement.**

Q3: If the particle is in a state $|\psi\rangle$, measurement of the variable (corresponding to) Ω will yield one of the eigenvalues ω with probability $P(\omega) \propto |\langle \omega | \psi \rangle|^2$. **The state of the system changes from $|\psi\rangle$ to $|\omega\rangle$ as a result of the measurement.**

C4: The state variables change with time according to Hamilton's equations:

$$\begin{aligned} \dot{x} &= \frac{\partial \mathcal{H}}{\partial p} \\ \dot{p} &= -\frac{\partial \mathcal{H}}{\partial x}. \end{aligned}$$

Q4: The state vector $|\psi(t)\rangle$ obeys the Schrodinger equation:

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

Here, H is the quantum Hamiltonian operator corresponding to the classical Hamiltonian \mathcal{H} .

6.1 Discussion of Postulates 1 -3

The first postulate **Q1** states that a quantum particle is described by $|\psi\rangle$ in a physical Hilbert space (which contains proper vectors normalizable to unity and improper vectors normalizable only to Dirac delta functions). Now, by definition, $|\psi\rangle$ has an infinite number of components in a given basis and here we clearly see the stark difference between the quantum and classical descriptions of a particle. The classical phase space (of finite dimension) is replaced by an infinite dimensional Hilbert space. So, why do we require a "particle" to be described by a vector in an infinite dimensional space?

The answer is best seen in the light of the conclusions of the double-slit experiment. There we saw that a photon needed to be described by a wave **function**. Thus, the states of quantum objects are to be described by wave functions and we know from classical analysis that a vector space of functions is an infinite dimensional space. The Fourier expansion of any well behaved function in a (infinite dimensional) sine - cosine basis is a classic instance of this.

The next important point related to **Q1** is the implication of the *principle of superposition* in quantum mechanics. Since kets form a linear vector space, if $|\psi\rangle$ and $|\psi'\rangle$ represent two possible states of a particle, so does $\alpha|\psi\rangle + \beta|\psi'\rangle$ where α and β are arbitrary scalars. The superposition principle is characteristic of any linear system and is quite familiar from classical physics. For instance, if $f(x)$ and $g(x)$ are two possible displacements of a string, so is $\alpha f(x) + \beta g(x)$. However there **is** a major difference in the interpretation of superposition between the classical and quantum viewpoints. In the case of the string for instance, the state $\alpha f(x) + \beta g(x)$ will have physical properties (shape, elastic energy stored etc.) very different from both $f(x)$ and $g(x)$. In quantum mechanics, $\alpha|\psi\rangle + \beta|\psi'\rangle$ will have attributes that *sometimes* resemble $|\psi\rangle$ and *at other times* resemble $|\psi'\rangle$. This ambiguity is an inescapable aspect of the theory and forces a probabilistic description. This also brings to the fore, the important conclusion that *results of observations are critically dependent on measurement*. Precisely, if we make a measurement on a quantum system that is in a state $\alpha|\psi\rangle + \beta|\psi'\rangle$, the act of measurement will force the system to *either* $|\psi\rangle$ or $|\psi'\rangle$ and which of the two states will actually result can be determined only probabilistically. The mechanism of this process is described by **Q2** and **Q3** and let us examine this in detail.

In classical mechanics, if a state (x, p) of a particle is given, we say that a dynamical variable ω (such as energy or momentum) has a value $\omega(x, p)$ in the sense that if the variable is measured we will obtain the value $\omega(x, p)$. The answer to the question of what is the corresponding statement in QM for a particle in a state $|\psi\rangle$ is provided by **Q2** and **Q3** in the following steps:

Step 1: Construct the corresponding quantum operator Ω as described in **Q2**.

Step 2: Find the eigenvalues ω_i and eigenvectors $|\omega_i\rangle$ of Ω in an orthonormal basis.

Step 3: Expand $|\psi\rangle$ in this basis: $|\psi\rangle = \sum_i |\omega_i\rangle \langle \omega_i | \psi \rangle$.

Step 4: The *probability* $P(\omega)$ that the result ω will obtain is proportional to the modulus squared of the projection of $|\psi\rangle$ along the eigenvector $|\omega\rangle$. ie. $P(\omega) \propto |\langle \omega | \psi \rangle|^2$.

Let us note the salient points that emerge from the above steps.

(1) There exist only probabilistic predictions for the result of a measurement of Ω . Furthermore, **the only possible values of Ω are its eigenvalues**. Since **Q2** demands that Ω be Hermitian, the eigenvalues are all real. This is how quantum mechanics ensures that all predicted observable values are real.

(2) The prescription $P(\omega) \propto |\langle \omega | \psi \rangle|^2$ implies that $|\langle \omega | \psi \rangle|^2$ is a relative probability. The absolute probability may be obtained as:

$$P(\omega) = \frac{|\langle \omega | \psi \rangle|^2}{\langle \psi | \psi \rangle}.$$

There is a subtle but important point here. Due to the normalization involved in obtaining the absolute probability, the relative probability distributions corresponding to the apparently distinct states $|\psi\rangle$ and $\alpha|\psi\rangle$ where α is a scalar, will lead to the same absolute probability distributions. Thus, corresponding to each physical state, there exists not one vector but a ‘ray’ or direction in Hilbert space. In other words, $|\psi\rangle$ and $\alpha|\psi\rangle$ are entirely equivalent in describing the physical state. This reminds us of the freedom we have in parallel translating vectors in a 2-d Cartesian space where a given vector is indistinguishable from any other vector of the same magnitude and direction in the space.

In QM this means that even with unit norm ($\langle \psi | \psi \rangle = 1$) we have the freedom to multiply a ket $|\psi\rangle$ by an arbitrary scalar of the form $e^{i\theta}$ without altering the physical state. Clearly, no such freedom exists in classical mechanics.

(3) It also follows from **Q2** and **Q3** that if $|\psi\rangle$ is an eigenstate $|\omega_i\rangle$ of the operator Ω **to begin with**, then the measurement of Ω is guaranteed to deliver the result ω_i . This is the only instance where a quantum measurement can be interpreted in a classical sense.

(4) For the sake of completeness, we discuss the superposition principle once more here. Let $|\omega_1\rangle$ and $|\omega_2\rangle$ be superposed to form a (normalized) state such as:

$$|\psi\rangle = \frac{\alpha|\omega_1\rangle + \beta|\omega_2\rangle}{(|\alpha|^2 + |\beta|^2)^{\frac{1}{2}}}.$$

Upon measurement of Ω we obtain either ω_1 or ω_2 with probabilities $\frac{|\alpha|^2}{(|\alpha|^2 + |\beta|^2)}$ and $\frac{|\beta|^2}{(|\alpha|^2 + |\beta|^2)}$ respectively. It bears repetition that this situation has no analog in classical mechanics.

(5) In order to extract information about any other observable Λ we repeat the whole process, finding the eigenspectrum of Λ . The whole point here is that in $|\psi\rangle$ is encoded all obtainable statistical information about the particle. To obtain this information with respect to a particular observable, we determine the eigenbasis of the corresponding operator and find the projection of $|\psi\rangle$ along all the eigenkets of the operator.

(6) There exists a natural way of going from the basis of one operator to another. Suppose we are working in the basis of the operator Ω in which:

$$|\psi\rangle = \sum_i |\omega_i\rangle \langle \omega_i | \psi \rangle .$$

We can move to the basis of another operator Λ directly as follows:

$$\langle \lambda_i | \psi \rangle = \sum_j \langle \lambda_i | \omega_j \rangle \langle \omega_j | \psi \rangle .$$

Thus all it takes to transform to the basis of a different operator Λ is to compute the scalars $\langle \lambda_i | \omega_j \rangle$.

To wind up our discussion on the first three postulates we mention four complications that can arise with our structure.

Complication 1: The Recipe $\Omega(X, P) = \omega(x \mapsto X, p \mapsto P)$ Is Ambiguous. In some cases there can arise ambiguity in the prescription to generate the quantum operator from the definition of the corresponding classical variable. For instance, if $\omega = xp$ it is not clear whether $\Omega = XP$ or $\Omega = PX$. The issue is that the classical variables commute while the operators do not. In this particular case, the resolution is to apply symmetrization and use $\Omega = \frac{XP+PX}{2}$. However, in a situation where symmetrization fails to resolve the ambiguity only experiments can guide the correct choice of the appropriate quantum operators.

Complication 2: The Operator Ω Is Degenerate. In this case Ω has repeated eigenvalues and in **Q3** the term for $P(\omega)$ gets replaced as $P(\omega) \propto | \langle \psi | \mathbb{P}_\omega | \psi \rangle |$ where \mathbb{P}_ω is the projection operator for the eigenspace with eigenvalue ω . With this **Q3** as stated originally becomes a special case for a non-degenerate operator.

Complication 3: The Eigenspectrum of Ω Is Continuous. Here the eigenvalues become continuous functions and the expansion of $|\psi\rangle$ in the eigenbasis of Ω becomes:

$$|\psi\rangle = \int |\omega\rangle \langle \omega | \psi \rangle d\omega .$$

The situation is the exact parallel of the transition to continuous random variables from discrete ones in probability theory and accordingly, we now interpret $P(\omega) = |\langle \omega | \psi \rangle|^2$ to be the **probability density** at ω . Indeed,

$$\int P(\omega) d\omega = \int |\langle \omega | \psi \rangle|^2 d\omega \tag{115}$$

$$= \int \langle \psi | \omega \rangle \langle \omega | \psi \rangle d\omega \tag{116}$$

$$= \langle \psi | I | \psi \rangle = \langle \psi | \psi \rangle = 1. \tag{117}$$

Complication 4: The Quantum Operator Ω Has No Classical Counterpart. There have been observed uniquely quantum phenomena like “spin” of particles that have no classical analogues whatsoever. When dealt with such situations the usual prescriptions for constructing quantum operators breaks down. The operators are constructed using intuition and semi-classical reasoning. In other words, QM is far from a complete theory and is more like a framework or a set of guidelines on how to deal with objects at the quantum level.

7 Measurements and Ensembles

From our discussion thus far, it is evident that the process of measurement and its implications are fundamentally different between classical and quantum mechanics. This section deals with this aspect as well as how this necessitates the use of ensembles for quantum measurements.

We recall from **Q3** that measurement of the variable Ω changes the state $|\psi\rangle$ to the eigenstate $|\omega\rangle$ corresponding to the eigenvalue ω obtained as a result of the measurement. This phenomenon is known as the *collapse or reduction of the state vector*.

The point to note here is that this collapse is unavoidable even in an ideal quantum measurement, whereas, in class mech. it is possible, *in principle*, to obtain the perfect measurement.

Given this state of affairs, how does one test QM? As a theory QM would be of little value without predictive power. Given the framework of QM, in order to test it, we must be able to:

- (1) Create particles in a well defined state $|\psi\rangle$.

(2) Check the probabilistic predictions at any time.

Let us now consider, in simplest terms, how this can be achieved. We begin with a particle in an arbitrary state $|\psi\rangle$ (unknown to us) and measure a variable Ω . If we get a non-degenerate eigenvalue ω , we are assured that we have in our hands the state $|\omega\rangle$ (which we know since we can solve the eigenvalue problem of Ω). It is important to note here that we had no information on $|\psi\rangle$ to begin with. Thus QM can tell us nothing about the state that we begin with but can make statements about the state that is obtained as a result of the measurement.

So now we have prepared the state $|\omega\rangle$. Let us now immediately measure another variable Ω . Indeed, we should be able to express ω as a linear combination of the eigenbasis of Ω and assume for the moment:

$$|\omega\rangle = \frac{1}{\sqrt{3}}|\lambda_1\rangle + \sqrt{\frac{2}{3}}|\lambda_2\rangle + 0 \cdot (\text{other eigenvectors})$$

The prediction of QM is that we should obtain either the eigenvalue λ_1 with probability $\frac{1}{3}$ **or** λ_2 with probability $\frac{2}{3}$. Clearly, if we obtain any other value, QM would be incorrect.

Now, even if we obtain λ_1 or λ_2 that would be insufficient to validate QM. Since probabilities are assigned for obtaining λ_1 or λ_2 , we need to repeat the measurement a large number of times to establish these probabilities (as in any statistical situation seen from a frequency point of view). However, since we seem to have little control over our states, in general, it is impossible to repeat the measurement on the one state that we have prepared.

This introduces the concept of an ensemble. Thus, any statement about the quantum measurement of a state implicitly assumes that we have carried out the measurement on an array of identically prepared states - in other words, an ensemble. Thus, an ensemble comprises a large number N of particles all in the same state, say, $|\omega\rangle$.

The concept of an ensemble is not exclusive to quantum mechanics. The idea is used very effectively in classical statistical mechanics, for instance. However the big difference between a classical ensemble and a quantum one is the following: Suppose, as a result of measuring an N particle ensemble for some quantity λ , we find that $N/3$ yield the value $\lambda = \lambda_1$ and $2N/3$ yield the value $\lambda = \lambda_2$. Now, if this were a classical ensemble it is acceptable (and natural) to conclude that $N/3$ of the particles were associated with $\lambda = \lambda_1$ **before** the measurement.

From the quantum point of view, of course, all the N particles were in identical states all represented by some $|\omega\rangle$ **before** the measurement and $N/3$ ended up in the state $|\lambda_1\rangle$ **as a result of** the measurement.

Since the ensemble is a large collection of identical particles, any observable associated with the ensemble can be treated as a random variable. It is then natural to inquire about the two important quantities associated with any random variable : the expectation (mean) value and the standard deviation.

7.1 Expectation (Mean) Value

Let Ω be an observable with eigenvalues ω_i and probabilities of obtaining the eigenvalues given by $P(\omega_i)$. As usual, we assume that we start off with a state $|\psi\rangle$ and measure Ω . The expectation value of Ω can be written from classical probability theory as:

$$\langle\Omega\rangle = \sum_i P(\omega_i) \omega_i \quad (118)$$

Recall that the probability of obtaining the state $|\omega_i\rangle$ starting from a state $|\psi\rangle$ is given by:

$$P(\omega_i) = |\langle\omega_i|\psi\rangle|^2 \quad (119)$$

Substituting Eqn 119 into Eqn 118 we can write:

$$\begin{aligned} \langle\Omega\rangle &= \sum_i |\langle\omega_i|\psi\rangle|^2 \omega_i \\ &= \sum_i \langle\psi|\omega_i\rangle \langle\omega_i|\psi\rangle \omega_i \end{aligned} \quad (120)$$

Using the relation $\Omega|\omega_i\rangle = \omega_i|\omega_i\rangle$ in the above equation we now write:

$$\langle\Omega\rangle = \sum_i \langle\psi|\Omega|\omega_i\rangle \langle\omega_i|\psi\rangle \quad (121)$$

We now recall from our discussion on the fundamental properties of eigen kets that :

$$|\omega_i\rangle\langle\omega_i| = I. \quad (122)$$

Substituting 122 in 121,

$$\langle\Omega\rangle = \langle\psi|\Omega|\psi\rangle. \quad (123)$$

The interesting point here is that in order to find the expectation value $\langle \Omega \rangle$ of an observable Ω over an ensemble, we **need not** solve the eigenvalue problem of Ω . Given the state of the ensemble $|\psi\rangle$ (indeed, all particles in the ensemble are assumed to be in identical states) and the representation of Ω in some basis, we can directly compute the expectation value.

7.2 The Uncertainty (Standard Deviation)

The other useful quantity associated with a random variable is the standard deviation which measures the average fluctuations of the random variable about the mean value. This precisely is what is called the **uncertainty** in the measurement of an observable in quantum mechanics. For an observable Ω the uncertainty is defined as:

$$\Delta\Omega = \langle (\Omega - \langle \Omega \rangle)^2 \rangle^{\frac{1}{2}} \quad (124)$$

If Ω has a discrete set of eigenvalues the variance (square of the standard deviation) can be written in terms of the probabilities of occurrence of each of the eigenvalues and the expectation value of Ω as:

$$(\Delta\Omega)^2 = \sum_i P(\omega_i) (\omega_i - \langle \Omega \rangle)^2 \quad (125)$$

In the case of a continuous spectrum for Ω , this becomes:

$$(\Delta\Omega)^2 = \int P(\omega) (\omega - \langle \Omega \rangle)^2 d\omega \quad (126)$$

As in the case of the expectation value, one does not need to know the eigenspectrum of Ω in order to calculate the uncertainty in a measurement of Ω .

7.3 Compatible and Incompatible Observables

From our discussion thus far, we are clear on the following point. If we make a measurement of an observable Ω on a particle in an arbitrary state $|\psi\rangle$, the particle ends up in an eigenstate $|\omega\rangle$ of Ω . The probability of this happenstance is given by $P(\omega) = |\langle \omega | \psi \rangle|^2$. Also, this provides us with a way of producing a particle in a state $|\omega\rangle$. All we need to do is make a measurement of Ω on the particle.

This leads to the question of multiple observables. We ask:

Can we take a particle in an arbitrary state $|\psi\rangle$ and produce a state with simultaneously well defined eigenvalues ω and λ for two observables Ω and Λ ?

The answer is yes if and only if the commutator bracket $[\Omega, \Lambda]$ (please see Eqn 22 for the definition) between the operators Ω and Λ vanishes and they do not have degenerate eigenspectra. The operators are then called compatible.

Importantly, the answer is no if the commutator bracket does not vanish. In other words, there cannot exist a state of a quantum particle with definite eigenvalues for two observables if the commutator bracket between the corresponding operators of the observables does not vanish. This is the origin of the Heisenberg Uncertainty Principle, the famous example of which is the canonical commutation relation between the position operator X and momentum operator P for a particle. The commutation rule is:

$$[X, P] = i\hbar. \quad (127)$$

Here $\hbar = h/2\pi$ where h is the Planck's constant and $i = \sqrt{-1}$. Since neither of these quantities is zero, the RHS of Eqn 127 cannot vanish and this implies that there does not exist a single state for which both X and P are **simultaneously** well defined.

In order to see how the uncertainty principle arises naturally from the formalism, let us conduct the following exercise. Consider an infinite-dimensional ket $|\psi\rangle$ expanded in the basis $|x\rangle$ of the position operator X . We write:

$$\begin{aligned} |\psi\rangle &= \int |x\rangle \langle x|\psi\rangle dx \\ &= \int |x\rangle \psi(x) dx. \end{aligned} \quad (128)$$

where we have used the result $\langle x|\psi\rangle = \psi(x)$ based on our discussion in Section 3.5. In an infinite-dimensional space, the inner product of a ket $|\psi\rangle$ with the basis vectors in the X basis just yields the function $\psi(x)$.

We call $\psi(x)$ the *wave function* in the X basis. Let us assume that the wave function is a Gaussian given by:

$$\psi(x) = A \exp \left[\frac{-(x-a)^2}{2\Delta^2} \right]. \quad (129)$$

In order to normalize the state, we write the inner product of $|\psi\rangle$ with itself as:

$$\begin{aligned}
 \langle \psi | \psi \rangle &= \int_{-\infty}^{\infty} \langle \psi | x \rangle \langle x | \psi \rangle dx \\
 &= \int_{-\infty}^{\infty} |\psi(x)|^2 dx \\
 &= \int_{-\infty}^{\infty} A^2 \exp \frac{-(x-a)^2}{\Delta^2} dx \\
 &= A^2 (\pi \Delta^2)^{\frac{1}{2}}.
 \end{aligned} \tag{130}$$

Hence, the normalized state can be written as:

$$\begin{aligned}
 \psi(x) &= \frac{1}{\sqrt{\langle \psi | \psi \rangle}} A \exp \left[\frac{-(x-a)^2}{2\Delta^2} \right] \\
 &= \frac{1}{(\pi \Delta^2)^{\frac{1}{4}}} \exp \left[\frac{-(x-a)^2}{2\Delta^2} \right].
 \end{aligned} \tag{131}$$

Now, from the probability interpretation of the wave function, it follows from eqn 131 that the probability of finding the particle between x and $x + dx$ is:

$$\begin{aligned}
 P(x) dx &= |\psi(x)|^2 dx \\
 &= \frac{1}{(\pi \Delta^2)^{\frac{1}{2}}} \exp \left[\frac{-(x-a)^2}{\Delta^2} \right] dx.
 \end{aligned} \tag{132}$$

It is clear from eqn 132 that the particle is to be most likely found around $x = a$ (the mean of the distribution) and that the probability of finding the particle drops off rapidly beyond a spread of Δ (the standard deviation). This statement can be quantified by computing the expectation value and the uncertainty for the corresponding operator X . We now proceed with this exercise.

It follows from our discussion in Section 3.5 that projecting $|\psi\rangle$ onto the X basis we can write:

$$\langle x | \psi \rangle = \psi(x). \tag{133}$$

Hence we obtain:

$$\begin{aligned}
\langle x|X|\psi \rangle &= \int_{-\infty}^{\infty} \langle x|X|x' \rangle \langle x'|\psi \rangle dx' \\
&= \int_{-\infty}^{\infty} x\delta(x-x')\psi(x') dx' \\
&= x\psi(x).
\end{aligned} \tag{134}$$

Using eqn 131 and eqn 134 the mean value of X can be computed as:

$$\begin{aligned}
\langle X \rangle &= \langle \psi|X|\psi \rangle \\
&= \int_{-\infty}^{\infty} \langle \psi|x \rangle \langle x|X|\psi \rangle dx \\
&= \int_{-\infty}^{\infty} \psi^*(x) x\psi(x) dx \\
&= \frac{1}{(\pi\Delta^2)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-a)^2}{\Delta^2}\right] x dx.
\end{aligned} \tag{135}$$

The above integral can be evaluated to give:

$$\langle X \rangle = a. \tag{136}$$

Next we calculate the uncertainty associated with X . The definition of uncertainty associated with an observable Ω is given by eqn 124. Hence we can write the uncertainty in X as the operator $\Delta X = \langle (X - \langle X \rangle)^2 \rangle^{\frac{1}{2}}$. Given our ket $|\psi \rangle$, the uncertainty in measuring X can be computed as:

$$\begin{aligned}
\Delta X &= [\langle \psi|(X - \langle X \rangle)^2|\psi \rangle]^{\frac{1}{2}} \\
&= [\langle \psi|X^2 - 2X\langle X \rangle + \langle X \rangle^2|\psi \rangle]^{\frac{1}{2}}.
\end{aligned} \tag{137}$$

However, we have :

$$\begin{aligned}
\langle \psi|-2X\langle X \rangle|\psi \rangle &= -2\langle X \rangle \langle \psi|X|\psi \rangle \\
&= -2\langle X \rangle \langle X \rangle \text{ (since } \langle \psi|X|\psi \rangle = \langle X \rangle) \\
&= -2\langle X \rangle^2.
\end{aligned} \tag{138}$$

Substituting eqn 138 into eqn 137 and using eqn 136 we can write:

$$\begin{aligned}
\Delta X &= [\langle \psi | X^2 - \langle X \rangle^2 | \psi \rangle]^{\frac{1}{2}} \\
&= [\langle \psi | X^2 - a^2 | \psi \rangle]^{\frac{1}{2}} \\
&= [\langle X^2 \rangle - a^2]^{\frac{1}{2}}.
\end{aligned} \tag{139}$$

Invoking our assumption of the explicit form of $|\psi \rangle$ as a Gaussian (eqn 131) and using the result corresponding to eqn 135 we write:

$$\begin{aligned}
\langle X^2 \rangle &= \langle \psi | X^2 | \psi \rangle \\
&= \int_{-\infty}^{\infty} \psi^*(x^2) x \psi(x) dx \\
&= \frac{1}{(\pi \Delta^2)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-a)^2}{\Delta^2}\right] x^2 \exp\left[-\frac{(x-a)^2}{\Delta^2}\right] dx \\
&= \frac{\Delta^2}{2} + a^2.
\end{aligned} \tag{140}$$

Substituting eqn 140 into eqn 139 we can finally write:

$$\Delta X = \frac{\Delta}{\sqrt{2}}. \tag{141}$$

Carrying out a similar exercise for the momentum operator P acting on $|\psi \rangle$ in the x basis, it can be shown that the uncertainty in P for the same $|\psi \rangle$ is :

$$\Delta P = \frac{\hbar}{\sqrt{2}\Delta}. \tag{142}$$

From eqns 141 and 142 we obtain the important result:

$$\Delta X \Delta P = \frac{\hbar}{2}. \tag{143}$$

The above equation is in fact the lower bound of the uncertainty principle. The equality obtains due to the fact that we have assumed our state $|\psi \rangle$ to be Gaussian. It is clear here that it is impossible to eliminate uncertainties in both the position X and the momentum P simultaneously since \hbar the Planck's constant is non-zero. Thus a measurement which aims for accuracy in X implies that the error (uncertainty) in P is large and vice versa.

8 The Schrodinger Equation

The discussion of the Schrodinger equation takes off from the postulate **Q4** in Section 6. The fourth postulate suggests that the time evolution of a quantum state obeys the Schrodinger equation. In other words, the equation(s) of motion of a quantum state is given by:

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H |\psi(t)\rangle \quad (144)$$

where H is the system hamiltonian.

The setting up of the equation is a relatively straightforward exercise if the corresponding classical hamiltonian is known. For example, if one considers the classical harmonic oscillator with the hamiltonian:

$$\mathcal{H} = \frac{p^2}{2m} + \frac{1}{2} m \omega^2 x^2 \quad (145)$$

the quantum counterpart is obtained by replacing the variables p and x with the corresponding operators discussed in detail in the previous sections. One obtains the quantum hamiltonian as:

$$H = \frac{P^2}{2m} + \frac{1}{2} m \omega^2 X^2. \quad (146)$$

8.1 General Approach to the Solution

8.1.1 The Time-Independent Hamiltonian Case

If the hamiltonian H is time-independent, the problem of solving the Schrodinger equation is very similar to the problem considered in Section 3.4.1. As in that case, we seek to construct the propagator $U(t)$ from the eigenvectors and eigenvalues of H . The solution can then be written as:

$$|\psi(t)\rangle = U(t) |\psi(0)\rangle. \quad (147)$$

The question now is the one of solving the eigenvalue problem for H . Let $|E\rangle$ be the normalized eigenkets of H . By definition:

$$H |E\rangle = E |E\rangle. \quad (148)$$

This is known as the time-independent Schrodinger equation. Once we have solved the eigenvalue problem for H we can express $|\psi\rangle$ in that eigenbasis as:

$$\begin{aligned} |\psi(t)\rangle &= \sum |E\rangle \langle E|\psi(t)\rangle \\ &= \sum a_E(t) |E\rangle. \end{aligned} \quad (149)$$

where we have defined $\langle E|\psi(t)\rangle = a_E(t)$. In order to extract a_E , we now act on both sides of eqn 149 with the operator $(i\hbar\frac{\partial}{\partial t} - H)$. We obtain:

$$\begin{aligned} \left(i\hbar\frac{\partial}{\partial t} - H\right) |\psi\rangle &= \sum \left(i\hbar\frac{\partial}{\partial t} - H\right) a_E(t) |E\rangle \\ &= (i\hbar\dot{a}_E - E a_E) |E\rangle \quad (\text{since } H|E\rangle = E|E\rangle). \end{aligned} \quad (150)$$

Now, the LHS of eqn 150 vanishes since the operator $(i\hbar\frac{\partial}{\partial t} - H)$ acting on $|\psi\rangle$ is nothing but a statement of the Schrodinger equation itself. Hence:

$$i\hbar\frac{d}{dt}a_E = E a_E. \quad (151)$$

It follows that:

$$a_E(t) = a_E(0) \exp\frac{-iEt}{\hbar}. \quad (152)$$

An important observation is that eqns 149 and 152 allow us to write:

$$\langle E|\psi(t)\rangle = \langle E|\psi(0)\rangle \exp\frac{-iEt}{\hbar}. \quad (153)$$

Finally, we can write the time evolution of $|\psi(t)\rangle$ as:

$$|\psi(t)\rangle = \sum |E\rangle \langle E|\psi(0)\rangle \exp\frac{-iEt}{\hbar}. \quad (154)$$

Once we recall that the propagator $U(t)$ is nothing but an operator that marches an initial state $|\psi(0)\rangle$ in time, we can extract the former from eqn 154 to write:

$$U(t) = \sum |E\rangle \langle E| \exp\frac{-iEt}{\hbar}. \quad (155)$$

A point here it that the normal modes given by:

$$|E(t)\rangle = |E\rangle \exp\frac{-iEt}{\hbar} \quad (156)$$

are called stationary states. They derive their name and are unique due to the fact that the probability distribution $P(\omega)$ for any random variable Ω is time-independent for such states. This can be seen as follows. Using our basic definition for the probability of obtaining a state $|\omega\rangle$ from a state $|\psi\rangle$ (see Section 6.1), for a stationary state we can write:

$$\begin{aligned}
 P(\omega, t) &= |\langle \omega | \psi(t) \rangle|^2 \\
 &= |\langle \omega | E(t) \rangle|^2 \\
 &= |\langle \omega | E \rangle \left(\exp \frac{-iEt}{\hbar} \right)|^2 \\
 &= |\langle \omega | E \rangle|^2 \\
 &= P(\omega, 0).
 \end{aligned} \tag{157}$$

Here we have used the fact that $\exp \frac{-iEt}{\hbar}$ is a complex number with modulus unity.

It is also interesting to note that the solution to the Schrodinger equation can be written down in one step if one were to view eqn 144 as an operator equation. Since H is an operator in Hilbert space it can be expected to have a matrix representation in a given basis for the space. Therefore, the propagator is nothing but the matrix exponential of H and can be readily written as:

$$\mathbf{U}(t) = \exp \frac{-i\mathbf{H}t}{\hbar} \tag{158}$$

Of course, the caveat here is that the operator (matrix) exponential series on the RHS of eqn 158 must be convergent.

We conclude this subsection with the following important observations. Since, by assumption, the hamiltonian H is a Hermitian operator, it follows that the propagator $U(t)$ is a unitary operator. From our discussion in Section 3, it then follows that the action of $U(t)$ may be viewed as a *rotation* in Hilbert space. In analogy with rotations in \mathbf{R}^n which leave the inner product between any two vectors invariant, we can then conclude that unitary evolution governed by the Schrodinger equation leaves the norm $\langle \psi | \psi \rangle$ invariant. Therefore, a normalized state remains normalized under unitary evolution.

The other consequence of viewing the time evolution of a quantum state as a rotation in Hilbert space is that one can choose a basis which “rotates” at the same rate as the evolving states. The states appear frozen in such a basis but now the operators (that were constant

matrices in the fixed basis) will now be time dependent. However, physical quantities will yield the same values as before in this picture since they are all associated with inner products that are invariant under rotations. This point of view wherein the operators are treated as time-dependent quantities is known as the *Heisenberg picture*. It is entirely equivalent to the treatment of the states as being time dependent, which is called the *Schrodinger picture*.

8.1.2 Time-Dependent Hamiltonian Case

There exists no specific strategy for generating solution when the hamiltonian is time-dependent. In situations where the hamiltonian can be split into two parts, say, $H(t) = H^0 + H^1(t)$ where H^0 is the dominating time-independent term, approximation schemes guided by perturbation theory can be used. The key point here is that in this case evaluating the propagator involves integrating the time-dependent hamiltonian over time in the matrix exponential. The non-commutativity of the hamiltonians at any two instants of time render this a delicate exercise and necessitate the introduction of a *time-ordered integral*. We do not delve further into this complication.

However, all the characteristics of unitary evolution survive this transition to the time-dependent case. For instance, the propagator satisfies the transitivity property:

$$U(t_3, t_2)U(t_2, t_1) = U(t_3, t_1). \quad (159)$$

Also, we have the following feature:

$$U^\dagger(t_2, t_1) = U^{-1}(t_2, t_1) = U(t_1, t_2). \quad (160)$$

We conclude this section with the note that the Schrodinger equation is usually solved by choosing a particular basis, commonly the X or P basis. Projecting the hamiltonian and the time derivative operator onto an appropriate basis, one ends up with a differential equation to be solved. Two one-dimensional examples are taken up in the next section.

9 Problems in One Dimension

9.1 The Free Particle

The simplest one dimensional problem is that of a free particle by which we mean a particle not moving under the influence of any potential. Consequently, the hamiltonian in this case has only contributions from the kinetic energy and can be written as:

$$H = \frac{P^2}{2m}.$$

Hence the Schrodinger equation is:

$$\begin{aligned} i\hbar \frac{d}{dt} |\psi\rangle &= H |\psi\rangle \\ &= \frac{P^2}{2m} |\psi\rangle. \end{aligned} \tag{161}$$

The solution may be written in the form:

$$|\psi\rangle = |E\rangle e^{-iEt/\hbar}. \tag{162}$$

We note that the form of the above solution, from the point of view of PDE solutions, is a consequence of assuming separation of variables and solving for the time dependent part. Now substituting Eqn 162 into Eqn 161 we obtain:

$$\begin{aligned} H|E\rangle &= E|E\rangle \\ &= \frac{P^2}{2m}|E\rangle \\ &= \frac{p^2}{2m}|E\rangle. \end{aligned} \tag{163}$$

It follows from Eqn 163 that :

$$\frac{p^2}{2m} - E = 0. \tag{164}$$

Hence the two allowed values for the momentum p are given by:

$$p = \pm\sqrt{2mE}. \tag{165}$$

Therefore, associated with each value of energy E , there exist two eigenstates given by:

$$|E, + \rangle = |p = \sqrt{2mE} \rangle \quad (166)$$

$$|E, - \rangle = |p = -\sqrt{2mE} \rangle . \quad (167)$$

Physically, the result is clear. In the one dimensional case, a particle with energy E can move either to the left or to the right with the same magnitude of the momentum $p = \sqrt{2mE}$. This much is identical to the situation in classical mechanics. However, in this quantum case, any state that is a linear combination of the eigenstates given by Eqn166 is **also** a solution to the problem! In other words, a particle of energy E has equal probability of being found moving in either direction with the same momentum.

9.2 Particle in a Box

This is a simple example where one clearly sees how energy emerges as a quantized dynamical variable. Consider the one dimensional motion of a particle confined to a box. The box naturally divides space into three regions as depicted in the following figure.

We will assume the following potential for the motion:

$$\begin{aligned} V(x) &= 0, & |x| < \frac{L}{2} \\ V(x) &= \infty, & |x| \geq \frac{L}{2}. \end{aligned} \tag{168}$$

Thus, the particle is confined to the region $-L/2 < x < L/2$ since it is confronted with an infinite potential barrier at the boundaries of the above region. Since the potential is time-independent, the wave function ψ (more precisely, the state $|\psi\rangle$) satisfies the time-independent Schrodinger equation written in the X basis given by:

$$\frac{d^2}{dx^2}\psi + \frac{2m}{\hbar^2}(E - V)\psi = 0. \tag{169}$$

We recognize that, from a mathematical standpoint, we ought to consider the solutions of Eqn169 for all regions in space (the regions represented by 1, 2 and 3 in the figure). However, the potential confines the particle to the box and hence it surely cannot be found in regions 1 or 3 at any time. Hence we may conclude that $\psi_1 = 0 = \psi_3$ where the subscripts identify the spatial region under consideration. We note here that the above heuristic argument can be rigorously established as well by actually solving Eqn 169 in regions 1 and 3 and then examining the solution in the limit $V \rightarrow \infty$.

Hence we only need to solve Eqn 169 in region 2. However, our solution ought to satisfy the boundary conditions given by $\psi_1 = 0 = \psi_3$. This is the crucial point which leads to quantization of energy.

In region 2 (ie. inside the box), since $V = 0$, Eqn 169 reduces to:

$$\frac{d^2}{dx^2}\psi + \frac{2m}{\hbar^2}E\psi = 0. \tag{170}$$

The solution may be written down in terms of the arbitrary constants A and B as:

$$\psi = A \exp(ikx) + B \exp(-ikx), \quad \text{where, } k = \sqrt{\frac{2mE}{\hbar^2}}. \tag{171}$$

Applying the boundary conditions that ψ vanishes at $x = \pm L/2$, we obtain the following set of equations, written in matrix form as:

$$\begin{pmatrix} e^{-ikL/2} & e^{ikL/2} \\ e^{ikL/2} & e^{-ikL/2} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{172}$$

The system of equations 172 has non-trivial solutions iff the determinant of the coefficient matrix vanishes. This leads to the condition:

$$k_n = \frac{n\pi}{L}, \quad n = 0, \pm 1, \pm 2, \dots \quad (173)$$

Hence, the solutions (eigenfunctions) can be written as:

$$\psi_n = \sqrt{\frac{2}{L}} \cos\left(\frac{n\pi x}{L}\right), \quad n = 1, 3, 5, 7, \dots \quad (174)$$

$$= \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right), \quad n = 2, 4, 6, 8, \dots \quad (175)$$

We note from the above that the trivial solution corresponding to $n = 0$ is unphysical since the wave function vanishes in this case and hence cannot represent the particle in the box.

Now, on substituting Eqn 173 into Eqn 171 we get the remarkable relation for the energy of the particle corresponding to the state n as:

$$E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{\hbar^2 \pi^2 n^2}{2mL^2}. \quad (176)$$

The above is a *quantization* condition for energy of the particle in the box. The key point is that this assertion that the admissible solutions to the Schrodinger equation correspond to fixed energy states is a consequence of the boundary conditions imposed on the solutions. In other words, if we require the particle to be bound within the box then the Schrodinger equation demands that the particle can exist only in certain energy states. This is in stark contrast to the situation in classical mechanics where such a particle restricted to move within a box can have any value for the energy.

Another important consequence of Eqn 176 is that the *ground state* of the particle (corresponding to $n = 1$) is not a state of rest (zero energy). This is also a radical point of departure from classical mechanics.

10 The Harmonic Oscillator

The harmonic oscillator is a quintessential example since it represents almost all periodic motion in the limit of small amplitudes. The problem is to solve quantum mechanically for the motion of an oscillator (say, a particle executing simple harmonic motion). This implies

that we need to solve for the state $|\psi\rangle$ of the oscillator that satisfies the Schrodinger equation:

$$i\hbar\frac{d}{dt}|\psi\rangle = H|\psi\rangle. \quad (177)$$

where H is the hamiltonian **operator** for the quantum harmonic oscillator (see Eqn 146) given by:

$$H = \frac{P^2}{2m} + \frac{1}{2}m\omega^2 X^2. \quad (178)$$

As has been repeatedly observed before, the idea is to obtain the propagator $U(t)$ for the problem since all dynamical information is contained in it. Indeed, $U(t)$ can be constructed once the eigenvalue problem for H is solved. The usual procedure in any problem is to project the abstract operator equation 177 onto the X basis, thereby obtaining an ordinary differential equation in x . Solving this DE and imposing appropriate boundary conditions on the solutions reveals the unique features of quantum dynamics.

In the case of the harmonic oscillator, there exists a clever algebraic technique introduced by Dirac that provides the solution to the problem without actually solving the Schrodinger equation! Apart from its sheer mathematical elegance the approach provides deep insights into the structure of quantum mechanics which have been crucial to further progress, for instance in generalizations to quantum field theories. We discuss the Dirac operator approach to the harmonic oscillator in the next subsection.

10.1 The Dirac Approach in the Energy Basis

We first establish that the eigenvalues of hamiltonian (energy) operator H cannot be negative. Given the form of H from Eqn 178, consider the action of H on an arbitrary ket $|\psi\rangle$ given by:

$$\begin{aligned} H|\psi\rangle &= \frac{1}{2m}P^2|\psi\rangle + \frac{1}{2}m\omega^2 X^2|\psi\rangle \\ &= \frac{1}{2m}P^\dagger P|\psi\rangle + \frac{1}{2}m\omega^2 X^\dagger X|\psi\rangle. \end{aligned} \quad (179)$$

Hitting the above equation on the left with $\langle \psi |$ we can write:

$$\begin{aligned}
\langle \psi | H | \psi \rangle &= \frac{1}{2m} \langle \psi | P^\dagger P | \psi \rangle + \frac{1}{2} m \omega^2 \langle \psi | X^\dagger X | \psi \rangle \\
&= \frac{1}{2m} \langle P \psi | P \psi \rangle + \frac{1}{2} m \omega^2 \langle X \psi | X \psi \rangle \\
&\geq 0.
\end{aligned} \tag{180}$$

The final step above follows from the fact that the norms of the states $|P\psi\rangle$ and $|X\psi\rangle$ have to be non-negative. In Eqn 180, if we now set $|\psi\rangle$ to be an eigenstate of H , it follows that the eigenvalues of H are non-negative.

Next, we recall the canonical commutation relation between the position operator X and the momentum operator P given by:

$$[X, P] = XP - PX = i\hbar. \tag{181}$$

The above relation may be viewed just as a statement of the Heisenberg uncertainty principle (see Eqn 127).

Now, **define** an operator a and its adjoint a^\dagger as follows:

$$a = \sqrt{\frac{m\omega}{2\hbar}} \left(X + \frac{iP}{m\omega} \right) \tag{182}$$

$$a^\dagger = \sqrt{\frac{m\omega}{2\hbar}} \left(X - \frac{iP}{m\omega} \right). \tag{183}$$

We note that, by definition, both a and its adjoint a^\dagger are non-Hermitian and hence cannot represent observables. However, the point is that they are related to H in a very nice way as we shall see below. Using Eqns 182 and 183 and keeping in mind that the correct order of multiplication is critical while dealing with operators, we can write:

$$\begin{aligned}
a^\dagger a &= \frac{m\omega}{2\hbar} \left(X - \frac{iP}{m\omega} \right) \left(X + \frac{iP}{m\omega} \right) \\
&= \frac{m\omega}{2\hbar} \left(X^2 + \frac{i}{m\omega} XP - \frac{i}{m\omega} PX + \frac{P^2}{m^2\omega^2} \right) \\
&= \frac{m\omega}{2\hbar} X^2 + \frac{1}{2m\omega\hbar} P^2 + \frac{i}{2\hbar} [X, P] \\
&= \frac{H}{\hbar\omega} - \frac{1}{2}.
\end{aligned} \tag{184}$$

A very similar calculation leads to the following result as well:

$$aa^\dagger = \frac{H}{\hbar\omega} + \frac{1}{2}. \tag{185}$$

In deriving the above, we have used the definition of H (Eqn 178) and the commutation relation (Eqn 181). From Eqn 184 we have the relation :

$$H = (a^\dagger a + 1/2) \hbar\omega. \quad (186)$$

Now, let us **define** the *number* operator N as:

$$N = a^\dagger a. \quad (187)$$

Indeed, N is Hermitian since:

$$\begin{aligned} N^\dagger &= (a^\dagger a)^\dagger \\ &= a^\dagger (a^\dagger)^\dagger \\ &= a^\dagger a \\ &= N. \end{aligned}$$

Now, from Eqn 186 we have:

$$H = (N + 1/2) \hbar\omega. \quad (188)$$

Since H is a linear function of N it follows that both can be diagonalized simultaneously. From our previous discussions we recall that this means that we can find kets that are eigenstates for both operators simultaneously. Let $|n\rangle$ be an eigenstate of N . Thus,

$$N|n\rangle = n|n\rangle.$$

Given the relation between H and N (Eqn 188) we can then write:

$$H|n\rangle = \left(n + \frac{1}{2}\right) \hbar\omega |n\rangle. \quad (189)$$

The next step reveals the point behind the whole exercise. Since we have established that $|n\rangle$ represents states that are simultaneously eigenstates for both H and N , it follows that $|n\rangle$ represents eigenstates of H . Now, comparing this statement with our operator form of the time independent Schrodinger equation in the energy basis, $H|E\rangle = E|E\rangle$, it is clear that the energy levels of the system (ie. eigenvalues of H) are now given by:

$$E_n = \left(n + \frac{1}{2}\right) \hbar\omega. \quad (190)$$

This is precisely the quantization condition on energy for the harmonic oscillator. The straightforward conclusion is that by introducing the operators a and a^\dagger we have solved the eigenvalue problem for H without having to solve a differential equation. Also, we see that the energy of the oscillator is *quantized*. That is, it can exist only in states the energy of each of which is given by Eqn 190.

However, at this point, the significance of n needs to be explored further. We reiterate that $|n\rangle$ represents eigenstates for both the *number* operator that we have defined earlier as well the hamiltonian. Thus, $|n\rangle$ represents the different allowed energy states of the oscillator. In what follows we address questions such as the possible allowed values for n and the transition of the oscillator between the allowed energy states induced by the operators that we have introduced.

Keeping in mind the definition of the commutator we first derive the following results.

$$\begin{aligned}
 [a^\dagger, a] &= (a^\dagger a - a a^\dagger) \\
 &= \frac{H}{\hbar\omega} - \frac{1}{2} - \left(\frac{H}{\hbar\omega} + \frac{1}{2} \right) \quad (\text{using Eqns 185 and 186}) \\
 &= -1.
 \end{aligned} \tag{191}$$

Recalling the definition of N (Eqn 187) consider:

$$\begin{aligned}
 [N, a] &= [a^\dagger a, a] \\
 &= (a^\dagger a a - a a^\dagger a) \\
 &= (a^\dagger a - a a^\dagger) a \\
 &= [a^\dagger, a] a \\
 &= -a. \quad (\text{using Eqn 191})
 \end{aligned} \tag{192}$$

We note that, in the penultimate step of the above derivation, we could legitimately pull the operator a outside the paranthesis to the **right** due to the fact that it was the last factor on the **right** in both the terms in the previous expression.

Likewise, we can also show:

$$[N, a^\dagger] = a^\dagger. \tag{193}$$

Using the above results Eqns 192 and 193 and the definition of the commutator which gives $[N, a] = Na - aN$, consider:

$$\begin{aligned}
Na|n\rangle &= ([N, a] + aN)|n\rangle \\
&= (-a + aN)|n\rangle \quad (\text{using 192}) \\
&= -a|n\rangle + na|n\rangle \quad (\text{since, } N|n\rangle = n|n\rangle, \text{ by defn}) \\
&= (n - 1)a|n\rangle .
\end{aligned} \tag{194}$$

Similarly we can show:

$$Na^\dagger|n\rangle = (n + 1)a^\dagger|n\rangle . \tag{195}$$

Note that since a is an operator, its action on $|n\rangle$ produces another ket. What Eqn 194 tells us is that this ket produced by the action of a on $|n\rangle$ is also an eigenstate of N since the state $a|n\rangle$ is regained (upto a constant factor) when N acts on it. A similar conclusion emerges from considering Eqn 195 which tells us that $a^\dagger|n\rangle$ is also an eigenstate of N .

The key point here is the following. From Eqn 194, we conclude that $a|n\rangle$ is an eigenstate of N **with eigenvalue** $(n - 1)$. Now, it is canonical that the eigenstate associated with an eigenvalue $(n - 1)$ is $|n - 1\rangle$. Hence $a|n\rangle$ must be some multiple of $|n - 1\rangle$ and we write:

$$a|n\rangle = c|n - 1\rangle . \quad (c \text{ is a constant}) \tag{196}$$

From normalizing $|n - 1\rangle$, we can show $c = \sqrt{n}$. Hence:

$$a|n\rangle = \sqrt{n}|n - 1\rangle . \tag{197}$$

A similar calculation yields:

$$a^\dagger|n\rangle = \sqrt{n + 1}|n + 1\rangle . \tag{198}$$

Let us now interpret our results. We recall that n represents the eigenvalues of the number operator N that was defined earlier. It is time now to recognize that n represents the number of quanta of energy in the oscillator and that N derives its name from this fact. At this point we invoke our first result in this subsection that the eigenvalues of the

hamiltonian operator are non-negative to conclude that the same holds for n . This follows from the linear relationship that we have established between H and N (Eqn188). Indeed, the oscillator cannot be in a state with a negative quantum of energy.

We now turn our attention to the operators a and a^\dagger whose action on an energy eigenstate (labelled by n) is given by Eqns 197 and 198. The operator a , acting on any given energy state produces the immediate lower energy state. Hence it is called the *annihilation* operator since it “takes away” a quantum of energy to push the oscillator to the $|n - 1 \rangle$ state. In contrast, the operator a^\dagger has the opposite effect, “pushing” the oscillator to the next higher energy state $|n + 1 \rangle$. Hence it is called the *creation* operator. These two operators provide the mathematical formalism for the oscillator to shift along the different (quantized) energy states.

Now, if we start from an energy eigenstate $|n \rangle$ and repeatedly apply the annihilation operator a to this state, we reach the lowest possible energy state of the oscillator or the *ground state*. This state corresponds to $n = 0$ and the corresponding eigenvalue for energy is given by $E_0 = \hbar\omega/2$ (see Eqn190). Alternatively, repeated application of the creation operator to the ground state yields the higher energy states.

Let us now step back to look at the bigger picture. The harmonic oscillator problem was solved in the operator approach without solving a differential equation in a particular basis and then applying appropriate boundary conditions. The quantization conditions on the energy of the oscillator were derived by an ingenious manipulation of operators, following Dirac. The other insight from this approach is the fundamental nature of the commutation relations between conjugate dynamical variables (in this case the X and P operators) in the process of quantization of a dynamical system. This is profound enough that the conditions to be satisfied by the matrix representation of the operators in the X basis (our second postulate **Q2**, please see Section 6) can well be replaced by the commutation relations (Eqn181)! These relations being basis independent (they are a statement about abstract operators), the power of replacing basis dependent conditions with the commutation relations is apparent. Furthermore, the route to quantization by prescribing commutation relations between the fundamental operators is very useful in the case of multi-dimensional systems.

10.2 The Power Series Approach

In this concluding sub-section, we sketch the approach to the harmonic oscillator problem by the brute force method of solving the Schrodinger equation explicitly in the X basis. In principle, this is nothing but the direct application of using power series methods to solve a second order ODE. In order to obtain normalized solutions, the power series solution must terminate and imposing this condition yields the quantization condition on the energy identical to Eqn 190. Under these circumstances, the eigenfunctions for the wave function ψ turn out to be the Hermite polynomials.

Given the harmonic oscillator hamiltonian (Eqn 178), the time-independent part of the Schrodinger equation (Eqn 177) in the energy eigenbasis can be written in operator form as:

$$\left(\frac{P^2}{2m} + \frac{1}{2}m\omega^2 X^2 \right) |E\rangle = E|E\rangle. \quad (199)$$

We note that the above equation is nothing but the general time-independent Schrodinger equation $H|E\rangle = E|E\rangle$ written out for the harmonic oscillator hamiltonian. As seen in previous problems, in order to project Eqn 199 onto the X basis (in other words, in order to obtain an ODE in x satisfied by the wave function ψ) we follow our *anstaz* of making the substitutions $X \rightarrow x$, $P \rightarrow -i\hbar d/dx$ and $|E\rangle \rightarrow \psi$ into Eqn 199 to obtain:

$$\left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2 \right) \psi = E\psi. \quad (200)$$

Re-arranging this equation we can write:

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2} \left(E - \frac{1}{2}m\omega^2 x^2 \right) \psi = 0. \quad (201)$$

An important point here is that we are interested in solutions to Eqn 201 that can be normalized to unity or the Dirac delta function. Per our earlier discussion (see Sections 2 and 6) only such kets constitute the *physical* Hilbert space and hence enter quantum mechanics.

Introducing the non-dimensional variable $x = by$, we can write Eqn 201 as:

$$\frac{d^2\psi}{dy^2} + \frac{2mEb^2}{\hbar^2} \psi - \frac{m\omega^2 b^4}{\hbar^2} y^2 \psi = 0. \quad (202)$$

Now, choosing $b = \sqrt{\frac{\hbar}{m\omega}}$ and defining another non-dimensional parameter $\epsilon = \frac{mEb^2}{\hbar^2} = \frac{E}{\hbar\omega}$ we can write Eqn 202 as:

$$\frac{d^2}{dy^2} + (2\epsilon - y^2) \psi = 0. \quad (203)$$

The next step is to examine the behavior of Eqn 203 in the limits $y \rightarrow \infty$ and $y \rightarrow 0$. In the limit $y \rightarrow \infty$, we may neglect the $2\epsilon\psi$ term to write the equation as:

$$\frac{d^2}{dy^2} - y^2\psi = 0. \quad (204)$$

The solution to Eqn 204, in the limit $y \rightarrow \infty$ can be obtained as:

$$\psi = Ay^m e^{\pm y^2/2}, \quad A- \text{an arbitrary constant.} \quad (205)$$

In this asymptotic limit, we choose the solution $\psi = Ay^m e^{-y^2/2}$ since the other possible solution grows exponentially.

Now, we examine the limit $y \rightarrow 0$. Since the $y^2\psi$ term can be dropped in this limit, Eqn 203 reduces to:

$$\frac{d^2}{dy^2} + 2\epsilon\psi = 0. \quad (206)$$

The solution can be written as:

$$\psi = A \cos(\sqrt{2\epsilon}y) + B \sin(\sqrt{2\epsilon}y). \quad (207)$$

Since we had dropped the $y^2\psi$ to obtain this solution, we ought to drop the y^2 and higher order terms in the solution as well. Thus, in the limit $y \rightarrow 0$, the solution should be of the form:

$$\psi = A + cy + O(y^2) \quad c- \text{a constant.} \quad (208)$$

From the consideration of the limits above, we deduce that the solution should be of the form:

$$\psi(y) = u(y)e^{-y^2/2}. \quad (209)$$

where u approaches the RHS of Eqn 208 in the limit $y \rightarrow 0$ and y^m in the limit $y \rightarrow \infty$. In order to determine $u(y)$, we now substitute from Eqn 209 into Eqn 203 to obtain:

$$\frac{d^2}{dy^2}u - 2y\frac{d}{dy}u + (2\epsilon - 1)u = 0. \quad (210)$$

We now assume power series solution to Eqn 210 of the form:

$$u(y) = \sum_{n=0}^{\infty} C_n y^n. \quad (211)$$

Following the standard procedure of substituting the assumed form of the power series into the equation to be solved yields the recurrence relation in this case as:

$$C_n + 2 = C_n \frac{(2n + 1 - 2\epsilon)}{(n + 2)(n + 1)}. \quad (212)$$

Now we are confronted with a major problem. It is that, when written out explicitly using co-efficients generated by the recurrence relation above, we obtain divergent solutions! However, if we impose the following constraint on ϵ given by:

$$\epsilon_n = \frac{2n + 1}{2}, \quad n = 0, 1, 2, \dots \quad (213)$$

the coefficient $C_n + 2$ and others dependent on it vanish. Now, with the choice $C_1 = 0$ when n is even and $C_0 = 0$ when n is odd, we obtain a finite polynomial of order n that satisfies the DE Eqn 203 as well as the required behavior in the limits $y \rightarrow \infty$ and $y \rightarrow 0$. We also note that, on imposition of the condition Eqn 213, the recurrence relation leads to polynomial solutions which turn out to be the *Hermite* polynomials. It is interesting to observe that these special functions, that were studied in their own right in mathematics long before the advent of quantum mechanics, emerged as the set of solutions to the harmonic oscillator problem.

The key point here is that Eqn 213 is a **quantization condition on energy**. Recalling our definition of the non-dimensional parameter $\epsilon = mEb^2/\hbar^2 = E/\hbar\omega$, we can re-write the condition on the allowed eigenvalues of the energy as:

$$E_n = \left(n + \frac{1}{2}\right) \hbar\omega. \quad (214)$$

Indeed, these are precisely the quantization conditions derived by the operator method in Eqn 190. Thus, by two distinct approaches, one reaches the remarkable conclusions that the energy levels of the harmonic oscillator are quantized and that the ground state of the quantum oscillator does not correspond to a state of zero energy.

This concludes our discussion of the harmonic oscillator.

References

- [1] Shankar R., *Principles of Quantum Mechanics*