

Fundamentals of Vibronic Spectroscopy

John A. Shelnutt

Advanced Materials Laboratory
Sandia National Laboratories
Albuquerque, NM 87185-1349

and

Department of Chemistry
University of Georgia
Athens, GA 30602-2556

January 19, 2009

Copyright © 2008 by John A. Shelnutt. All rights reserved.

PREFACE

This is a course devoted to the development of a theoretical formalism that can be used generally to describe the interaction between a radiation field (*e.g.*, light) and a collection of charged particles (*e.g.*, molecules). Energy measurements on such a system, *i.e.*, a molecule interacting with light, constitute the pursuit of most spectroscopic techniques.

In general, the radiation field is electromagnetic radiation consisting of γ -rays, X-rays, UV-visible-NIR light, infrared light, microwaves, radio waves, or other particle fields. However, we will limit ourselves to photon fields in this text.

The collection of particles can be as small as a single charged particle, or as large as a molecule or condensed matter such as a crystal or amorphous solid.

The spectroscopic measurement process is generally an effort to detect the number of photons (intensity) of light as a function of their energy $E = h\nu$ (or, equivalently, the frequency or wavelength) of the photons leaving the interaction region. This spectroscopic information (*i.e.*, the energy spectrum) is generally used to obtain information about the system of particles being probed by the radiation field. (*e.g.*, molecular information)

In this course in molecular spectroscopy, we will first develop a general formalism that is adequate for a description of most spectroscopic methods. Most conventional spectroscopic texts present a hodge-podge of different formalisms comprising a different approach for each type of spectra. Instead, here we develop a general formalism that can be applied in a general way. In addition, this theoretical approach is subject to few limitations on its applicability and can be used to actually calculate spectra in most cases. We will apply this approach in particular to vibrational spectra (infrared and Raman spectra) and to electronic spectra when vibrational structure is evident in the optical absorption and Raman excitation spectra. Other spectroscopic techniques will also be examined in less detail, but the methods for applying the theoretical treatment to these spectroscopic techniques are the same.

This book is intended for first- or second-year graduate students in chemistry or chemical physics. We first develop the quantum theory that is the basis of spectroscopy and proceed to the development of quantum electrodynamics just to the point needed for spectroscopic theory. The idea is to fully develop the quantum mechanical basis for the interaction of light with molecules, not to lead the reader into a full-blown study of quantum electrodynamics. With this solid theoretical foundation, we then develop a general vibro-electronic (vibronic) approach for treating and calculating vibronic spectra.

This vibronic approach described in the text provides the formalism required for simulating the spectra obtained in most spectroscopic techniques, and, in the weak vibronic interaction limit, it gives explicit formulae for spectral simulation and prediction. The vibronic effects on the UV-visible absorption and Raman spectra are examined in detail as examples, but it is easily seen how to apply this vibronic formalism to a variety of other spectroscopic methods and techniques.

In the later chapters, dissipative phenomena in spectra are introduced in a novel way using results from the study of multiplicative stochastic processes and the averaged time-evolution equations. Using this treatment of dissipation, for example, the line widths in spectral bands are introduced in a natural way.

CONTENTS

Preface		2
Contents		4
Chapter 1	Introduction and Background	6
	1.1 Quantum Mechanics	6
	1.2 Background	7
	1.3 Postulate I of Quantum Theory	8
	1.4 Vector Spaces	10
	1.5 Hilbert Space	20
	1.6 Function Space	24
	1.7 Postulate II of Quantum Theory	26
	1.8 Linear Operators	27
	1.9 Eigenvalues and Eigenvectors	29
	1.10 Postulate IV of Quantum Theory	32
	1.11 Postulate V of Quantum Theory	36
	1.12 Compatible observables	37
	1.13 Heisenberg Uncertainty Principle	40
	1.14 Alternative Postulates of Quantum Theory	47
	1.15 Postulate III of Quantum Theory	49
	1.16 Time Evolution Operator	53
	1.17 Constants of Motion	55
	1.18 Hamiltonian of a Molecule	57
Chapter 2	Interaction Hamiltonian for a Radiation Field with a Molecule	60
	2.1 Classical Mechanics	60
	2.2 Reformulations of Classical Equations of Motion	61
	2.3 Energy Formulation of the Equation of Motion	62
	2.4 Lagrange's Equations of Motion	65
	2.5 Hamilton's Equations of Motion and the Hamiltonian	74
	2.6 Electrodynamics and Maxwell's Equations	77
	2.7 Classical Hamiltonian of a Charged Particle Interacting with an Electromagnetic Field	78
	2.8 The Quantum Mechanical Hamiltonian Operator for a Charged Particle in a Field	85
	2.9 The Hamiltonian of a Molecule in an Electromagnetic Field	86
	2.10 Gauge Invariance and the Electromagnetic Wave Equations	87
	2.11 The Hamiltonian of the Free Electromagnetic Field	95
	2.12 Quantum Harmonic Oscillators and the Electromagnetic Field	100
	2.13 The Hamiltonian of a Molecule in an Electromagnetic Field	121
Chapter 3	Perturbative Treatment of the Interaction Hamiltonian	124
	3.1 Perturbation Theory and Time Evolution of the System	124
	3.2 Transition Probabilities	126
	3.3 Evaluation of the Matrix Elements of the Interaction	131
	3.4 The Long Wavelength Approximation	136

	3.5	The Dipole Approximation	136
	3.6	Absorption Spectroscopy	141
	3.7	Fermi's Golden Rule	141
	3.8	Two-Photon Absorption Spectroscopy	141
	3.9	Spontaneous Raman Spectroscopy	141
	3.10	Resonance Fluorescence	145
	3.11	The Raman Scattering Tensor	148
Chapter 4		The Molecular Energy Eigenstates	150
	4.1	The Born-Oppenheimer Approximation	150
	4.2	Adiabatic Born-Oppenheimer Approximation	150
	4.3	The Crude Born-Oppenheimer Approximation	154
	4.4	The Electronic States	156
	4.5	Vibrational States (Harmonic Oscillator)	156
	4.6	Vibronic Coupling and the Approximate Vibronic States	160
Chapter 5		Vibronic Spectroscopy	166
	5.1	UV-Visible Absorption Spectra with Vibrational Structure	166
	5.2	Intensity Borrowing in Absorption Spectra	173
	5.3	Infrared Absorption Spectra	177
	5.4	Vibrational Raman Scattering Tensor	179
	5.5	Rayleigh Scattering	181
	5.6	Stokes Raman Scattering	181
	5.7	Interference Effects and Intensity Borrowing in Resonance Raman Spectra	185
	5.8	Anti-Stokes Raman Scattering	187
	5.9	Overtone and Combination Line Raman Intensities	192
	5.10	Non-resonance Raman Scattering	197
	5.11	Multi-mode Phenomena	197
Chapter 6		Solutions of Radiation Field Problem	198
	6.1	Plane Waves	
	6.2	Wave Packets	
Chapter 7		Dissipative Phenomena and Line Widths	7
	7.1	Reduced (Time-Averaged) Schrödinger Equation	
	7.2	Reduced Density Matrix	
Chapter 8		Transformations and Averaging in the Laboratory Frame	8

CHAPTER 1

Introduction and Background

1.1 Quantum Mechanics

In spectroscopy, our goal is generally to describe a system, like that illustrated in Figure 1, composed of a radiation field, typically photons, and a collection of charged particles interacting with the electromagnetic field. The aim of most spectroscopic techniques is to measure the energy spectrum of such a system with the ultimate goal of learning about the collection of particles, such as their molecular structure. The radiation field is generally electromagnetic in nature, *i.e.*, UV, visible, NIR, and infrared light, X-rays, γ -rays, microwave, and radio waves. We will limit ourselves to the photons of an electromagnetic field in this text. The particles could be the electrons and nuclei of a molecule, or more complicated condensed phases. In general, this problem should be treated entirely quantum mechanically, and that is our primary task.

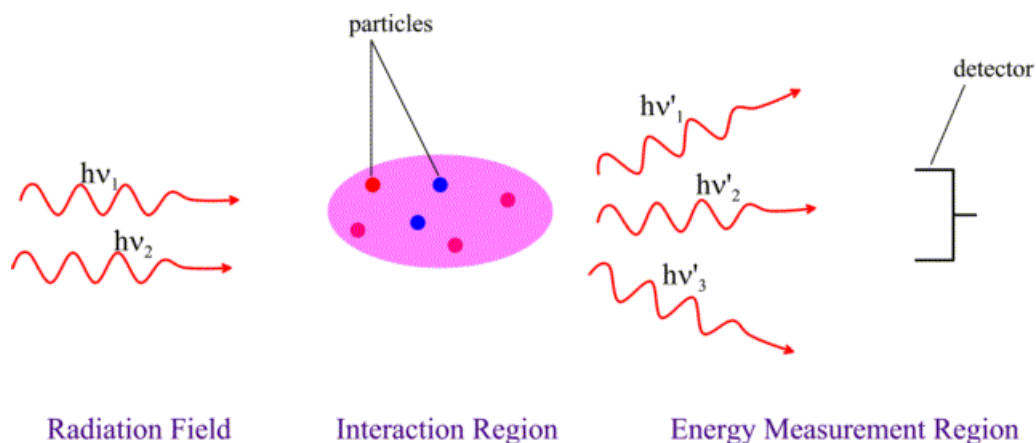


Figure 1. Spectroscopy is generally an energy spectrum measurement for a system composed of a radiation field and a collection of charged particles.

To begin, we review the postulates of quantum mechanics. These five postulates are the basic assumptions about how nature is organized and how we can use mathematics to derive predictions about future events given information about current conditions. More, specifically, these postulates describe how mathematics can be used to *determine the outcome of measurements of physical observables and to make future predictions on the basis of the present values of a set of physical observables.*

The choice of these basic postulates is somewhat arbitrary. Those chosen are the ones most easily applied to solve many chemical and physical problems. However, other sets of basic postulates of quantum mechanics could be chosen. As an example, we will also list a more basic set of postulates, which are equivalent to the conventional ones, that we will be applying to

develop the spectroscopic methods in this course--our primary goal in this text. This more basic postulate set will illustrate how fundamental quantum mechanics is to understanding the way our world works and how intimately interconnected the mathematical methods of quantum theory are to the measurement process.

1.2 Background

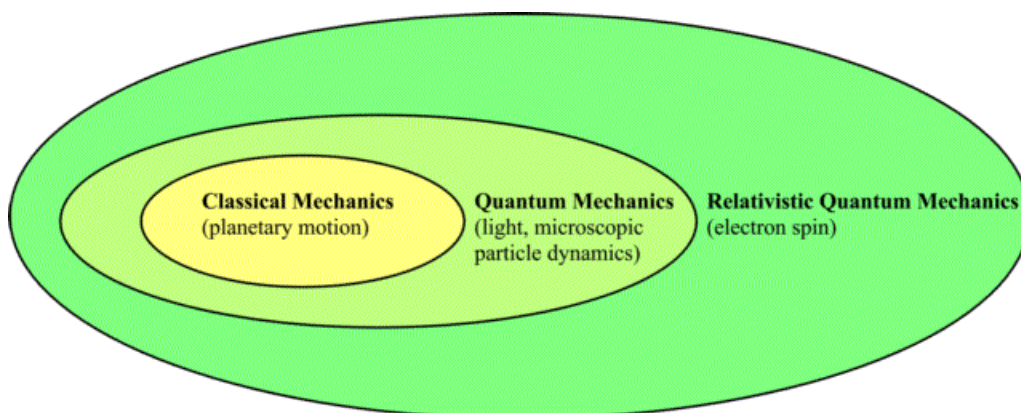


Figure 2. All observable phenomena.

With a sufficient *quantum mechanical description* of a system of particles (electrons and nuclei for chemical problems), one can calculate any property of the system or predict results of any experiment designed to measure an observable property of the system.

The system can be as large as the universe or as small as a single particle. Of course, the larger the system, the more complicated the mathematical description. Therefore, we will usually consider a simple particle or a small group of particles (*e.g.*, a molecule), possibly including interactions of this subset of the universe with the rest of the universe by adding external forces. The latter is an example of one kind of *reduction of the problem* to a manageable level of description. Another kind of reduction comes when we wish to measure, not the microscopic behavior or a complete description of the system of particles, but some macroscopic property of the collection of particles. The latter kind of reduction of the description involves the use of *quantum statistical mechanics*, which is not to be covered in the present course. Sometimes useful approximate methods are appropriate, and these methods represent another type of reduction. For example, for low velocities a non-relativistic quantum mechanical treatment is sufficient. In other cases, a relativistic quantum theory is required. All of the postulates except postulate III, which gives the time evolution of the system, are the same for relativistic and non-relativistic quantum mechanics. In other cases, a reduction to the classical mechanics limit of the quantum theory is appropriate. Using these powerful reductive computational methods, in principle, one can calculate “any quantity of chemical interest.” In practice, quantum chemical calculations are often very complex, requiring large and fast computers.

We will first introduce the concepts of quantum theory, gradually developing the

mathematical skills necessary for the use these concepts in spectroscopy. Later, we will use quantum mechanical concepts to solve a number of simple general problems of interest to us as chemists, from simple 1-particle, 1-dimensional problems to more complicated 3-dimensional problems with many particles (molecules).

One obstacle to learning to think quantum mechanically (which we must do in order to become competent chemists) is its bizarre, non-intuitive nature. To develop intuition for quantum mechanical systems, one must first understand the origin of the postulates of quantum mechanics. Secondly, after understanding their necessity, one must learn from the mathematics of quantum mechanics a new way of thinking about nature. Unlike classical mechanics, we have no “natural” or “common sense” understanding of quantum mechanical phenomena. The purpose of this course is to understand the necessity of a quantum mechanical description of nature, and spectroscopic techniques in particular, and to learn how to think in quantum mechanical terms about such problems. In order to accomplish this, we must work out the physical consequences of the quantum theory in enough instances of spectroscopic interest to gain at least a rudimentary intuition for happenings at the microscopic level. This acquired intuition will serve to guide our use of experimental spectroscopic techniques to obtain structural and chemical information.

1.3 Postulate I of Quantum Theory

We will develop the necessary background in classical mechanics, as it is needed for understanding each of the basic postulates of quantum mechanics.

Postulate I: For a system of particles with n generalized positional coordinates q_1, \dots, q_n , the state at time t is specified completely by a normalized state vector (wavefunction or eigenfunction):

$$|\psi_t\rangle \equiv \psi_t(q_1, q_2, \dots, q_n)$$

in Hilbert space (extended to continuous eigenvalues).

The state vector is in general an ∞ -dimensional complex vector of unit length, *e.g.*, an infinite series of complex numbers. Hilbert space is just an infinite dimensional vector space for which an inner product is defined (inner product space). This is necessary so that the norm of the vector is defined. The q_i in the state vector are independent of time. (In contrast, the coordinates are a function of time in classical mechanics.) Here, the state vector is time dependent, and, in making it time dependent, we are implicitly assuming the Schrödinger picture of quantum mechanics.

The physical implication of Postulate I is that *all possible information about the system of particles is contained in the state vector*. Viewed another way, the state vector tells what can be known about the system of particles. The postulate does not answer several important questions, such as: (1) how do we extract information about the system (*e.g.*, position, momentum) from

the state vector, (2) how is state vector itself to be interpreted, and, (3) how are physical measurements related to state vector? These questions are answered by Postulates II, III, IV, and V and a Corollary to Postulate I.

The physical interpretation of the state vector is given by the corollary to Postulate I (question 2 above).

Corollary: The probability **P** of finding the particle of the system associated with the state vector $|\psi_t\rangle$ at points between q_1, q_2, \dots, q_n and $q_1 + dq_1, \dots, q_n + dq_n$ is given by:

$$P = P(q_1, q_2, \dots, q_n) dq_1 \dots dq_n = \psi_t^* \psi_t dq_1 \dots dq_n = |\psi_t(q_1, q_2, \dots, q_n)|^2 dq_1 \dots dq_n,$$

where $P(q_1 \dots q_n) = |\psi_t(q_1, q_2, \dots, q_n)|^2$ is known as the probability density.

Note then that the quantity defined by the *inner product* of the state vector with itself $\langle \psi_t | \psi_t \rangle$:

$$\langle \psi_t | \psi_t \rangle \equiv \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \psi_t^*(q_1, q_2, \dots, q_n) \psi_t(q_1, q_2, \dots, q_n) dq_1, \dots, dq_n = 1$$

Mathematically, this follows from the fact that the state vector is postulated to be *normalized*, but physically the integral represents the probability of finding the particles of the system in the universe, *i.e.*, all of space, and this must be unity.

For example, suppose the *state vector* is a Gaussian function as shown in Figure 3. Then the area of the indicated thin section is $|\psi_t(q')|^2 dq$, and the integral is unity. The Gaussian wavefunction has the form

$$\psi(q) = \frac{1}{(2\pi\sigma^2)^{1/4}} e^{-\frac{(q-\mu)^2}{4\sigma^2}},$$

so the wavefunction squared has the form of the normal distribution,

$$|\psi(q)|^2 = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-\frac{(q-\mu)^2}{2\sigma^2}},$$

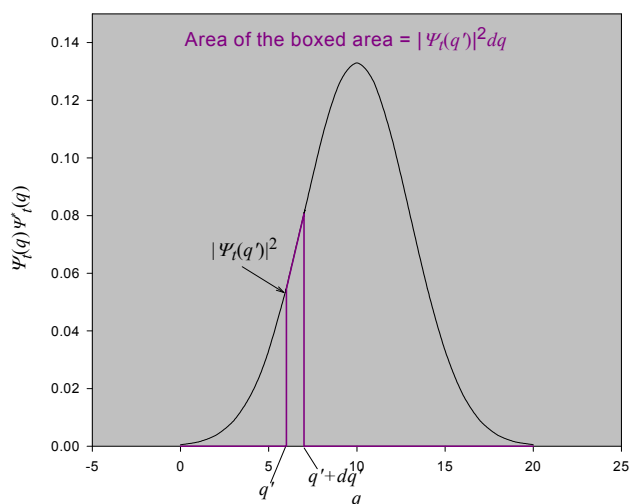


Figure 3. A Gaussian statefunction.

and the integral from $-\infty$ to $+\infty$ is 1 (using $\int_0^{\infty} e^{-a^2x^2} dx = \frac{(\pi)^{1/2}}{2a}$).

There are grave differences between the classical and quantum concepts of the state of the system. Classically, the state is completely defined by measuring exactly and simultaneously the values of all coordinates and their conjugate momenta for all the particles of the system at $t = 0$. Further, if the forces acting on the particles are also known then the state (*i.e.*, the coordinates and momenta) is known exactly for all times.

In quantum mechanics, the state of the system is specified by the state vector, which moreover has only a probabilistic interpretation. That is, only probabilities for measuring various values of the coordinates and momenta can be given. Furthermore, the coordinates q_j themselves have a different interpretation in quantum mechanics. The coordinates no longer represent the positions of the particles, but instead represent positions in coordinate space at which a probability field is defined. The coordinates in quantum mechanics are independent of time, whereas in classical mechanics the coordinates of the particles may vary in time. However, the probability of finding a particle in a certain location does vary in time because the state vector varies with time. Another postulate (Postulate III) tells how the state vector evolves with time and gives quantum mechanics its predictive capability for future measurements.

1.4 Vector Spaces

Before we further develop the quantum theory, let us consider the mathematical nature of the state vector, that is, a vector in a vector space called Hilbert space. Hilbert space is a special kind of vector space, which we define below.

Definition. A vector space, V , is defined as a collection of objects, called vectors, which satisfy the following rules of combination,

1. *Closure under addition:* if \vec{u} and \vec{v} are two vectors in the space, then $\vec{u} + \vec{v}$ is a *unique* vector also in the space (*i.e.*, $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ is unique).
2. *Additive associativity:* $(\vec{v} + \vec{u}) + \vec{w} = \vec{v} + (\vec{u} + \vec{w})$.
3. *Additive commutivity:* $\vec{v} + \vec{u} = \vec{u} + \vec{v}$.
4. *Additive identity:* $\vec{v} + \vec{0} = \vec{v}$, $\forall \vec{v}$.
5. *Additive inverses:* $\forall \vec{v}$ an inverse exist, $-\vec{v}$, such that $\vec{v} + (-\vec{v}) = \vec{0}$.
6. *Closure under scalar multiplication:* \forall scalars c and vectors \vec{v} , there is a unique vector $c\vec{v}$ also in V .
7. *Scalar multiplication is:*
 - a. Distributive: $c(\vec{v} + \vec{u}) = c\vec{v} + c\vec{u}$
 - b. Distributive: $(c + c')\vec{v} = c\vec{v} + c'\vec{v}$
 - c. Associative: $(cc')\vec{v} = c(c'\vec{v})$

An example. Consider the following example of a vector space. Consider the collections of n real members,

$$\vec{v} = (v_1, \dots, v_n)$$

where v_1, \dots, v_n are called the components of the vector, and n is the dimension of the vector. These entities make up a vector space. Vector addition is defined by addition of the components, *i.e.*,

$$\vec{v} + \vec{u} = (v_1 + u_1, v_2 + u_2, \dots, v_n + u_n)$$

All seven properties of vectors can be shown to hold for such quantities. For example, with scalar multiplication defined by $c\vec{v} = (cv_1, cv_2, \dots, cv_n)$, then property 7a can be shown to hold. Explicitly,

$$\begin{aligned} c(\vec{v} + \vec{u}) &= c(v_1 + u_1, \dots, v_n + u_n) \\ c(\vec{v} + \vec{u}) &= (cv_1 + cu_1, \dots, cv_n + cu_n) \\ c(\vec{v} + \vec{u}) &= (cv_1, \dots, cv_n) + (cu_1, \dots, cu_n) \\ c(\vec{v} + \vec{u}) &= c\vec{v} + c\vec{u}. \end{aligned}$$

Linear independence. One of the most important concepts in quantum mechanics stemming from the vector nature of the state is *linear independence*, which means that no vector of a group of linearly independent vectors can be represented as a linear combination of the others. Because of closure under addition, a linear combination of vectors defined by

$$\vec{u} = \sum_{i=1}^n c_i \vec{v}_i$$

where the c_i are a set of scalar coefficients is also a vector. The set of n vectors, \vec{v}_i , are said to be linearly independent, if and only if,

$$\sum_{i=1}^n c_i \vec{v}_i = \vec{0} \Rightarrow c_1 = c_2 = \dots = c_n = 0.$$

Proof: On the contrary, suppose that one of the c_i is non-zero, say $c_j \neq 0$, then

$$\sum_{i=1}^n c_i \vec{v}_i = \vec{0}$$

and

$$\sum_{i \neq j}^n c_i \vec{v}_i = -c_j \vec{v}_j.$$

Thus, $\vec{v}_j = -\frac{1}{c_j} \sum_{i \neq j}^n c_i \vec{v}_i$, and \vec{v}_j is a linear combination of the other vectors, and therefore the group of vectors is not linearly independent.

Basis. Another useful vector space concept is that of a *basis* for a vector space. A set of linearly independent vectors that span the entire space is a basis for the space. The set spans the space, if every vector can be written as a linear combination of the set of vectors. If the spanning set of vectors $\{\vec{u}_1 \dots \vec{u}_n\}$ form a basis for the space, (i.e., they are linearly independent), then any vector \vec{v} can be represented as a linear combination of the \vec{u}_i ,

$$\vec{v} = \sum_{i=1}^n c_i \vec{u}_i.$$

Moreover, the c_i are unique.

Proof: Suppose there are two linear representations of \vec{v} , then

$$\vec{v} = \sum_{i=1}^n c_i \vec{u}_i \quad \text{and} \quad \vec{v} = \sum_{i=1}^n c'_i \vec{u}_i.$$

Subtracting,

$$\vec{v} - \vec{v} = \vec{0} = \sum_{i=1}^n (c_i - c'_i) \vec{u}_i$$

However, linear independence implies that $c_i - c'_i = 0$ for $\forall i$, and thus, $c_i = c'_i$ for all i , which proves that the c_i are unique.

Any two basis sets for the same space must have the same number of vectors. This number is defined as the dimension of the vector space. Thus, the vector space in our example is an n -dimensional vector space. A set of vectors larger than n can span the space, but only n of these vectors can be linearly independent.

Isomorphism. Isomorphism of vector spaces is another useful concept, which, in quantum mechanics, means that we can choose the most convenient of a group of isomorphic vector spaces, because all properties will be the same regardless of the space chosen. This will be shown later to insure the equivalence of wave mechanics (function space) and matrix mechanics (Hilbert space).

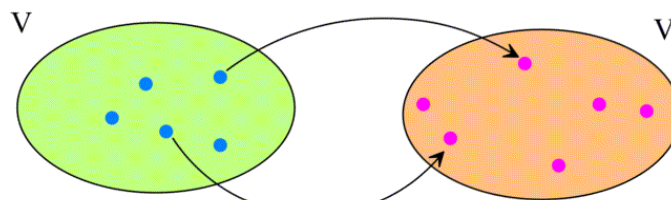


Figure 4. One-to-one vector mapping from vector space V to V' .

Two vector spaces, V and V' , are said to be *isomorphic*, if (a) for every vector in V , there

is one and only one corresponding vector in V' , and (b) the laws of combination are the same for addition and scalar multiplication.

Inner product spaces. In some vector spaces, an additional operation for the combination of vectors, called the *inner product*, is defined. Hilbert space is such an *inner product space*. In quantum mechanics, the state vector inner product, at least with itself at this point in our discussion, must be defined because the norm of the state vectors $\langle \psi_t | \psi_t \rangle$ must equal 1. For our n -dimensional vector-space example, with vectors $\vec{v} = (v_1 \dots v_n)$, the inner product is defined by

$$(\vec{v}, \vec{u}) = \sum_{i=1}^n v_i u_i (= |\vec{u}| |\vec{v}| \cos \theta),$$

and this n -dimensional, inner product space is said to be Euclidean if the norm is real and non-negative. The inner product of a vector with itself is $(\vec{v}, \vec{v}) = \sum_{i=1}^n v_i^2$,

the length of the vector squared. The norm or length of the vector is then $(\vec{v}, \vec{v})^{1/2} = |\vec{v}| \geq 0$.

Complex inner product vector spaces. If the components of the vectors are complex numbers, then we define the inner product as:

$$(\vec{u}, \vec{v}) = \sum_{i=1}^n u_i^* v_i,$$

or in the state-vector notation (converting the sum to an integral),

$$\langle \psi_t | \chi_t \rangle = \int \dots \int \psi_t^*(q \dots) \chi_t(q \dots) dq_1 dq_2 \dots dq_n,$$

and thus the norm of the vector $|\psi_t\rangle$ is

$$\langle \psi_t | \psi_t \rangle = \int \dots \int \psi_t^*(q \dots) \psi_t(q \dots) dq_1 dq_2 \dots dq_n,$$

so the lengths of the vectors defined in this way will non-negative and real (Euclidean). (This will be shown below for state vectors in Hilbert space.) This follows from the fact that a complex number times its complex conjugate is non-negative real,

$$(a - bi)(a + bi) = a^2 + b^2 \geq 0.$$

Thus,

$$(\vec{u}, \vec{u})^{1/2} = \left(\sum_{i=1}^n u_i^* u_i \right)^{1/2} = \left\{ \sum_{j=1}^n (u_j^r + i u_j^i)^* (u_j^r + i u_j^i) \right\}^{1/2}$$

$$= \left\{ \sum_{j=1}^n (u_j^r - iu_j^i)(u_j^r + iu_j^i) \right\}^{1/2}.$$

And,

$$(\bar{u}, \bar{u})^{1/2} = \left\{ \sum_{j=1}^n (u_j^{r2} + u_j^{i2}) \right\}^{1/2} = \left\{ \sum_{j=1}^n |u_j|^2 \right\}^{1/2} \geq 0.$$

However, note that $(\bar{v}, \bar{u}) \neq (\bar{u}, \bar{v})$, *i.e.*, the vector product does not commute. Instead,

$$\begin{aligned} (\bar{u}, \bar{v}) &= \sum_{i=1}^n u_i^* v_i = \sum_{i=1}^n (u_i v_i^*)^* = \left(\sum_{i=1}^n u_i v_i^* \right)^* = \left(\sum_{i=1}^n v_i^* u_i \right)^* = (\bar{v}, \bar{u})^* \\ (\bar{u}, \bar{v}) &= (\bar{v}, \bar{u})^* \end{aligned}$$

$$\begin{aligned} \text{Note: } (a_j + ib_j)^* &= u_j^* & (a'_j + ib'_j) &= v_j \\ (a_j + ib_j)^* (a'_j + ib'_j) &= a_j a'_j + b_j b'_j + i(a_j b'_j - a'_j b_j) \\ \text{and } (a'_j + ib'_j)^* &= v_j^* & (a_j + ib_j) &= u_j \\ (a'_j + ib'_j)^* (a_j + ib_j) &= a_j a'_j + b_j b'_j + i(a'_j b_j - a_j b'_j) \\ \therefore u_j^* v_j &\neq v_j^* u_j & \Rightarrow & (\bar{u}, \bar{v}) \neq (\bar{v}, \bar{u}) \end{aligned}$$

Scalar multiplication with complex numbers must also be defined for the inner product.

$$\begin{aligned} (c\bar{u}, \bar{v}) &= \sum_{j=1}^n c^* u_j^* v_j = c^* \sum_{j=1}^n u_j^* v_j = c^* (\bar{u}, \bar{v}) \\ (c\bar{u}, \bar{v}) &= c^* (\bar{u}, \bar{v}). \end{aligned}$$

And,

$$(\bar{u}, c\bar{v}) = \sum_{j=1}^n u_j^* c v_j = c \sum_{j=1}^n u_j^* v_j = c (\bar{u}, \bar{v}).$$

The formal definition of an inner product space requires that the inner product obey relationships (1)-(3) below.

1. $(\bar{u}, \bar{v}) = (\bar{v}, \bar{u})^*$ (shown above),
2. $(a\bar{v} + b\bar{u}, \bar{w}) = a^* (\bar{v}, \bar{w}) + b^* (\bar{u}, \bar{w})$, and
3. $(\bar{v}, \bar{v}) \geq 0$.

Summarizing our results so far for vector spaces, we have defined the rules obeyed by

vectors, which are the members or entities composing a vector space. We also defined the operations that are required in any vector space, namely, vector addition and scalar multiplication. We showed some of the properties that special sets of vectors may possess, *i.e.*, a set of vectors (1) may be linearly independent (one vector of the set may not be represented as linear combination of the others), (2) may span the space (all vectors in the space may be represented as a linear combination the vectors of the set), and (3) may form a basis (linearly independent set of vectors that span the entire space). We also have defined the dimension of the space as the number of vectors in a basis for the space and the requirements for two spaces to be isomorphic. We have further defined some special kinds of vector spaces such as the inner product space, which has an additional operation for combination of vectors, called an inner or vector product. This is a way that two vectors can be “multiplied”. Specifically, in our n -dimension vector space composed of sets of n real numbers, $\vec{v} = (v_1, v_2, \dots, v_n)$, the inner product has been defined, and this space is said to be Euclidean because the norm (or length) of vector is a non-negative, real number. The norm is defined as $(\vec{v}, \vec{v})^{1/2}$, that is, the square root of the sum of squares of the elements. Three-dimensional coordinate space is such a space. The length of the position vector in coordinate spaces is $\vec{r} = r = (\vec{r}, \vec{r}) = (x^2 + y^2 + z^2)^{1/2} = |\vec{r}|$. Also, note that the norm of \vec{v} is indeed non-negative, real. Finally, we have introduced the idea of (n -dimensional) complex vector spaces, composed of sequences of complex numbers $\vec{v} = (a_1 + b_1i, a_2 + b_2i, \dots, a_n + b_ni)$. The reason for focusing on these types of vector spaces is that Hilbert space is such a complex vector space (∞ -dimensional). Hilbert space is also an inner product space, and we have defined the inner product in an n -dimensional complex vector space. Consequently, we can make the association, $\langle u | v \rangle \leftrightarrow (\vec{u}, \vec{v})$, that is, the complex inner product is the same as the vector product of state vectors of Hilbert space. (The extension to infinite-dimensional spaces will be dealt with shortly.)

Other useful relationships for the inner product. The Schwartz inequality states that “the product of the lengths of two vectors is greater than the magnitude of the inner product.”

Proof: Let $\vec{w} = a\vec{u} - b\vec{v}$, where a and b are arbitrary,

then $|\vec{w}|^2 \geq 0$, thus,

$$|a\vec{u} - b\vec{v}|^2 \geq 0.$$

$$\begin{aligned} \text{Also } (a\vec{u} - b\vec{v}, a\vec{u} - b\vec{v}) &= (a\vec{u}, a\vec{u}) + (b\vec{v}, b\vec{v}) - (b\vec{v}, a\vec{u}) - (a\vec{u}, b\vec{v}) \\ &= a^* a (\vec{u}, \vec{u}) + b^* b (\vec{v}, \vec{v}) - b^* a (\vec{v}, \vec{u}) - a^* b (\vec{u}, \vec{v}) \\ &= |a|^2 |\vec{u}|^2 + |b|^2 |\vec{v}|^2 - a^* b (\vec{u}, \vec{v}) - b^* a (\vec{v}, \vec{u}). \end{aligned}$$

Now choose: $a = |\vec{v}| (\vec{v}, \vec{u})^{1/2}$ and $b = |\vec{u}| (\vec{v}, \vec{u})^{1/2}$.

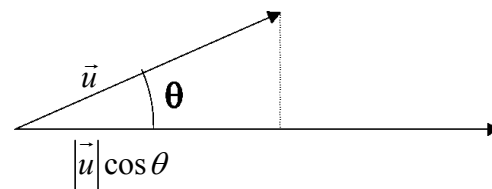
Remembering that $(\vec{u}, \vec{v})^{1/2} (\vec{v}, \vec{u})^{1/2} = [(\vec{u}, \vec{v})(\vec{u}, \vec{v})^*]^{1/2} = (|\vec{u}, \vec{v}|^2)^{1/2} = |\vec{u}, \vec{v}|$, and substituting we have

$$\begin{aligned} |\vec{u}|^2 |\vec{v}|^2 |(\vec{u}, \vec{v})| + |\vec{u}|^2 |\vec{v}|^2 |(\vec{v}, \vec{u})| - |\vec{v}| |\vec{u}| |(\vec{u}, \vec{v})(\vec{v}, \vec{u})| - |\vec{v}| |\vec{u}| |(\vec{v}, \vec{u})(\vec{u}, \vec{v})| &\geq 0 \\ |\vec{u}|^2 |\vec{v}|^2 |(\vec{u}, \vec{v})| + |\vec{u}|^2 |\vec{v}|^2 |(\vec{v}, \vec{u})| - |\vec{u}| |\vec{v}| |(\vec{u}, \vec{v})|^2 - |\vec{u}| |\vec{v}| |(\vec{v}, \vec{u})|^2 &\geq 0 \end{aligned}$$

$$2|\vec{u}||\vec{v}|(\vec{u}, \vec{v})[|\vec{u}||\vec{v}| - |(\vec{u}, \vec{v})|] \geq 0$$

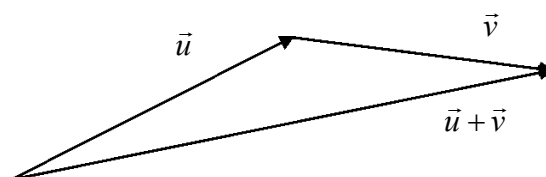
$$|\vec{u}||\vec{v}| \geq |(\vec{u}, \vec{v})| \quad (\text{Schwartz inequality})$$

$$\vec{u} \cdot \vec{v} = (\vec{u}, \vec{v}) = |\vec{u}||\vec{v}| \cos \theta \leq |\vec{u}||\vec{v}|$$



Additional inner product relationships are:

1. $|c\vec{v}| = |c||\vec{v}|$.
2. $|\vec{v}| > 0$, except for $\vec{v} = \vec{0}$.
3. $|\vec{u} + \vec{v}| \leq |\vec{u}| + |\vec{v}|$. (Triangle Inequality)



Orthonormality and complete sets. Two vectors, \vec{u} and $\vec{v} \neq \vec{0}$, are said to be orthogonal, if and only if,

$$(\vec{u}, \vec{v}) = 0.$$

A set of vectors are orthonormal, if and only if,

$$(\vec{v}_i, \vec{v}_j) = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases} \quad \forall \text{ vectors } \vec{v}_i \text{ and } \vec{v}_j \in \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}.$$

An example is the set of unit vectors in the x , y , and z directions in coordinate space, $\{\hat{i}, \hat{j}, \hat{k}\}$.

Any vector with a finite length $|\vec{v}|$ can be normalized to unity according to,

$$\hat{v} = \frac{\vec{v}}{|\vec{v}|}.$$

Proof: $(\hat{v}, \hat{v}) = \left(\frac{\vec{v}}{|\vec{v}|}, \frac{\vec{v}}{|\vec{v}|} \right) = \frac{1}{|\vec{v}|^2} (\vec{v}, \vec{v}) = \frac{|\vec{v}|^2}{|\vec{v}|^2} = 1$

$$\therefore |\hat{v}| = [(\hat{v}, \hat{v})]^{1/2} = 1$$

Orthonormal Basis. An orthonormal set of n vectors $\{\hat{v}_1, \hat{v}_2, \dots, \hat{v}_n\}$ that span an n -dimensional space is called an orthonormal basis. The set of vectors is necessarily linearly independent, since the set spans the space any vector \vec{u} can be expressed as a linear combination of the \vec{v}_j as

$$\vec{u} = \sum_{j=1}^n c_j \hat{v}_j,$$

and,

$$(\hat{v}_i, \vec{u}) = \sum_{j=1}^n (\hat{v}_i, c_j \hat{v}_j) = \sum_{j=1}^n c_j (\hat{v}_i, \hat{v}_j) = \sum_{j=1}^n c_j \delta_{ij}$$

$$(\hat{v}_i, \vec{u}) = c_i$$

Thus, the set is linearly independent since

$$\sum_{j=1}^n c_j \hat{v}_j = 0 \quad \Rightarrow \quad c_1 = \dots = c_n = 0.$$

Theorem: A finite orthonormal set of vectors $\{\hat{v}_1, \hat{v}_2, \dots, \hat{v}_k\}$ is linearly independent.

A set of k finite orthonormal vectors form a basis (span and are linearly independent) for a k -dimensional space. In addition, for an n -dimensional space, any set of n orthonormal vectors forms a basis.

Completeness (in a finite dimensional space). A set of vectors is a complete set if not contained in any larger orthonormal set in that space, *i.e.*, if it is a basis for the space. Any vector can be represented as a linear combination of a complete set.

Bessel's inequality is useful because it shows us how to generate a complete set of vectors that span a space. Let $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ be any finite orthonormal set in a complex inner product space. If \vec{u} is any vector in the space, and if

$$c_i = (\hat{v}_i, \vec{u}),$$

then

$$\sum_{i=1}^n |c_i|^2 \leq |\vec{u}|^2 \quad (\text{Bessel's inequality}).$$

Also,

$$\vec{u}' = \vec{u} - \sum_{i=1}^n c_i \hat{v}_i$$

is orthogonal to each \vec{v}_i and to the entire space spanned by the set. Remember everything we

learn about vectors is now also true for the states of a physical system, so this gives us a way to construct orthogonal states of the system and a set of states that form an orthonormal basis. In addition, notice that in the statement of Bessel's inequality we have not specified that the vector space is finite.

$$\begin{aligned}
\text{Proof: } 0 \leq |\vec{u}'|^2 &= (\vec{u}', \vec{v}') = (\vec{u} - \sum_{i=1}^n c_i \hat{v}_i, \vec{u} - \sum_{i=1}^n c_i \hat{v}_i) \\
&= (\vec{u}, \vec{u}) - \sum_{i=1}^n c_i^* (\hat{v}_i, \vec{u}) - \sum_{i=1}^n c_i (\vec{u}, \hat{v}_i) + \sum_{j=1}^n \sum_{i=1}^n c_i^* c_j (\hat{v}_i, \hat{v}_j) \\
&= |\vec{u}|^2 - \sum_{i=1}^n c_i^* c_i - \sum_{i \neq j} c_i^* c_j + \sum_{j=1}^n \sum_{i=1}^n c_i^* c_j \delta_{ij} \\
&= |\vec{u}|^2 - \sum_{i=1}^n |c_i|^2 \\
0 \leq |\vec{u}'|^2 &= |\vec{u}|^2 - \sum_{i=1}^n |c_i|^2,
\end{aligned}$$

giving

$$\sum_{i=1}^n |c_i|^2 \leq |\vec{u}|^2 \quad (\text{Bessel's Inequality}).$$

To prove the second part, *i.e.*, \vec{u}' is orthogonal to space spanned by $\{\hat{v}_j\}$, we must show that $(\vec{u}', \hat{v}_j) = 0$, for all j .

$$\begin{aligned}
\text{Proof: } (\vec{u}', \hat{v}_j) &= (\vec{u} - \sum_i c_i \hat{v}_i, \hat{v}_j) \\
(\vec{u}', \hat{v}_j) &= (\vec{u}, \hat{v}_j) - \sum_i c_i^* (\hat{v}_i, \hat{v}_j) \\
(\vec{u}', \hat{v}_j) &= (\vec{u}, \hat{v}_j) - \sum_i c_i^* \delta_{ij} \\
(\vec{u}', \hat{v}_j) &= c_j^* - c_j^* = 0 \quad \forall j.
\end{aligned}$$

Thus, \vec{u}' is orthogonal to the space spanned by $\{\hat{v}_1, \hat{v}_2, \dots, \hat{v}_n\}$.

If the vector space is only n -dimensional, then (\vec{u}, \hat{v}_j) , the component of \vec{u}' along each orthogonal "direction" in the space, is 0,

$$\vec{u}' = \sum_{j=1}^n (\vec{u}, \hat{v}_j) \hat{v}_j = \sum_{j=1}^n 0 \cdot \hat{v}_j = \vec{0},$$

i.e., \vec{u}' is the zero vector.

Note also that for an n -dimensional space,

$$\begin{aligned}\vec{u} = \vec{0} &\Rightarrow \vec{u} = \sum_{i=1}^n c_i \hat{v}_i \\ |\vec{u}|^2 &= \sum_{i,j=1}^n c_j^* c_i (\hat{v}_j, \hat{v}_i) = \sum_{i,j} c_j^* c_i \delta_{ij} \\ |\vec{u}|^2 &= \sum_{j=1}^n c_j^* c_j = \sum_{j=1}^n |c_j|^2\end{aligned}$$

or,

$$\sum_{j=1}^n |c_j|^2 = |\vec{u}|^2.$$

That is, for a complete orthonormal set, Bessel's Inequality becomes an equality.

What if the space is larger than n -dimensional? To examine this case, let the space be k -dimensional, where $k > n$, also let $\{\hat{v}_1, \dots, \hat{v}_n, \dots, \hat{v}_k\}$ be an orthogonal basis for the space, then

$$\begin{aligned}\hat{u}' &= \sum_{j=1}^k (\vec{u}', \hat{v}_j) \hat{v}_j = \sum_{j=1}^n (\vec{u}', \hat{v}_j) \hat{v}_j + \sum_{j=n+1}^k (\vec{u}', \hat{v}_j) \hat{v}_j \\ \hat{u}' &= \sum_{j=1}^k (\vec{u}', \hat{v}_j) \hat{v}_j \text{ is not necessarily } = \vec{0}.\end{aligned}$$

Therefore, this vector \vec{u}' could be used to extend the set of vectors by one, and, using similar vectors, we could ultimately extend the orthonormal spanning the n -dimensional subspace to a complete set for the k -dimensional space. This is the basis of the *Gram-Schmidt orthogonalization procedure*.

The Gram-Schmidt orthogonalization procedure always allows us to construct an orthonormal set of vectors from a finite set of linearly independent vectors. They are:

$$\begin{aligned}\vec{u}_1 &= \vec{v}_1 \\ \vec{u}_2 &= \vec{v}_2 - \frac{(\vec{u}_1, \vec{v}_2) \vec{u}_1}{|\vec{u}_1|^2} \\ \vec{u}_3 &= \vec{v}_3 - \frac{(\vec{u}_1, \vec{v}_3) \vec{u}_1}{|\vec{u}_1|^2} - \frac{(\vec{u}_2, \vec{v}_3) \vec{u}_2}{|\vec{u}_2|^2} \\ \vec{u}_{m+1} &= \vec{v}_{m+1} - \sum_{i=1}^m \frac{(\vec{u}_i, \vec{v}_{m+1}) \vec{u}_i}{|\vec{u}_i|^2}\end{aligned}$$

This orthogonal set can be normalized easily.

Parseval's Equation is another important relationship in quantum mechanics.

$$(\bar{w}, \bar{u}) = \sum_{i=1}^n (\bar{w}, \hat{v}_i)(\hat{v}_i, \bar{u}) \quad (\text{Parseval's equation})$$

is true, if $\{\hat{v}_j\}$ is a finite orthonormal basis for an n -dimensional space, and \bar{u} and \bar{w} are in the space. In state vector notation, Parseval's equation becomes,

$$\langle \psi_1 | \psi_2 \rangle = \sum_i \langle \psi_1 | \psi_i \rangle \langle \psi_i | \psi_2 \rangle.$$

This expression suggests that for a complete set we can formally write

$$\sum_i |\psi_i\rangle \langle \psi_i| = 1.$$

This expression is actually an operator sometimes called the completeness relation.

1.5 Hilbert Space

Postulate I states that the state of a system is completely described by a vector in Hilbert space, $|\Psi_i(q_1 \dots q_n)\rangle$, which gives the probability amplitude for finding the system in a element of volume of configuration space, $dq_1 \dots dq_n$. We are now in a position to define Hilbert space. Hilbert space is a vector space of composed of vectors, whose elements are an infinite set of complex numbers, $|u\rangle = (u_1 + iu'_1, u_2 + iu'_2, \dots)$. There is also an inner product defined and that is required to be finite, so that $\langle u | v \rangle < \infty$. To see what this definition means we must answer the following questions:

1. What is an ∞ -dimensional vector?
2. How is the inner product defined for ∞ -dimensional complex vectors?
3. What conditions on the inner product are required to form the inner product space?
4. Can we insure a finite norm for the vectors (required for a probabilistic interpretation of the state vector).

We start by extending our concepts of a vector space to infinite dimensional vectors. Instead of sequences of n complex numbers, we consider infinite sequences such as,

$$\vec{v} = (\dots v_i \dots).$$

We want the ∞ -dimensional space to be an inner-product space, so that a norm will be defined in the usual way as,

$$|\vec{v}| = (\vec{v}, \vec{v})^{1/2},$$

and we want the norm to be a finite, non-negative, real number. This is important because our state vector must be normalized (usually to unity). If the vector \vec{v} has a finite norm, we can always normalize, it to unity, by

$$\hat{v} = \frac{\vec{v}}{|\vec{v}|}.$$

This is possible if we define the inner product by analogy with the n -dimensional space as:

$$(\vec{v}, \vec{u}) = \sum_{i=1}^{\infty} v_i^* u_i,$$

and therefore the norm is given by,

$$|\vec{v}| = (\vec{v}, \vec{v}) = \left[\sum_{i=1}^{\infty} v_i^* v_i \right]^{1/2} = \left[\sum_{i=1}^{\infty} |v_i|^2 \right] \geq 0.$$

However, not all sequences of complex numbers have a finite norm. Nonetheless, we can show that all ∞ sequences that do possess a finite norm, as defined above, form a vector space. To show this we must demonstrate that all of the properties of vector spaces hold for such vectors. Most importantly, we must show closure under addition. That is we must prove that if \vec{u} and \vec{v} have finite norms, *i.e.*, are vectors in the space, we must show that

$$\vec{u} + \vec{v} = (u_1 + v_1, \dots, u_i + v_i, \dots)$$

also has a finite norm, or

$$\sum_{i=1}^{\infty} |u_i + v_i|^2 < \infty.$$

We start by showing that for two complex numbers,

$$|u_i + v_i|^2 \leq 2 \left[|u_i|^2 + |v_i|^2 \right].$$

Proof: We start with the triangle inequality applied to the complex components of the vectors \vec{u} and \vec{v} . The triangle inequality can be applied since complex numbers can be represented as two-dimensional vectors. Thus,

$$\begin{aligned} & |u_i + v_i| \leq |u_i| + |v_i| \\ \therefore & |u_i + v_i|^2 \leq |u_i|^2 + |v_i|^2 + 2|u_i||v_i|, \end{aligned}$$

but we can also write,

$$\begin{aligned} & (|u_i| - |v_i|)^2 \geq 0 \\ \therefore & |u_i|^2 + |v_i|^2 - 2|u_i||v_i| \geq 0 \\ \text{or} & |u_i|^2 + |v_i|^2 \geq 2|u_i||v_i| \\ \therefore & |u_i + v_i|^2 \leq |u_i|^2 + |v_i|^2 + 2|u_i||v_i| \leq 2[|u_i|^2 + |v_i|^2]. \end{aligned}$$

Using this result,

$$\begin{aligned} \sum_{i=1}^{\infty} |u_i + v_i|^2 & \leq 2 \left[\sum_{i=1}^{\infty} |u_i|^2 + \sum_{i=1}^{\infty} |v_i|^2 \right] = 2[(\vec{u}, \vec{u}) + (\vec{v}, \vec{v})] < \infty \\ (\vec{u} + \vec{v}, \vec{u} + \vec{v}) & = |\vec{u} + \vec{v}|^2 < \infty \\ |\vec{u} + \vec{v}| & < \infty, \end{aligned}$$

that is, the norm is finite, and therefore $\vec{u} + \vec{v}$ is a vector within the space and closure is proven.

Closure under scalar multiplication is also easily proven:

$$|c\vec{v}| = (c\vec{v}, c\vec{v})^{1/2} = \left[c^* c \sum_{i=1}^{\infty} |v_i|^2 \right]^{1/2} = \left[|c|^2 |\vec{v}|^2 \right]^{1/2} < \infty$$

Therefore,

$$|c||\vec{v}| < \infty.$$

The inner product of two ∞ -dimensional vectors with finite norms is also finite, *i.e.*, the inner product exists.

$$(\vec{u}, \vec{v}) = \sum_i u_i^* v_i,$$

since

$$|u_i|^2 + |v_i|^2 \geq 2|u_i||v_i| \geq 2|u_i^* v_i|,$$

where the last inequality follows from:

$$\begin{aligned}
u_i &= a + bi \\
v_i &= a' + b'i \\
|u_i^* v_i| &= |(a - bi)(a' + b'i)| = |(aa' + bb') + i(ab' - a'b)| \\
|u_i^* v_i| &= \{(aa' + bb')^2 + i(ab' - a'b)^2\}^{1/2} \\
|u_i||v_i| &= (a^2 + b^2)^{1/2}(a'^2 + b'^2)^{1/2} = [(a^2 + b^2)(a'^2 + b'^2)]^{1/2} \\
|u_i||v_i| &= [a^2 a'^2 + b^2 b'^2 + a^2 b'^2 + a'^2 b^2]^{1/2} \\
|u_i^* v_i| &= [a^2 a'^2 + b^2 b'^2 + 2aa'bb' + a'^2 b^2 + a^2 b'^2 - 2aa'bb']^{1/2} \\
|u_i||v_i| &= |u_i^* v_i|.
\end{aligned}$$

Thus, we have,

$$(\vec{u}, \vec{v}) = \sum_i u_i^* v_i \leq \sum_i |u_i^* v_i| \leq \frac{1}{2} \sum_i (|u_i|^2 + |v_i|^2) < \infty,$$

showing that the inner product is finite.

The three properties required for an inner product space can also be shown to hold.

The concept of an orthonormal basis (complete set of orthonormal vectors) is maintained in Hilbert space; and there are an infinite number of orthonormal vectors in the basis, $\{\hat{v}_1, \dots, \hat{v}_i, \dots\}$. The projection of a vector \vec{u} along the i^{th} unit vector is (\hat{v}_i, \vec{u}) .

We can also define the projection of \vec{u} onto an orthonormal set $\{\dots, \hat{v}_i, \dots\}$ as,

$$\sum_{i=1}^{\infty} (\hat{v}_i, \vec{u}) \hat{v}_i.$$

If this orthonormal set is complete, *i.e.*, an orthonormal basis, then from our proof of Bessel's inequality for the case of completeness,

$$\vec{u} = \sum_{i=1}^{\infty} c_i \hat{v}_i = \sum_{i=1}^{\infty} (\hat{v}_i, \vec{u}) \hat{v}_i,$$

That is, the projection of \vec{u} onto an orthonormal basis is the vector itself.

One can also generalize Parseval's Equation to infinite-dimensional vector spaces. Let $\{\hat{v}_1, \dots, \hat{v}_i, \dots\}$ be an orthonormal basis, then

$$(\bar{u}, \bar{w}) = \sum_{i=1}^{\infty} (\bar{u}, \hat{v}_i)(\hat{v}_i, \bar{w}).$$

Using the different notation for state vectors,

$$\bar{u} \equiv |u\rangle \text{ and } (u \equiv \langle u|.$$

Parseval's equation becomes,

$$\langle u|w\rangle = \sum_{i=1}^{\infty} \langle u|v_i\rangle \langle v_i|w\rangle = \langle u|\left\{ \sum_{i=1}^{\infty} |v_i\rangle \langle v_i| \right\}|w\rangle,$$

and we identify the quantity in $\{ \}$ as unity, giving the completeness relation for infinite dimensional spaces

$$\sum_{i=1}^{\infty} |v_i\rangle \langle v_i| = 1.$$

1.6 Function Space

Function vector spaces. We don't normally think of functions as forming a vector space, but in certain instances, this is the case. In quantum mechanics, the state vector can often be represented as a function, called a wavefunction or statefunction.

Consider the set of complex-valued functions of a real variable x . Further, limit the functions to those that are integrable and square-integrable, *i.e.*,

$$\int f(x) dx < \infty,$$

and

$$\int |f(x)|^2 dx < \infty,$$

or, in other words, these integrals exist. We can see that the latter inequality is the continuous analog of the discrete condition, $\sum_{i=1}^{\infty} |v_i|^2 < \infty$, that the norm exists for Hilbert space vectors.

Functions that are integrable and square-integrable form an inner-product vector space. For example, we can show closure under addition, *i.e.*,

$$h(x) = f(x) + g(x)$$

must be integrable and square-integrable for $h(x)$ to be a vector in this space. First,

$$\int h(x)dx = \int \{f(x) + g(x)\}dx = \int f(x)dx + \int g(x)dx < \infty,$$

so $h(x)$ is integrable. For square-integrability,

$$\begin{aligned} \text{Proof: } \int |h(x)|^2 dx &= \int |f(x) + g(x)|^2 dx = \int \{f(x) + g(x)\}^* \{f(x) + g(x)\} dx \\ &= \int \{|f|^2 + |g|^2 + f^*g + g^*f\} dx \\ &= \int |f|^2 + |g|^2 + 2\text{Re}(f^*g) dx \leq \int |f|^2 + |g|^2 + 2|f^*g| dx \\ &\leq \int \{|f|^2 + |g|^2 + 2|f||g|\} dx \\ \int |h(x)|^2 dx &\leq \int 2\{|f|^2 + |g|^2\} dx = 2\int |f|^2 dx + 2\int |g|^2 dx < \infty, \end{aligned}$$

since $2|f||g| \leq |f|^2 + |g|^2$, and the latter follows from $(|f| - |g|)^2 = |f|^2 + |g|^2 - 2|f||g| \geq 0$. Therefore, the sum of two square-integrable functions is square integrable. Similarly, we can show the other six properties for vector addition and scalar multiplication hold for such functions.

Inner Product Function Spaces. We can also define the inner product in function space as:

$$(f, g) = \int_{-\infty}^{\infty} f^*(x)g(x)dx,$$

as a generalization of

$$\sum_x f_x^* g_x,$$

where x is a continuous index. Further, we can show that (f, g) exists and satisfies the requirement for inner product space, namely,

1. $(f, g) = \int f^* g dx = \int (g^* f)^* dx = (g, f)^*$
2. $(cf + dg, h) = c^*(f, h) + d^*(g, h)$
3. $(f, f) = \int f^* f dx = \int |f|^2 dx \geq 0$ (non-negative, real and finite).

Finally, we can generalize our one-variable function space made up of vectors $\{f_i(x)\}$ to n generalized coordinate variables, q_1, \dots, q_n , so that the vectors are now functions of the q_i 's, i.e.,

the set of functions,

$$\{F_i(q_1, \dots, q_n)\}.$$

The inner product of such vectors is defined by,

$$\{F(q_1, \dots, q_n), G(q_1, \dots, q_n)\} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} G(q_1, \dots, q_n) F(q_1, \dots, q_n) dq_1 \cdots dq_n.$$

This inner product space is isomorphous to the Hilbert space of state vectors. In Postulate I, we identified the state of the system as a state vector $|\psi_t(q_1, \dots, q_n)\rangle$. $|\psi_t(q_1, \dots, q_n)\rangle$ is now seen to be a function in a function space in n generalized coordinates, *i.e.*,

$$|\psi_t(q_1, \dots, q_n)\rangle \leftrightarrow F(q_1, \dots, q_n, t)$$

The inner product is now written as,

$$\langle \psi_t | \chi_t \rangle \leftrightarrow (\psi_t(q_1, \dots, q_n), \chi_t(q_1, \dots, q_n)) = \int \cdots \int \psi_t^*(q_1 \dots q_n) \chi_t(q_1 \dots q_n) dq_1 \dots dq_n.$$

Notice that the inner product of the state vector $|\psi_t\rangle$ with itself is the norm (squared),

$$\langle \psi_t | \psi_t \rangle = \int \cdots \int \psi_t^*(q_1 \dots q_n) \psi_t(q_1 \dots q_n) dq_1 \dots dq_n$$

which is finite, and can be made to equal unity. Thus, we can use either a vector in function space or a vector in Hilbert space (infinite sequences of complex numbers) to represent the state of the system. Thus, in function space, we are working in the wave mechanics of Schrödinger. In Hilbert space, we are working in the matrix mechanics of Heisenberg. Both representations of quantum mechanics are equivalent because the vector spaces are equivalent. Now that we understand what Hilbert space is, we are ready for the second postulate of quantum mechanics.

1.7 Postulate II of Quantum Theory

Postulate II defines how the state vectors are related to physically observable properties of the system.

Postulate II. For each *physical observable* A , there is a Hermitian operator \hat{A} with a complete orthonormal set of eigenfunctions (eigenvectors), $|\alpha_i\rangle$, and, further, the only possible values that can be obtained for a measurement of A are the eigenvalues of \hat{A} ,

$$a_1, a_2, \dots, a_n, \dots$$

The physical observables are just the usual dynamical variables of classical mechanics, things like momentum, energy, position, *etc.* Mathematically, we must introduce the concept of the Hermitian operator, a linear operator that has real eigenvalues. The eigenvalues must of course be real because they are the possible values of the dynamical variables. We must also acquaint ourselves with the eigenvalue problem for both continuous and discrete eigenvalues. Moreover, then there is the question of how we obtain \hat{A} the operator corresponding to an observable A .

1.8 Linear operators

An operator is a set of instructions that change one vector into another. An operator \hat{A} is said to “operate on” a vector $|\alpha\rangle$, when written as a product,

$$\hat{A}|\alpha\rangle.$$

The result of \hat{A} operating on $|\alpha\rangle$ is another vector $|\beta\rangle$, *i.e.*,

$$|\beta\rangle = \hat{A}|\alpha\rangle.$$

$|\beta\rangle$ is, in general, not in the same space as $|\alpha\rangle$.

However, we will restrict ourselves to linear operators, which have the property that they map a vector in the space V into another vector in V (the same space). A linear operator is defined by the two conditions:

- 1 $\hat{L}(|\alpha\rangle + |\beta\rangle) = \hat{L}|\alpha\rangle + \hat{L}|\beta\rangle$ Distributive under vector addition.
- 2 $\hat{L}(c|\alpha\rangle) = c(\hat{L}|\alpha\rangle)$ Commutative under scalar multiplication.

The derivative operator provides an example of a linear operator in function space. Let $\hat{L} = \frac{d}{dx}$ and $|\alpha\rangle = f(x)$ and $|\beta\rangle = g(x)$ are vectors in the space. Then

$$\hat{L}(|\alpha\rangle + |\beta\rangle) = \frac{d}{dx}(f(x) + g(x)) = \frac{d}{dx}f(x) + \frac{d}{dx}g(x),$$

$$\hat{L}(|\alpha\rangle + |\beta\rangle) = \hat{L}|\alpha\rangle + \hat{L}|\beta\rangle.$$

Thus, $\frac{d}{dx}$ is a linear operator in function space. It is not a Hermitian operator, however.

A Hermitian operator is defined as a self-adjoint operator, *i.e.*, an operator that satisfies the relationship,

$$\hat{A} = \hat{A}^\dagger$$

where the symbol \dagger represents the adjoint of the operator. The adjoint of an operator is defined in terms of the inner product relationship,

$$(\hat{L}\bar{u}, \bar{v}) = (\bar{u}, \hat{L}^\dagger\bar{v}).$$

Thus, the adjoint of \hat{L} , \hat{L}^\dagger , is the operator which acts on the vector on the right to keep the inner product the same as when \hat{L} acts on the vector on the left. Thus, for a *Hermitian* operator \hat{L} , the same operator acts on the vector on the left and on the right and keeps the inner product the same.

An example of a Hermitian operator is $\hat{p}_x \equiv -i\hbar \frac{d}{dx}$. For two vectors in function space, $\bar{u} = f(x)$ and $\bar{v} = g(x)$, which of course must be integrable and square integrable, we have

$$\begin{aligned} (\hat{p}_x f, g) &= \int_{-\infty}^{\infty} \left\{ -i\hbar \frac{d}{dx} f(x) \right\}^* g(x) dx \\ (\hat{p}_x f, g) &= +i\hbar \int_{-\infty}^{\infty} \left\{ \frac{d}{dx} f(x) \right\}^* g(x) dx \\ \{ \text{since } \frac{d}{dx} f(x)^* g(x) &= \frac{d}{dx} f^*(x) g(x) + f^*(x) \frac{d}{dx} g(x) \} \\ (\hat{p}_x f, g) &= i\hbar \left\{ \int_{-\infty}^{\infty} \frac{d}{dx} [f^*(x) g(x)] dx - \int_{-\infty}^{\infty} f^*(x) \frac{d}{dx} g(x) dx \right\} \\ &= i\hbar \left\{ f^*(x) g(x) \Big|_{-\infty}^{+\infty} - \int_{-\infty}^{\infty} f^*(x) \frac{d}{dx} g(x) dx \right\} \\ &= i\hbar f^*(x) g(x) \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} f^*(x) (-i\hbar) \frac{d}{dx} g(x) dx \\ (\hat{p}_x f, g) &= i\hbar f^*(x) g(x) \Big|_{-\infty}^{\infty} + (f, \hat{p}_x g) \end{aligned}$$

Now since $\int_{-\infty}^{\infty} |f(x)|^2 dx$ is finite, $f(x) \rightarrow 0$ as $x \rightarrow \pm\infty$, and the same is true for $f(x)$, $g(x) \rightarrow 0$ as $x \rightarrow \pm\infty$. Therefore, we have, $(\hat{p}_x f, g) = (f, \hat{p}_x g)$. However, by the definition of the adjoint,

$$(\hat{p}_x f, g) = (f, \hat{p}_x^\dagger g),$$

thus, in this case we have $\hat{p}_x = \hat{p}_x^\dagger$, *i.e.*, \hat{p}_x is Hermitian.

We will find out later that \hat{p}_x is the operator corresponding to the x -component of the momentum of a particle. It is necessary that \hat{p}_x , or, in general, the operator for any other observable be Hermitian, because only then will the eigenvalues, *i.e.*, the outcomes of any measurement, be real numbers. The classical x momentum,

$$mv_x \rightarrow -i\hbar \frac{\partial}{\partial x}.$$

To get the operator corresponding to a classical mechanical variable that is a function of the position and linear momentum, $f(x, p_x)$, we just replace x in f with the operator \hat{x} = “multiplication by x ” and replace $p_x = m\dot{x}$ by the operator $\hat{p}_x = -i\hbar \frac{\partial}{\partial x}$ in the function f , $f(\hat{x}, \hat{p}_x)$.

1.9 Eigenvalues and eigenvectors

The second postulate of quantum theory states that the only values that can be obtained for a measurement of the observable represented by the operator \hat{A} are the eigenvalues of \hat{A} , $a_1, a_2, \dots, a_n, \dots$. The eigenvalues are the numbers, a_i that satisfy the equations

$$\hat{A}_i |\alpha_i\rangle = a_i |\alpha_i\rangle.$$

This equation must be solved to find the eigenvalues, a_i , which in general may be complex numbers, and the eigenvectors $|\alpha_i\rangle$. The interpretation of the eigenvector is the subject of later postulates. As an example, we can find the eigenvalues and eigenvectors of \hat{T}_x defined by,

$$\hat{T}_x = \frac{\hat{p}_x^2}{2m}$$

where $\hat{p}_x = -i\hbar \frac{d}{dx}$. Substituting gives

$$\hat{T}_x = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2},$$

the quantum mechanical *kinetic energy* operator. In other words, we want to find solutions $\{E_i\}$ and $\{|\varepsilon_i\rangle\}$, that satisfy,

$$\hat{T}_x |\varepsilon_i\rangle = E_i |\varepsilon_i\rangle.$$

To do this, consider functions $|\varepsilon_k\rangle$ of the form

$$|\varepsilon_k\rangle = A \sin(kx) + B \cos(kx).$$

Applying the operator,

$$\begin{aligned} \hat{T}_x |\varepsilon_k\rangle &= -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} (A \sin kx + B \cos kx), \\ &= -\frac{\hbar^2}{2m} \{A(-k^2) \sin kx + B(-k^2) \cos kx\} \\ &= +\frac{k^2 \hbar^2}{2m} \{A \sin kx + B \cos kx\}. \end{aligned}$$

Thus, we have

$$\hat{T}_x |\varepsilon_k\rangle = \left(\frac{k^2 \hbar^2}{2m} \right) |\varepsilon_k\rangle,$$

and, except for normalization concerns, we have the eigenvalues of the kinetic energy operator. That is, according to Postulate II, the possible values for the kinetic energy of a particle of mass m when not acted upon by any forces are

$$\frac{k^2 \hbar^2}{2m}.$$

The eigenvalues are real because \hat{T}_x is Hermitian.

We now demonstrate that the eigenvalues of a Hermitian operator are real. Let \hat{A} be a Hermitian operator with eigenvalues a_1, \dots, a_n, \dots , and eigenvectors $\{|\alpha_i\rangle\}$. Then, $\hat{A}|\alpha_i\rangle = a_i|\alpha_i\rangle$ is satisfied, and

$$\langle \alpha_i | \hat{A} | \alpha_i \rangle = a_i \langle \alpha_i | \alpha_i \rangle = a_i$$

if the $|\alpha_i\rangle$ are normalized. Since \hat{A} is Hermitian, $\hat{A} = \hat{A}^\dagger$, substituting, we also have

$$\begin{aligned}\langle \alpha_i | \hat{A} | \alpha_i \rangle &= \langle \alpha_i | \hat{A}^\dagger | \alpha_i \rangle \\ &= \langle \hat{A} \alpha_i | \alpha_i \rangle,\end{aligned}$$

since $(A^\dagger)^\dagger = A$ is always true. Continuing, we have

$$\begin{aligned}\langle \alpha_i | \hat{A} | \alpha_i \rangle &= \langle a_i \alpha_i | \alpha_i \rangle \\ &= a_i^* \langle \alpha_i | \alpha_i \rangle = a_i^*.\end{aligned}$$

Thus, $a_i = a_i^*$, implying that a_i is real.

Now, if we have solved the eigenvalue problem for our operator \hat{A} and obtained a complete set of orthonormal eigenvectors $\{|\tilde{\alpha}_i\rangle\}$, then the state of the system can also be represented as a linear combination of the eigenvectors of \hat{A} as

$$|\psi_t\rangle = \sum_{i=1}^{\infty} c_i |\alpha_i\rangle,$$

and the coefficients c_k can be found by taking the inner product of the state vector $|\psi_t\rangle$ with each of the $|\alpha_k\rangle$,

$$\begin{aligned}\langle \alpha_k | \psi_t \rangle &= \left\langle \alpha_k \left| \sum_{i=1}^{\infty} c_i |\alpha_i\rangle \right. \right\rangle = \sum_{i=1}^{\infty} c_i \langle \alpha_k | \alpha_i \rangle \\ \langle \alpha_k | \psi_t \rangle &= \sum_{i=1}^{\infty} c_i \delta_{ik} = c_k\end{aligned}$$

Thus, $c_k = \langle \alpha_k | \psi_t \rangle$, and substitution gives

$$|\psi_t\rangle = \sum_{i=1}^{\infty} |\alpha_i\rangle \langle \alpha_i | \psi_t \rangle,$$

which is just as easily obtained by using the completeness relationship,

$$\sum_{i=1}^{\infty} |\alpha_i\rangle \langle \alpha_i| = 1,$$

by placing operating with the identity operator on the state vector.

1.10 Postulate IV of Quantum Theory

We will come back to Postulate III later. Postulate IV now tell us how to use the eigenvalues and eigenvectors of an observable to obtain the probability of finding the system to have the value a_i when a measurement of the observable corresponding to the operator \hat{A} is made.

Postulate IV. If the observable A , with eigenvalues of \hat{A} given by $\{a_i\}$ and eigenvectors given by $\{|\alpha_i\rangle\}$, is measured for the system in the state $|\psi_t\rangle$, then the probability of obtaining a_i at time t is given by

$$P_t(a_i) = \sum_k |\langle \alpha_k | \psi_t \rangle|^2$$

$$\{ \{P_t(a_i)\} = \int_k |\langle \alpha(k) | \psi_t \rangle|^2 dk \text{ for continuous eigenvalues.} \}$$

The summation over k includes all eigenvalues $a_k = a_i$ (or $a(k) = a$).

Let us first consider the case, for which all eigenvalues a_i are non-degenerate, *i.e.*, $a_i \neq a_k, \forall k, i$. In this case, the probability reduces to

$$P_t(a_i) = |\langle \alpha_i | \psi_t \rangle|^2 = |c_i|^2.$$

The probability of measuring the system and finding any one of the a_i , that is

$$\sum_i P_t(a_i) = \sum_i |\langle \alpha_i | \psi_t \rangle|^2$$

should be unity. Because the $|\alpha_i\rangle$ form a complete orthonormal set, we can expand $|\psi_t\rangle$ as

$$|\psi_t\rangle = \sum_{k=1}^{\infty} c_k(t) |\alpha_k\rangle,$$

and therefore,

$$\sum_{i=1}^{\infty} P_t(a_i) = \sum_{i=1}^{\infty} \sum_{k=1}^{\infty} |c_k(t) \langle \alpha_i | \alpha_k \rangle|^2 = \sum_{i=1}^{\infty} \sum_{k=1}^{\infty} |c_k(t) \delta_{ik}|^2$$

$$= \sum_{i=1}^{\infty} |c_i(t)|^2.$$

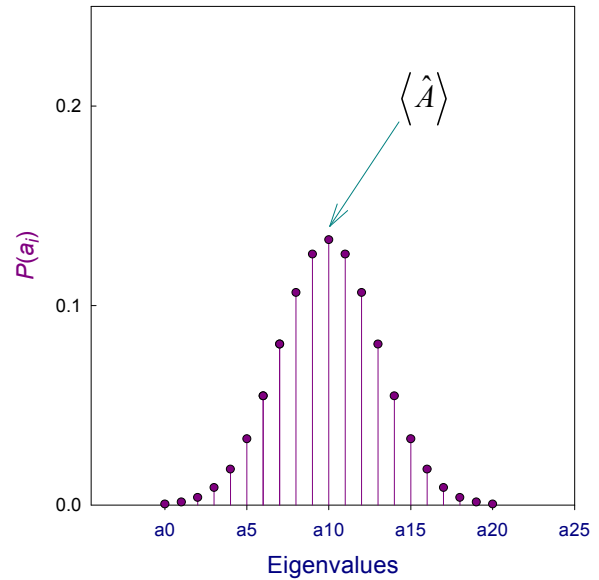
Finally, because the wavefunction $|\psi_t\rangle$ is normalized, *i.e.*,

$$\langle \psi_t | \psi_t \rangle = \sum_{i=1}^{\infty} \sum_{k=1}^{\infty} c_i^*(t) c_k(t) \langle \alpha_i | \alpha_k \rangle = \sum_{i=1}^{\infty} |c_i(t)|^2 = 1,$$

thus we have

$$\sum_{i=1}^{\infty} P_i(a_i) = 1.$$

Figure 5. A probability distribution for the discrete eigenvalues of the observable A . For this symmetric distribution, the expectation value and the most likely value, a_{10} in this case, are the same. This need not be the case for other distributions.



This result shows that our definition of $P_i(a_i)$ is reasonable for a probability, since the probability of finding the system in a possible state should be one. Furthermore,

$$0 \leq |\langle \alpha_i | \psi_t \rangle|^2 = |c_i(t)|^2 \leq 1,$$

i.e., the probability of each individual eigenvalue is a positive number between zero and one, as required for a probabilistic interpretation.

Implications of Postulate IV. Postulate IV brings out a fundamental difference between classical and quantum mechanics. It indicates that we cannot know the outcome of a measurement with certainty. Instead, only the probability of the outcome of a measurement can be given. This is a fundamentally different situation from classical mechanics. As a corollary, Postulate IV dictates that separate identical measurements carried out on identical systems will not necessarily give the same values each time we measure the observable. Nevertheless, Postulate IV gives quantum mechanics strong predictive capabilities, since we know how to calculate the probability of each particular value for the observable, and therefore, we know or can obtain the probability distribution for the observable.

In particular, we can calculate the average value of the probability distribution for a measurement of an observable. This average value is sometimes called the *expectation value*. The expectation value for an observable represented by the operator \hat{A} is written as $\langle \hat{A} \rangle$ and is defined as for any probability distribution as

$$\langle \hat{A} \rangle \equiv \sum_{i=1}^{\infty} P(a_i) a_i,$$

or as

$$\langle \hat{A} \rangle = \int_k a(k) P(a(k)) dt$$

for continuous eigenvalues. Thus, $\langle \hat{A} \rangle$ is just the weighted sum of the a_i 's.

In terms of the wavefunction $|\psi_t\rangle$, using Postulate IV we have, for discrete eigenvalues,

$$\langle \hat{A} \rangle = \sum_{k=1}^{\infty} |\langle \alpha_k | \psi_t \rangle|^2 a_k$$

or, for continuous eigenvalues,

$$\langle \hat{A} \rangle = \int_k |\langle \alpha(k) | \psi_t \rangle|^2 a(k) dk.$$

Another formula for $\langle \hat{A} \rangle$ can be obtained by considering the quantity,

$$\langle \psi_t | \hat{A} | \psi_t \rangle = \langle \psi_t | \hat{A} \left(\sum_{i=1}^{\infty} |\alpha_i\rangle \langle \alpha_i| \right) | \psi_t \rangle,$$

where the factor in parenthesis is just 1. Evaluating the right side of the equation, we have

$$\begin{aligned} \langle \psi_t | \hat{A} | \psi_t \rangle &= \langle \psi_t | \sum_{i=1}^{\infty} \hat{A} |\alpha_i\rangle \langle \alpha_i | \psi_t \rangle \\ &= \langle \psi_t | \sum_{i=1}^{\infty} a_i |\alpha_i\rangle \langle \alpha_i | \psi_t \rangle \\ &= \sum_{i=1}^{\infty} \langle \psi_t | \alpha_i \rangle \langle \alpha_i | \psi_t \rangle a_i \\ &= \sum_{i=1}^{\infty} \langle \alpha_i | \psi_t \rangle^* \langle \alpha_i | \psi_t \rangle a_i, \end{aligned}$$

and finally,

$$\langle \psi_t | \hat{A} | \psi_t \rangle = \sum_{i=1}^{\infty} |\langle \alpha_i | \psi_t \rangle|^2 a_i$$

which is recognized as just $\langle \hat{A} \rangle$. The quantity $\langle \psi_t | \hat{A} | \psi_t \rangle$ is sometimes called a diagonal *matrix element*; it is diagonal because the state vector is the same on both sides of the operator.

We can also calculate higher moments of the probability distribution for the measurement of \hat{A} . Especially useful is the second moment or mean square deviation, σ , given by

$$\sigma^2 = \sum_{i=1}^{\infty} P(a_i) (a_i - \langle \hat{A} \rangle)^2 \equiv (\Delta \hat{A})^2.$$

By analogy with the analysis immediately above,

$$\begin{aligned} (\Delta \hat{A})^2 &= \langle (\hat{A} - \langle \hat{A} \rangle)^2 \rangle = \langle \psi | (\hat{A}^2 - 2\langle \hat{A} \rangle \hat{A} + \langle \hat{A} \rangle^2) | \psi \rangle \\ &= \langle \psi | \hat{A}^2 | \psi \rangle - 2\langle \hat{A} \rangle \langle \psi | \hat{A} | \psi \rangle + \langle \hat{A} \rangle^2 \langle \psi | \psi \rangle, \end{aligned}$$

and we have the useful relationship,

$$(\Delta \hat{A})^2 = \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2$$

$\Delta \hat{A}$ is called the uncertainty in the measurement of observable A, and it is used in the Heisenberg uncertainty principle developed later.

Problem: Prove that if $(\Delta \hat{A})^2 = 0$ for the state function $|\psi\rangle$, then $|\psi\rangle$ is an eigenvector of \hat{A} .

$$(\Delta \hat{A})^2 = 0 = \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2,$$

so

$$\langle \hat{A}^2 \rangle = \langle \hat{A} \rangle^2$$

Now,

$$\langle \hat{A}^2 \rangle = \langle \psi | \hat{A}^2 | \psi \rangle = \langle \hat{A} \rangle^2,$$

therefore,

$$\begin{aligned} \langle \hat{A} \psi | \hat{A} \psi \rangle &= \langle \hat{A} \rangle \langle \psi | \hat{A} \psi \rangle = \langle \hat{A} \rangle \langle \hat{A} \psi | \psi \rangle, \\ \hat{A} | \psi \rangle &= \langle \hat{A} \rangle | \psi \rangle. \end{aligned}$$

Therefore, $|\psi\rangle$ is an eigenvector of \hat{A} with eigenvalue $\langle\hat{A}\rangle$. Also, prove the converse.

Ehrenfest's theorem shows that quantum mechanics reduces to classical mechanics for systems for which the uncertainty is negligible. The theorem is essentially Newton's Law written in terms of expectation values,

$$\frac{d\langle\vec{p}\rangle}{dt} = -\langle\vec{\nabla}V\rangle \quad (\text{Ehrenfest's theorem}),$$

where $\langle\vec{p}\rangle$ represents the expectation value of the momentum, $\langle\psi|\vec{p}|\psi\rangle$, $-\vec{\nabla}V$ is the force and $\langle\psi|(-\vec{\nabla}V)|\psi\rangle$ its expectation value. Ehrenfest's theorem also justifies writing a quantum mechanical operator as the same function of position and momentum operators, $f(\hat{q}_1, \dots, \hat{q}_n, \hat{p}_1, \dots, \hat{p}_n)$, as is done for the classical variable, $f(q_1, \dots, q_n, p_1, \dots, p_n)$.

1.11 Postulate V of Quantum Theory

We are now ready for Postulate V, which explains how the state function $|\psi_t\rangle$ is influenced by the measurement of an observable.

Postulate V. If a measurement of the observable corresponding to the operator \hat{A} yields a value a_i , then the state of the system immediately after the measurement is given by

$$|\psi_t\rangle = \sum_{k=1}^n c_i^k |\alpha_i^k\rangle \quad (n\text{-fold degenerate case}),$$

where the sum is over all states k with the degenerate eigenvalue, a_i , and the linear coefficients c_i^k are chosen so that $|\psi_t\rangle$ is normalized.) If the eigenvalues are all non-degenerate, then

$$|\psi_t\rangle = |\alpha_i\rangle \quad (\text{non-degenerate case}).$$

In the non-degenerate case, the postulate simply states that if a measurement of A yields a_k , then the state of the system immediately after the measurement is completed is

$$|\psi_t\rangle = |\alpha_k\rangle,$$

i.e., the eigenvector corresponding to the eigenvalue that was measured.

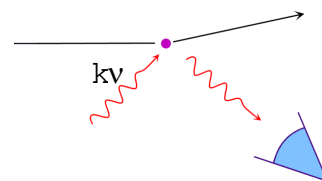
In general, Postulates IV and V mean that a measurement can have a number of possible outcomes although the state of the system is completely defined. However, once the measurement has been made, the state vector “collapses” to the eigenvector corresponding to the value obtained. This postulate just formalizes the experimental fact that as soon as a measurement of an observable has been made, an subsequent measurement of the same observable gives the same result (value), provided the measurement is performed immediately. That is, if the measurement is made before the system has time to evolve. The postulate just expresses the experimental fact that, at the quantum level, the measurement process itself requires an interaction with the system that cannot be neglected.

1.12 Compatible observables

In classical mechanics, the measurement of the coordinates (q) does not affect the measurement of the momenta (p) of the system. This is not true for quantum mechanics. In quantum theory, the coordinates and momenta of classical mechanics are replaced by quantum mechanical operators,

$$q_i \rightarrow q_i$$

$$p_i \rightarrow -i\hbar \frac{\partial}{\partial q_i}$$



and such operators, in general, may not possess simultaneous eigenfunctions.

Problem: Suppose we have the eigenvectors of the i^{th} coordinate, that is, we have solved the eigenvalue problem,

$$\hat{q}_i |\theta_i\rangle = q_i |\theta_i\rangle$$

what if

$$|\theta_i\rangle = \sum_{i=1}^{\infty} c_i |\rho_i\rangle,$$

where the $|\rho_i\rangle$ satisfy the momentum eigenvalue problem,

$$\hat{p}_i |\rho_i\rangle = p_i |\rho_i\rangle.$$

Assume also that

$$|\rho_i\rangle = \sum_{i=1}^{\infty} c'_i |\theta_i\rangle$$

Consider two operators \hat{A} and \hat{B} with eigenfunctions $\{|\alpha_i\rangle\}$ and $\{|\beta_i\rangle\}$, assuming non-degenerate eigenvalues a_i and b_i . Then a measurement of A will give one of the a_i , say a_k , and change the state of the system to $|\psi_i\rangle = |\alpha_k\rangle$. Now, in general, a subsequent measurement of B will give the eigenvalue b_j with probability,

$$P(b_i) = \left| \langle \beta_j | \psi_t \rangle \right|^2 = \left| \langle \beta_j | \alpha_k \rangle \right|^2.$$

And, of course, an immediate second measurement of A, gives a_k with probability

$$P_{2^{nd}}(a_k) = \left| \langle \alpha_k | \beta_j \rangle \right|^2 = \left| \langle \alpha_k | \alpha_k \rangle \right|^2 = 1.$$

Thus, the probability will be unity for only one value of b_j , if and only if $|\alpha_k\rangle$ corresponds to one of the eigenvectors of \hat{B} , say $|\beta_j\rangle$. Thus, we can know both A and B with certainty (probability = 1) simultaneously, if and only if the eigenvectors of A are also the eigenvectors of B.

The observable A and B are said to be *compatible*, if and only if \hat{A} and \hat{B} possess a complete, orthonormal set of simultaneous eigenfunctions.

Commuting Observables. Suppose we have two compatible observables A and B, then

$$\hat{A}\hat{B}|\beta_i\rangle = b_i\hat{A}|\beta_i\rangle.$$

Further, since $|\beta_i\rangle$ is also an eigenvector of A, say $|\beta_i\rangle = |\alpha_k\rangle$, then

$$\hat{A}\hat{B}|\beta_i\rangle = b_i\hat{A}|\alpha_k\rangle = b_ia_k|\beta_i\rangle.$$

Similarly, operating with \hat{B} and \hat{A} in the reverse order, we have

$$\hat{B}\hat{A}|\beta_i\rangle = \hat{B}\hat{A}|\alpha_k\rangle = a_k\hat{B}|\alpha_k\rangle = a_k\hat{B}|\beta_i\rangle = a_kb_i|\beta_i\rangle.$$

Now, since a_k and b_i are just real numbers,

$$a_kb_i = b_ia_k,$$

thus,

$$\hat{A}\hat{B}|\beta_i\rangle = \hat{B}\hat{A}|\beta_i\rangle$$

$$(\hat{A}\hat{B} - \hat{B}\hat{A})|\beta_i\rangle = 0.$$

So, we can write the operator equation,

$$\hat{A}\hat{B} - \hat{B}\hat{A} = 0,$$

and we say that operators that obey such an equation commute. The quantity $\hat{A}\hat{B} - \hat{B}\hat{A}$ occurs so often in quantum mechanics we define the commutator $[\hat{A}, \hat{B}]$,

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}.$$

Consequently, if we have simultaneous eigenvectors for \hat{A} and \hat{B} , then $[\hat{A}, \hat{B}] = 0$, and the converse is true as well. Thus, we can state the condition that A and B are compatible observables in these equivalent ways. A and B are compatible, if and only if,

1. \hat{A} and \hat{B} possess, a complete, orthonormal set of simultaneous eigenvectors, or equivalently,
2. $[\hat{A}, \hat{B}] = 0$.

The conjugate coordinates and momenta are used to specify the classical state of a system of particles. Are \hat{q}_i and \hat{p}_i compatible? Equivalently, do \hat{q} and \hat{p} commute? Consider the effect of the product of these operators on an arbitrary state vector $|\psi\rangle$,

$$\hat{q}\hat{p} = q \left(-i \frac{\hbar \partial}{\partial q} \right) |\psi\rangle = -i\hbar q \frac{\partial \psi(q)}{\partial q} = +i\hbar \psi(q) - i\hbar \frac{\partial}{\partial q} (q\psi(q)),$$

where we have used

$$\frac{\partial}{\partial q} (q\psi(q)) = \psi(q) + q \frac{\partial \psi(q)}{\partial q}.$$

Therefore,

$$q \left(-i \frac{\hbar \partial}{\partial q} \right) |\psi\rangle = i\hbar |\psi(q)\rangle - i\hbar \frac{\partial}{\partial q} (q\psi(q)).$$

Now,

$$\hat{p}\hat{q} = \left(-i \frac{\hbar \partial}{\partial q} \right) q |\psi\rangle = -i\hbar \frac{\partial}{\partial q} (q\psi(q)).$$

Subtracting the last equation from the next to the last, we see

$$\hat{q}\hat{p}|\psi\rangle - \hat{p}\hat{q}|\psi\rangle = i\hbar|\psi\rangle$$

$$(\hat{q}\hat{p} - \hat{p}\hat{q})|\psi\rangle = i\hbar|\psi\rangle$$

which gives the operator relationship

$$[\hat{q}, \hat{p}] = i\hbar.$$

Therefore, \hat{p} and \hat{q} do not commute, and, therefore, they are not compatible. This means that the coordinates and momenta cannot be simultaneously known with probability one. Thus, if we measure q exactly we can give only the probability less than unity for a subsequent and immediate measurement of p .

1.13 Heisenberg Uncertainty Principle

We now derive a general relationship that describes how the measurement of one observable influences our ability to know another. We do this for a particular class non-compatible observables. This relationship is expressed in terms of the mean square deviation in the probability distribution of the two observables: $(\Delta\hat{A})^2$ and $(\Delta\hat{B})^2$.

Consider two non-commuting Hermitian operators \hat{A} and \hat{B} , such that

$$[\hat{A}, \hat{B}] = i\hat{C}, \quad \text{where } \hat{C} \text{ is Hermitian.}$$

Note this includes the commutator relationship we just derived for the coordinates and momenta. We now define new observables, *i.e.*, Hermitian operators, that have zero average values. Their operators are:

$$\hat{A}' = \hat{A} - \langle \hat{A} \rangle \hat{I},$$

$$\hat{B}' = \hat{B} - \langle \hat{B} \rangle \hat{I},$$

$$\hat{C}' = \hat{A}\hat{B} + \hat{B}\hat{A},$$

and we also define the operator \hat{D}' , which is not Hermitian,

$$\hat{D}' = \hat{A}' + (a + bi)\hat{B}',$$

where a and b are real numbers. Notice that $\langle \hat{A}' \rangle = \langle \psi | \{ \hat{A} - \langle \hat{A} \rangle \hat{I} \} | \psi \rangle = \langle \hat{A} \rangle - \langle \hat{A} \rangle = 0$, and, similarly, $\langle \hat{B}' \rangle = 0$.

First, we will show \hat{A}' , \hat{B}' , and \hat{C}' are Hermitian. (1) We start by showing that $a\hat{B}$ is Hermitian if \hat{B} is Hermitian and a is real, *i.e.*, if $\hat{B}^\dagger = \hat{B}$, then

$$[a\hat{B}]^\dagger = a\hat{B}.$$

First, consider

$$(\vec{u}, a\hat{B}\vec{v}) = a(\vec{u}, \hat{B}\vec{v}) = a(\hat{B}^\dagger\vec{u}, \vec{v}) = a\frac{a^*}{a^*}(\hat{B}^\dagger\vec{u}, \vec{v})$$

$$(\vec{u}, a\hat{B}\vec{v}) = \frac{a^*}{a^*}(a\hat{B}^\dagger\vec{u}, \vec{v}),$$

and since $\hat{B}^\dagger = \hat{B}$ because \hat{B} is Hermitian we have

$$(\vec{u}, a\hat{B}\vec{v}) = ((a\hat{B})\vec{u}, \vec{v}),$$

since a is real and thus $a^* = a$. Comparing with the definition of a self-adjoint operator,

$$(\vec{u}, a\hat{B}\vec{v}) = ((a\hat{B})^\dagger\vec{u}, \vec{v}),$$

and we see that $[a\hat{B}]^\dagger = a\hat{B}$, so $a\hat{B}$ is Hermitian if \hat{B} is Hermitian and a is real, *i.e.*, a real number times a Hermitian operator is a Hermitian operator.

(2) Next, we show that \hat{I} is Hermitian, where \hat{I} is the identity operator defined by

$$\hat{I}\vec{u} \equiv \vec{u}.$$

Consider

$$(\hat{I}^\dagger\vec{u}, \vec{v}) = (\vec{u}, \hat{I}\vec{v}) = (\vec{u}, \vec{v}) = (\hat{I}\vec{u}, \vec{v})$$

Again, comparing with the definition of a Hermitian operator, we find $\hat{I}^\dagger = \hat{I}$, and the identity operator is Hermitian.

(3) Now let \hat{A} and \hat{B} be Hermitian, and show that $\hat{A} + \hat{B}$ is Hermitian.

$$(\vec{u}, (\hat{A} + \hat{B})\vec{v}) = (\vec{u}, \hat{A}\vec{v}) + (\vec{u}, \hat{B}\vec{v})$$

$$= (\hat{A}, \bar{u}\bar{v}) + (\hat{B}\bar{u}, \bar{v}),$$

since \hat{A} and \hat{B} are Hermitian. Continuing,

$$(\bar{u}, (\hat{A} + \hat{B})\bar{v}) = ((\hat{A} + \hat{B})\bar{u}, \bar{v}),$$

and

$$(\bar{u}, (\hat{A} + \hat{B})\bar{v}) = ((\hat{A} + \hat{B})^\dagger \bar{u}, \bar{v}).$$

Therefore, we have

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

and $\hat{A} + \hat{B}$ is Hermitian, if \hat{A} and \hat{B} are. We could have shown this more simply by using the fact that the adjoint operation is distributive for linear operators,

$$(L_1 + L_2)^\dagger = L_1^\dagger + L_2^\dagger = (L_1 + L_2)$$

We are now set to show that the primed operators are Hermitian. Starting with \hat{A}' ,

$$\begin{aligned} \hat{A}'^\dagger &= (\hat{A} - \langle A \rangle \hat{I})^\dagger = \hat{A}^\dagger + (-\langle A \rangle \hat{I})^\dagger, \\ &= \hat{A} + (-\langle A \rangle \hat{I})^\dagger, \end{aligned}$$

since \hat{A} is Hermitian, and

$$= \hat{A} + (-\langle A \rangle \hat{I}),$$

since a real number times a Hermitian operator is Hermitian. We now have

$$\hat{A}'^\dagger = \hat{A} + (-\langle A \rangle \hat{I})$$

$$\hat{A}'^\dagger = (\hat{A} - \langle A \rangle \hat{I})^\dagger = \hat{A}',$$

and \hat{A}' is thus Hermitian. Similarly, \hat{B}' can be shown to be Hermitian.

We now show that \hat{C}' is Hermitian, using $(L_1 L_2)^\dagger = L_2^\dagger L_1^\dagger$ and $(L_1 + L_2)^\dagger = L_1^\dagger + L_2^\dagger$.

$$\begin{aligned}\hat{C}'^\dagger &= (\hat{A}\hat{B} + \hat{B}\hat{A})^\dagger \\ &= (\hat{A}\hat{B})^\dagger + (\hat{B}\hat{A})^\dagger \\ &= \hat{B}^\dagger \hat{A}^\dagger + \hat{A}^\dagger \hat{B}^\dagger \\ &= \hat{A}\hat{B} + \hat{B}\hat{A}\end{aligned}$$

since \hat{A} and \hat{B} are Hermitian. Therefore, $\hat{C}'^\dagger = \hat{C}'$ is Hermitian.

\hat{D}' is not Hermitian; nevertheless, $\hat{D}'|\psi\rangle$ is a vector in Hilbert space, because $\hat{D}'|\psi\rangle = \hat{A}'|\psi\rangle + (a+bi)\hat{B}'|\psi\rangle$, *i.e.*, $\hat{D}'|\psi\rangle$ is a linear combination of vectors in Hilbert space. As a consequence, the vector $\hat{D}'|\psi\rangle$ has a finite norm, *i.e.*,

$$0 \leq \langle \hat{D}'\psi | \hat{D}'\psi \rangle < \infty$$

Substituting,

$$\begin{aligned}\langle [\hat{A}' + (a+bi)\hat{B}']\psi | [\hat{A}' + (a+bi)\hat{B}']\psi \rangle &\geq 0 \\ \langle \hat{A}'\psi | \hat{A}'\psi \rangle + \langle (a+bi)\hat{B}'\psi | (a+bi)\hat{B}'\psi \rangle + \langle (a+bi)\hat{B}'\psi | \hat{A}'\psi \rangle + \langle \hat{A}'\psi | (a+bi)\hat{B}'\psi \rangle &\geq 0 \\ \langle \psi | \hat{A}'^2 | \psi \rangle + (a+bi)^*(a+bi)\langle \psi | \hat{B}'^2 | \psi \rangle + (a+bi)^*\langle \psi | \hat{B}'\hat{A}' | \psi \rangle + (a+bi)\langle \psi | \hat{A}'\hat{B}' | \psi \rangle &\geq 0 \\ \langle \hat{A}'^2 \rangle + (a^2 + b^2)\langle \hat{B}'^2 \rangle + a\langle \psi | (\hat{A}'\hat{B}' + \hat{B}'\hat{A}') | \psi \rangle + bi\langle \psi | \hat{A}'\hat{B}' - \hat{B}'\hat{A}' | \psi \rangle &\geq 0.\end{aligned}$$

Using $(\hat{A}'\hat{B}' + \hat{B}'\hat{A}') = \hat{C}'$ and $\hat{A}'\hat{B}' - \hat{B}'\hat{A}' = i\hat{C}$, we have

$$\langle \hat{A}'^2 \rangle + (a^2 + b^2)\langle \hat{B}'^2 \rangle + a\langle \hat{C}' \rangle + b\langle i(\hat{A}'\hat{B}' - \hat{B}'\hat{A}') \rangle \geq 0,$$

but we also have

$$\hat{A}'\hat{B}' - \hat{B}'\hat{A}' \equiv [\hat{A}', \hat{B}'] = (\hat{A} - \langle \hat{A} \rangle \hat{I})(\hat{B} - \langle \hat{B} \rangle \hat{I}) - (\hat{B} - \langle \hat{B} \rangle \hat{I})(\hat{A} - \langle \hat{A} \rangle \hat{I})$$

$$\begin{aligned}
&= \hat{A}\hat{B} - \hat{B}\hat{A} - \langle \hat{A} \rangle \langle \hat{B} \rangle - \langle \hat{B} \rangle \langle \hat{A} \rangle + \langle \hat{A} \rangle \langle \hat{B} \rangle + \langle \hat{A} \rangle \langle \hat{B} \rangle + \langle \hat{B} \rangle \langle \hat{A} \rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle \\
&= \hat{A}\hat{B} - \hat{B}\hat{A} = [\hat{A}, \hat{B}] = i\hat{C}.
\end{aligned}$$

Therefore,

$$\langle \hat{A}'^2 \rangle + (a^2 + b^2) \langle \hat{B}'^2 \rangle + a \langle \hat{C}' \rangle - b \langle \hat{C}' \rangle \geq 0$$

Now, a and b are arbitrary constants, so we can choose,

$$a = -\frac{\langle \hat{C}' \rangle}{2\langle \hat{B}'^2 \rangle} \text{ and } b = +\frac{\langle \hat{C}' \rangle}{2\langle \hat{B}'^2 \rangle}.$$

Substituting,

$$\begin{aligned}
\langle \hat{A}'^2 \rangle + \left(\frac{\langle \hat{C}' \rangle^2}{4\langle \hat{B}'^2 \rangle^2} + \frac{\langle \hat{C}' \rangle^2}{4\langle \hat{B}'^2 \rangle^2} \right) \langle \hat{B}'^2 \rangle - \frac{\langle \hat{C}' \rangle^2}{2\langle \hat{B}'^2 \rangle} - \frac{\langle \hat{C}' \rangle^2}{2\langle \hat{B}'^2 \rangle} &\geq 0, \\
\langle \hat{A}'^2 \rangle - \left(\frac{\langle \hat{C}' \rangle^2 + \langle \hat{C}' \rangle^2}{4\langle \hat{B}'^2 \rangle} \right) &\geq 0, \\
\langle \hat{A}'^2 \rangle &\geq \left(\frac{\langle \hat{C}' \rangle^2 + \langle \hat{C}' \rangle^2}{4\langle \hat{B}'^2 \rangle} \right) \\
\langle \hat{A}'^2 \rangle \langle \hat{B}'^2 \rangle &\geq \frac{\langle \hat{C}' \rangle^2}{4} + \frac{\langle \hat{C}' \rangle^2}{4}.
\end{aligned}$$

We now cast this expression in terms of the uncertainties or mean square deviations of the distributions for these observables, *i.e.*,

$$(\Delta \hat{A}')^2 = \langle \hat{A}'^2 \rangle - \langle \hat{A}' \rangle^2,$$

However, remember that we constructed \hat{A}' and \hat{B}' such that $\langle \hat{A}' \rangle = \langle \hat{B}' \rangle = 0$; therefore,

$$(\Delta\hat{A}')^2 = \langle \hat{A}'^2 \rangle = \langle \psi | (\hat{A} - \langle A \rangle \hat{I}) | \psi \rangle = \langle \hat{A}^2 \rangle - 2\langle A \rangle^2 + \langle A \rangle^2 = \langle \hat{A}^2 \rangle - \langle A \rangle^2 = \langle \Delta\hat{A} \rangle^2,$$

and similarly,

$$(\Delta\hat{B}')^2 = \langle \hat{B}'^2 \rangle = (\Delta\hat{B})^2.$$

Using these relationships between the uncertainty of \hat{A} and \hat{A}' and \hat{B} and \hat{B}' , the product of uncertainties becomes,

$$(\Delta\hat{A})^2 (\Delta\hat{B})^2 \geq \frac{\langle \hat{C}' \rangle^2}{4} + \frac{\langle \hat{C} \rangle^2}{4} \quad (\text{Uncertainty relationship}).$$

Simplifying,

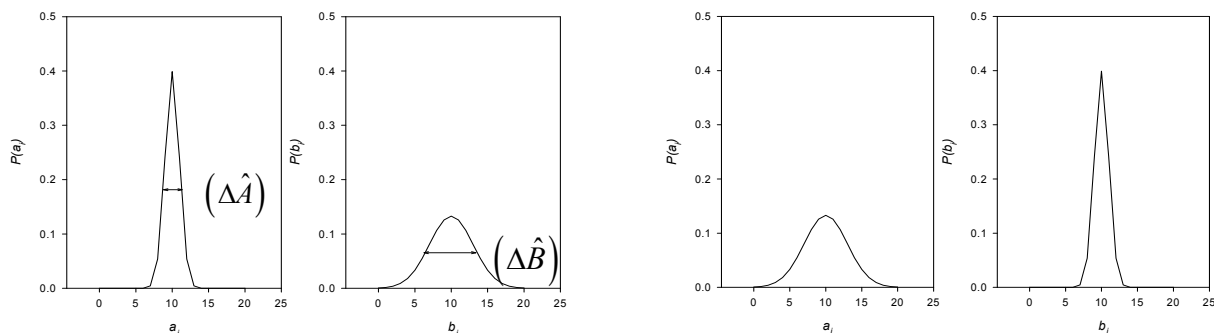
$$(\Delta\hat{A})^2 (\Delta\hat{B})^2 \geq \frac{\langle \hat{C} \rangle^2}{4}$$

is also true. Finally, taking the square root of both sides of the inequality, we have

$$(\Delta\hat{A})(\Delta\hat{B}) \geq \frac{\langle \hat{C} \rangle}{2} \quad (\text{Heisenberg Uncertainty Principle}).$$

Now, remember that $\hat{C} \equiv -i[\hat{A}, \hat{B}]$; therefore, if \hat{A} and \hat{B} don't commute, then the magnitude $\langle \hat{C} \rangle$ determines how well we can simultaneously know the observables A and B. The uncertainty in A ($\Delta\hat{A}$) and the uncertainty in B ($\Delta\hat{B}$), cannot both be zero, if $\langle \hat{C} \rangle$ is non-zero.

Proof: Suppose $(\Delta\hat{A}) \rightarrow 0$, then $(\Delta\hat{B}) \geq \frac{\langle \hat{C} \rangle}{2} \left(\frac{1}{(\Delta\hat{A})} \right) \rightarrow \infty$.



Therefore, the better we know A, the worse is the uncertainty in B, and *vice versa*.

Specific physical examples of such incompatible variables that obey the Heisenberg Uncertainty Principle are conjugate coordinates and momenta. We showed earlier that \hat{q} and

$\hat{p} \equiv -i\hbar \frac{\partial}{\partial q}$ satisfy the commutator relation:

$$[\hat{q}, \hat{p}] = i\hbar.$$

This relationship falls into the form of our assumed relationship between the operators \hat{A} and \hat{B} , *i.e.*, $[\hat{A}, \hat{B}] = i\hat{C}$, or, in our case, $[\hat{q}, \hat{p}] = i(\hbar\hat{I})$, thus, $\hat{C} = \hbar\hat{I}$. Then, \hat{p} and $\langle \hat{A} \rangle$ satisfy the uncertainty relationship,

$$(\Delta\hat{p})(\Delta\hat{q}) \geq \frac{\langle \hat{C} \rangle}{2} = \frac{\langle \psi | \hbar\hat{I} | \psi \rangle}{2} = \frac{\hbar}{2} \langle \psi | \psi \rangle = \frac{\hbar}{2}$$

$$(\Delta\hat{p})(\Delta\hat{q}) \geq \frac{\hbar}{2}.$$

This relationship holds for any generalized coordinate and momentum, for example, the angle of rotation, θ , and the angular p_θ momentum.

Let's now review these results for the uncertainties of such observables. We showed that two observables that obey a commutator relation

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

where \hat{A} , \hat{B} , and \hat{C} are Hermitian, also obey the Heisenberg Uncertainty relationship

$$(\Delta\hat{A})(\Delta\hat{B}) \geq \frac{\langle \hat{C} \rangle}{2}.$$

Furthermore, any generalized coordinate and its conjugate momentum obey the particular form of this expression where $\langle \hat{C} \rangle$ is \hbar ,

$$(\Delta\hat{q})(\Delta\hat{p}) \geq \frac{\hbar}{2}.$$

This relationship implies that if we make a perfect measurement of the position of a particle, which can be done in principle, then the momentum of the particle is necessarily completely undetermined, *i.e.*, the width of the probability distribution of the momentum is

infinitely broad. Therefore, since classically to define the state of the system we must give both the coordinates and momenta at some time, we really can never know the classical state of a system. Further, in classical mechanics we must have the coordinates and momenta at $t = 0$, *i.e.*, the initial conditions, so that we can use $\vec{F} = m\vec{a}$ (or an equivalent expression) to make future predictions of the state of the system.

Thus, quantum mechanics tells us:

1. that the classical state of a system is indeterminate, and
2. that because we can only give the probability distribution for the initial state, we can only give the probability of the system occupying some state, defined by $\{q_1, q_2, \dots, q_n\}$ and $\{p_1, p_2, \dots, p_n\}$, at some later time t . (This situation is similar to that of statistical mechanics. The classical state at some later time is given probabilistically by integrating the equations of motion ($\vec{F} = m\vec{a}$, or equivalently Hamilton's Equations of Motion) for each possible initial state of the system.)

Problem Set

1. Show $\hat{p}_x \equiv -i\hbar \frac{d}{dx}$ is Hermitian.
2. Prove Ehrenfest's Theorem $\frac{d\langle \vec{p} \rangle}{dt} = -\langle \vec{\nabla} V \rangle$.
3. Prove that the eigenvalues of a Hermitian operator are real numbers.
4. Show that $|\varepsilon_k\rangle = A \sin(kx) + B \cos(kx)$ for $k = 1, \dots, n, \dots$ are eigenfunctions of the x -component of the kinetic energy of a particles of mass m , $\hat{T}_x = \frac{\hat{p}_x^2}{2m}$.
5. Show that if $(\Delta \hat{A})^2 = 0$ for the state $|\psi\rangle$, then $|\psi\rangle$ is an eigenfunction of \hat{A} . In addition, show the converse.
6. Show that observables A and B are compatible, if and only if, \hat{A} and \hat{B} possess a complete orthonormal set of simultaneous eigenvectors, or equivalently, $[\hat{A}, \hat{B}] = 0$
7. Show that \hat{q} and \hat{p} don't commute and are, thus, not compatible.

1.14 Alternative Postulates of Quantum Theory

For the postulates, developed so far, alternative ones are more fundamentally connected to the measurement process. The measurement process, no matter how complex, can always be reduced to a series of "Yes-No" experiments. For example, a multi-channel spectral analyzer can be thought of as many individual yes-no experiments—one for each channel. The individual channels define a yes-no experiment and together they define an energy or position spectrum. The yes-no experiment for each channel is "did the particle land in this channel or not?" Yes or No? If yes, then the particles energy is a certain value because of the position at which it was detected. The position is related to the energy by the dispersion of the

spectrometer.

The question “did the particle land in this channel or not?” is called a *proposition* (which is either true or false). The quantum physics is in the fundamental relationships between such propositions. Let’s assume some very basic relationships or axioms hold for these propositions, for example.

I. Let a and b be two proposition of physical system, then “ a implies b ” ($a \subseteq b$) means that if a is true, then b is also (definition of $a \subseteq b$). The relation called “ \leq ” satisfies,

1. $a \subseteq a$
2. $a \subseteq b ; b \subseteq a$, then $a = b$
3. $a \subseteq b ; b \subseteq c$, then $a \subseteq c$

The set of all propositions (yes-no experiments) is partially ordered by these reasonable relationships.

II. There exist a proposition, $\bigcap_{i \in I} a_i$, which has the property that if

$$x \subseteq a_i \text{ for all } i \in I, \text{ then } x \subseteq \bigcap_I a_i,$$

at least in principle (problem for complementary observables). For example, $x = a \cap b$ is true implies that a is true and b is true. The latter axiom implies the absurd proposition

$$\phi = \bigcap_{a \in L} a,$$

is always false.

III. Ortho complementation. For every a , there exist a' (“not a ”) so that,

1. $(a')' = a$
2. $a' \cap a = \phi$
3. $a \leq b \Rightarrow b' \leq a'$

The axioms listed are reasonable for the discussion of physical systems. They form what in mathematics is called a “complete, ortho-complemented lattice”. Axioms I, II, and III formalize a set of empirical relations exhibited when making measurements on physical systems. They are also similar to the rules of ordinary logic.

Up to now there are no specific properties of the physical systems to which the axioms apply; these are included in two more axioms that are needed to describe the relationships between proposition for physical systems.

A. Atomicity (Lattice of classes of subsets of propositions).

1. For any proposition $a \neq \phi$, there exist a “minimal proposition” P such that $P \subseteq a$.
2. If Q is a point, then $a \leq x \leq a \cup Q \Rightarrow x = a$ or $x = a \cup Q$.

Taken together I, II, III, and A define the lattice of the propositions for a classical physical system, a lattice that is said to be a “ σ -complete, atomic Boolean lattice”. For quantum mechanics, we need an additional axiom.

P. Weak Modularity. We postulate that $a \subseteq b$ implies that a is compatible with b ($a \leftrightarrow b$).

One can show that our proposition system, defined as a lattice of propositions that satisfies axioms I, II, III, A, P, can be represented as a lattice of subspaces in some linear vector space. If we choose the complex numbers as the elements of our vector, we get conventional quantum mechanics in Hilbert space. That is, the lattice of subspaces of our proposition system is isomorphous with Hilbert space (matrix mechanics). We have already shown that Hilbert space is isomorphous with an inner-product function space (wave mechanics). It is also true that the kinematic structure that is entirely independent of the state of the system is expressed in the structure of this proposition system. The equivalence between this proposition system and Hilbert space carries over to equivalence of various elements of the two systems. In particular, elementary propositions are represented by projection operators in Hilbert space, intersection \leftrightarrow set intersection in Hilbert space, unions \leftrightarrow closed linear span of subspaces, *etc.* This discussion of an alternate set of postulates for quantum mechanics is of course incomplete, but is included to illustrate how basic the structure of quantum mechanics is and how closely related to the measurement process it is.

1.15 Postulate III of Quantum Theory

Postulate III tells us how the quantum state of a system $|\psi_t\rangle$ changes with time.

Postulate III. For every physical system, there exist a linear Hermitian operator, called the Hamiltonian, which determines the time variation of the state vector $\psi(\vec{q}, t)$ through the time-dependent Schrödinger equation,

$$\hat{H}(q_1, q_2, \dots, q_n; t)\psi(q_1 \dots q_n; t) = i\hbar \frac{\partial \psi(q_1, \dots, q_n; t)}{\partial t},$$

during any time interval in which the system is not disturbed by a measurement.

Properties of the Hamiltonian. For most of the systems we will be interested in \hat{H} will be independent of time and will be the same as the total energy. We will later discuss some of

the conditions under which \hat{H} is just the total energy. [Essentially, these conditions are that $H \neq H(t) \Rightarrow (1) V \neq V(t; \dot{q}_j)$ and (2) $\vec{r}_i \neq \vec{r}_i(t)$].

Our next concern is how to construct the Hamiltonian operator, \hat{H} . We will see later that the Hamiltonian is defined in classical mechanics by

$$H \equiv \sum_{i=1}^n p_i \dot{q}_i - L.$$

The corresponding operator is constructed by replacing q_i and p_i by the corresponding operators:

$$\hat{q}_i \leftrightarrow q_i,$$

that is, “multiplication by q_i ”, and

$$\hat{p}_i \leftrightarrow -i\hbar \frac{\partial}{\partial q_i}$$

in the classical Hamiltonian. The Hamiltonian also determines the time evolution of the system in classical mechanics. Because of its importance generally and because we need to find the Hamiltonian for the interaction between a collection of charged particles and the electromagnetic radiation field, we will develop the classical Hamilton’s equations of motion shortly. But, we first further develop the dynamical consequences of Postulate III and Schrödinger’s equation.

We must solve the Schrödinger equation for the state of the system as a function of time. As is the custom with partial differential equations, we assume the different variables can be factored. Thus, we start by assuming $|\psi_i\rangle = \chi(q_1 \dots q_n) \varphi(t)$, that is the statefunction is a product of a time function and a spatial function. Substituting,

$$\hat{H}(q_1 \dots q_n) \chi(q_1 \dots q_n) \varphi(t) = i\hbar \frac{\partial \chi(q_1 \dots q_n) \varphi(t)}{\partial t},$$

$$\varphi(t) \hat{H}(q_1 \dots q_n) \chi(q_1 \dots q_n) = \left(i\hbar \frac{\partial \varphi(t)}{\partial t} \right) \chi(q_1 \dots q_n),$$

$$\frac{1}{\chi(q_1 \dots q_n)} \hat{H}(q_1 \dots q_n) \chi(q_1 \dots q_n) = \frac{i\hbar}{\varphi(t)} \frac{\partial \varphi(t)}{\partial t} = E = \text{constant}.$$

The only way the left and right sides can be equal for all values of q_1, \dots, q_n , and t is if each side

equals the same constant, E . This then gives two equations, one independent of t ,

$$\hat{H}(q_1 \dots q_n) \chi(q_1 \dots q_n) = E \chi(q_1 \dots q_n)$$

(Time-independent Schrödinger Equation),

and another that is independent of the coordinates,

$$i\hbar \frac{\partial \varphi(t)}{\partial t} = E \varphi(t).$$

Since \hat{H} is a linear Hermitian operator, it is an observable, and the values that this observable can have are the eigenvalues of \hat{H} . Thus, the real numbers E that satisfy the time-dependent Schrödinger equation, which is the eigenvalue equation for \hat{H} , also determine the time evolution of the system. Generally, the eigenvalue problem will have more than one solution, so we label the eigenvalue E with a subscript, E_i ,

$$\hat{H} \chi_i = E_i \chi_i,$$

and thus we have different time dependence according to the specific eigenvalue of \hat{H} , that is,

$$i\hbar \frac{\partial \varphi_i(t)}{\partial t} = E_i \varphi_i(t).$$

The solutions of Schrödinger's equation will be just the product of the eigenfunctions that satisfy these two independent equations

$$|\psi(q_1 \dots q_n; t)\rangle = \chi_i(q_1 \dots q_n) \varphi_i(t)$$

Solution of time eigenvalue problem. We can easily solve for the time factor by noting that,

$$i\hbar \frac{\partial \varphi_j(t)}{\partial t} = E_j \varphi_j(t)$$

is satisfied by

$$\varphi_j(t) = e^{-iE_j t/\hbar}$$

To show this, note that

$$\begin{aligned}\frac{i\hbar\partial\varphi_j(t)}{\partial t} &= i\hbar\frac{\partial e^{-iE_j t/\hbar}}{\partial t} = i\hbar\left(-\frac{iE_j}{\hbar}\right)e^{-iE_j t/\hbar} \\ &= E_j e^{-iE_j t/\hbar} \\ \frac{i\hbar\partial\varphi_j(t)}{\partial t} &= E_j\varphi_j(t).\end{aligned}$$

Thus, the general solutions of Schrödinger's equation can be written as

$$|\psi_j(q_1 \dots q_n; t)\rangle = e^{-iE_j t/\hbar} |\chi_j(q_1 \dots q_n)\rangle.$$

Thus, when $\hat{H} \neq \hat{H}(t)$ we only have to solve for the eigenvalues and eigenfunctions of \hat{H} , *i.e.*, solve the time-independent Schrödinger equation,

$$\hat{H}|\chi_i\rangle = E_i|\chi_i\rangle,$$

to find the state $|\psi_j\rangle = |\varphi_j(t)\rangle|\chi_j(q)\rangle$ at any time t and the energy values that can be found upon a measurement of the energy. In addition, Postulates IV and V apply to the energy of the system E , and, in most cases, \hat{H} is the Hermitian operator that corresponds to E .

Is $\varphi_j(t)$ normalized? To answer this question, consider,

$$\begin{aligned}\langle\varphi_j|\varphi_j\rangle &= \varphi_j^*(t)\varphi_j(t) = |\varphi_j(t)|^2 = e^{+iE_j t/\hbar} e^{-iE_j t/\hbar} = e^{-i(E_j - E_j)t/\hbar} \\ &= e^0 = 1.\end{aligned}$$

Thus, the time-dependent wavefunction $|\psi\rangle = \varphi_i(t)\chi_i(q_1 \dots q_n)$ will be normalized provided the spatially dependent part $\chi_i(q_1 \dots q_n)$ is normalized. Notice that there is no integration over time; integration in $\langle\psi_t|\psi_t\rangle$ is over all coordinates, not t .

The states,

$$|\psi_j\rangle = e^{-iE_j t/\hbar} |\chi_j(q_1 \dots q_n)\rangle,$$

are called *stationary states*, since only the phase factor, $e^{-iE_j t/\hbar}$, of magnitude one changes with time. The coordinate dependent part does not change with time.

Since we have now solved for the time evolution of the system, we can ask the general question of whether probability is conserved when a state varies in time by the time-dependent Schrödinger equation. It of course must be conserved, and we now show this.

Conservation of probability. We must show that when $|\psi_t\rangle$ is normalized at some time t it remains normalized for all times. That is,

$$\begin{aligned} \frac{d}{dt}\langle\psi_t|\psi_t\rangle &= \left\langle\frac{\partial\psi_t}{\partial t}|\psi_t\right\rangle + \left\langle\psi_t\left|\frac{\partial\psi_t}{\partial t}\right.\right\rangle \\ &= \left\langle-\frac{i}{\hbar}\hat{H}\psi_t|\psi_t\right\rangle + \left\langle\psi_t\left|-\frac{i}{\hbar}\hat{H}\psi_t\right.\right\rangle \\ &= \frac{i}{\hbar}\left(\langle\hat{H}\psi_t|\psi_t\rangle - \langle\psi_t|\hat{H}\psi_t\rangle\right) \\ &= \frac{i}{\hbar}\left(\langle\psi_t|\hat{H}\psi_t\rangle - \langle\psi_t|\hat{H}\psi_t\rangle\right) \\ &= \frac{d}{dt}\langle\psi_t|\psi_t\rangle = 0, \end{aligned}$$

and, therefore, $\langle\psi_t|\psi_t\rangle = \text{constant}$. Thus, if that constant is initially one, then the state will always be normalized and probability is conserved.

1.16 Time Evolution Operator

Let us now define an operation $\hat{U}(t)$, which changes the state vector at $t=0$ to the state vector at time t . That is,

$$|\psi_t\rangle = \hat{U}(t)|\psi_0\rangle$$

Notice that $\hat{U}(0) = \hat{I}$, the identity operator. $\hat{U}(t)$ “evolves” the system at $t=0$ to the system at time t , i.e., it “moves state vector from an initial orientation in Hilbert space to its orientation at time t .”

Let us now find an expression for $\hat{U}(t)$. It will probably involve \hat{H} , since \hat{H} determines the time evolution via Schrödinger’s equation. Substituting for $|\psi_t\rangle$,

$$\hat{H}(\hat{U}(t)|\psi_0\rangle) = i\hbar \frac{\partial}{\partial t} \hat{U}(t)|\psi_0\rangle$$

or

$$\left(\hat{H}\hat{U}(t) - i\hbar \frac{\partial}{\partial t} \hat{U}(t) \right) |\psi_0\rangle = 0, \text{ for all } |\psi_0\rangle$$

Therefore, the time evolution operator satisfies the operator equation,

$$\hat{H}\hat{U}(t) = i\hbar \frac{\partial}{\partial t} \hat{U}(t).$$

The solution can be written formally as

$$\hat{U}(t) = e^{-i\hat{H}t/\hbar} \quad (\text{when } H \neq H(t)),$$

where $e^{-i\hat{H}t/\hbar}$ is an operator defined by the linear combination operators,

$$\hat{U}(t) = \sum_{n=0}^{\infty} \frac{(-t/\hbar \hat{H})^n}{n!} \equiv e^{-i\hat{H}t/\hbar} \quad (\text{power series expansion}).$$

Problem: Show that $\hat{U}(t)$ is linear and unitary ($\hat{U}^\dagger = \hat{U}^{-1}$), and thus not Hermitian.

If the time-independent Schrödinger equation has been solved for the eigenfunctions $|\psi_k\rangle$ and eigenvalues E_k then the time-evolution operator satisfies the eigenvalue equation

$$\hat{U}(t)|\psi_k\rangle = e^{-iE_k t/\hbar} |\psi_k\rangle$$

We can prove this using the definition of $\hat{U}(t)$ as an expansion.

We can get a matrix representation of $\hat{U}(t)$ by multiplying on the left by another of the stationary states and integrating,

$$\langle \psi_j | \hat{U}(t) | \psi_k \rangle = \langle \psi_j | e^{iE_j t/\hbar} e^{-iE_k t/\hbar} | \psi_k \rangle$$

$$U_{jk}(t) = e^{-i(E_k - E_j)t/\hbar} \delta_{jk},$$

thus the matrix is diagonal and for $t = 0$,

$$U_{jk}(t) = \delta_{jk},$$

i.e., for energy eigenstates, $\hat{U}(t)$ is just the unit matrix (operator) at $t = 0$.

1.17 Constants of Motion

Classically, constants of motion are properties that do not change with time, such as the total energy or angular momentum. The constants of motion have time derivatives that vanish, such as

$$\frac{d(T+V)}{dt} = \frac{dE}{dt} = 0 \quad (\text{total energy}),$$

or

$$\frac{dL}{dt} = 0 \quad (\text{angular momentum}).$$

The angular momentum is a constant of motion when the torque $\vec{N} = 0$.

Is there a quantum mechanical analog of the classical constants of motion? Yes! The quantum mechanical analog is that the time derivative of the expectation value vanishes.

$$\frac{d}{dt} \langle \hat{A} \rangle = 0$$

i.e., the average value of measurements of the observable is constant.

For an observable A , we can derive an expression for the time dependence of the expectation value using Schrödinger's equation.

$$\begin{aligned} \frac{d}{dt} \langle \hat{A} \rangle &\equiv \frac{d}{dt} \langle \psi_t | \hat{A} | \psi_t \rangle \\ &= \left\langle \frac{\partial \psi_t}{\partial t} \middle| \hat{A} \psi_t \right\rangle + \left\langle \psi_t \middle| \frac{\partial \hat{A}}{\partial t} \right\rangle \psi_t + \left\langle \psi_t \middle| \hat{A} \middle| \frac{\partial \psi_t}{\partial t} \right\rangle \\ &= \left\langle -\frac{i}{\hbar} \hat{H} \psi_t \middle| \hat{A} \psi_t \right\rangle + \left\langle \psi_t \middle| \frac{\partial \hat{A}}{\partial t} \right\rangle \psi_t + \left\langle \psi_t \middle| \hat{A} \left(-\frac{i}{\hbar} \hat{H} \psi_t \right) \right\rangle \end{aligned}$$

$$\begin{aligned}
&= \frac{i}{\hbar} \langle \psi_t | \hat{H} \hat{A} | \psi_t \rangle - \frac{i}{\hbar} \langle \psi_t | \hat{A} \hat{H} | \psi_t \rangle + \left\langle \frac{\partial \hat{A}}{\partial t} \right\rangle \\
&= \frac{i}{\hbar} \langle \psi_t | (\hat{H} \hat{A} - \hat{A} \hat{H}) | \psi_t \rangle + \left\langle \frac{\partial \hat{A}}{\partial t} \right\rangle \\
\frac{d}{dt} \langle \hat{A} \rangle &= \frac{i}{\hbar} \langle [\hat{H}, \hat{A}] \rangle + \left\langle \frac{\partial \hat{A}}{\partial t} \right\rangle
\end{aligned}$$

Thus, if $\frac{\partial \hat{A}}{\partial t} = 0$, then $\left\langle \frac{\partial \hat{A}}{\partial t} \right\rangle = 0$. Furthermore, if \hat{A} is also compatible with \hat{H} , then $[\hat{H}, \hat{A}] = 0$. Under these conditions, $\frac{d}{dt} \langle \hat{A} \rangle$ will vanish, and we say that the observable corresponding to \hat{A} is a constant of the motion. In particular, the energy is a constant of the motion if $\hat{H} \neq \hat{H}(t)$ because $[\hat{H}, \hat{H}] = 0$, *i.e.*,

$$\frac{d}{dt} \langle \hat{H} \rangle = 0 \quad (\text{quantum mechanical})$$

$$\frac{dH}{dt} = 0 \quad (\text{classical mechanical}).$$

However, notice that it is not necessary that \hat{A} commute with \hat{H} since we can have $[\hat{H}, \hat{A}] \neq 0$, while $\langle [\hat{H}, \hat{A}] \rangle = 0$. Under this condition the observable is still a constant of the motion.

We have completed our review of quantum theory, so it is helpful to compare the classical and quantum systems.

	Classical	Quantum
<i>State of System</i>	$\left\{ \begin{array}{l} q = q(t) \\ p = p(t) \end{array} \right\}$	$ \psi_t(q_1 \dots q_n)\rangle$
<i>Initial Conditions</i>	$\left\{ \begin{array}{l} q = q(0) \\ p = p(0) \end{array} \right\}$	$ \psi_0(q_1 \dots q_n)\rangle$
<i>Hamiltonian of System</i>	$H(q, p, t)$	$\hat{H}(\hat{q}, \hat{p}, t)$
<i>Time development</i>	$\left\{ \begin{array}{l} \frac{\partial H}{\partial P} = \dot{q} \\ -\frac{\partial H}{\partial q} = \dot{p} \end{array} \right\}$	$\hat{H} \psi_t\rangle = i\hbar \frac{\partial \psi_t\rangle}{\partial t}$

Hamilton's equations of motion will be derived shortly.

We have concentrated on the Schrödinger Picture of quantum mechanics, but other pictures are possible. The interaction picture is especially useful for perturbative treatments of interactions between solved non-interacting subsystems.

	States	Operators	Eigenvectors
<i>Schrödinger</i>	<i>Moving</i>	<i>Stationary</i>	<i>Stationary</i>
<i>Heisenberg</i>	<i>Stationary</i>	<i>Moving</i>	<i>Moving</i>
<i>Interaction</i>	<i>Moving</i>	<i>Moving</i>	<i>Moving</i>

To develop intuition at the quantum level we need to solve many problems of chemical interest as possible. It is our primary task for this course to solve problems of interest in spectroscopy.

1.18 Hamiltonian for a Molecule

Our task in the Chapter 2 is to obtain, from basic principles, the Hamiltonian for the interaction of a radiation field with a collection of charged particles, typically a molecule. We will now write down an expression for the molecular part of this Hamiltonian, using in advance the knowledge that the Hamiltonian is just the total energy of the molecule. The procedures will be put on a firm theoretical basis in the next chapter.

For a molecule made up of N electrons and nuclei, the Hamiltonian operator is needed in order to find the possible energy levels and stationary states of the molecule. The classical molecular Hamiltonian of an isolated molecule is given by

$$H = T + V$$

where

$$V = \sum_{j>k}^N \frac{Z_j Z_k}{r_{jk}}$$

where Z_j is the charge of the j^{th} electron or nucleus and \vec{r}_{jk} is the vector distance between the j^{th} and k^{th} particle. The kinetic energy is given by

$$T = \frac{1}{2} \sum_{j=1}^N m_j v_j^2 = \frac{1}{2} \sum_{j=1}^N \frac{p_j^2}{m_j} = \frac{1}{2} \sum_{j=1}^N \frac{1}{m_j} (p_{xj}^2 + p_{yj}^2 + p_{zj}^2)$$

To get the Hamiltonian operator for the molecule, we make the replacements

$$p_{xj} \rightarrow \hat{p}_{xj} = -i\hbar \frac{\partial}{\partial x_j}$$

and

$$\vec{r}_{jk} \rightarrow \hat{\vec{r}}_{jk} = \vec{r}_{jk},$$

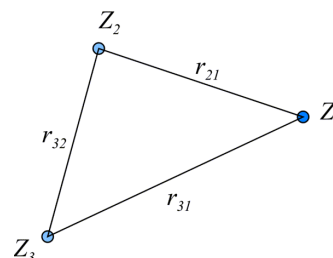
that is, multiplication by the vector position. Thus, the expression for the potential energy operator is the same,

$$\hat{V} = \sum_{j>k}^N \sum_{k=1}^{j-1} \frac{Z_j Z_k}{r_{jk}},$$

and the kinetic energy operator is

$$\hat{T} = -\frac{1}{2} \sum_{j=1}^N \frac{\hbar^2}{m_j} \left(\frac{\partial^2}{\partial x_j^2} + \frac{\partial^2}{\partial y_j^2} + \frac{\partial^2}{\partial z_j^2} \right)$$

$$\hat{T} = -\frac{\hbar^2}{2} \sum_{j=1}^N \frac{\vec{\nabla}_j \cdot \vec{\nabla}_j}{m_j} = -\frac{\hbar^2}{2} \sum_{j=1}^N \frac{\nabla_j^2}{m_j}.$$



Therefore, the molecular Hamiltonian is

$$\hat{H}_{mol} = \hat{T} + \hat{V}$$

or

$$\hat{H}_{mol} = -\frac{\hbar^2}{2} \sum_{j=1}^N \frac{\nabla_j^2}{m_j} + \sum_{j>k}^N \frac{Z_j Z_k}{r_{jk}}$$

The solution of the molecular Schrödinger equation,

$$\hat{H}_{mol} |\psi_i\rangle = E_i |\psi_i\rangle,$$

is very complicated in general and requires approximate separation of the electronic and vibrational motion. Then, the coupling of the vibrational and electronic states by the vibro-electronic (vibronic) interaction must be dealt with to obtain the vibronic states that satisfy the molecular Schrödinger equation. We do this in Chapter 4.