Postulates of Quantum Theory

Quantum Theory can be formulated according to a few *postulates* (i.e., theoretical principles based on experimental observations). The goal of this section is to introduce such principles, together with some mathematical concepts that are necessary for that purpose.**R1(190)** To keep the notation as simple as possible, expressions are written for a 1-dimensional system. The generalization to many dimensions is usually straightforward.

$\boxed{Postulate\ 1}$: *Any system can be described by a [wave function](#) $\psi(t,x)$, where t is a parameter representing the time and x represents the coordinates of the system. Function $\psi(t,x)$ must be continuous, single valued and square integrable.**R1(57)***

Note 1: As a consequence of Postulate 4, we will see that $P(t,x) = \psi^*(t,x)\psi(t,x)dx$ represents the probability of finding the system between x and $x+dx$ at time t.

$\boxed{Postulate\ 2}$: *Any observable (i.e., any measurable property of the system) can be described by an operator. The operator must be linear and hermitian.*

What is an *operator* ? What is a *linear operator* ? What is a *hermitian operator*?

Definition 1: An operator $\hat{O}$ is a mathematical entity that transforms a function $f(x)$ into another function $g(x)$ as follows,**R4(96)**

$$\hat{O}f(x) = g(x),$$

where f and g are functions of x.

Definition 2: R1(190) An operator $\hat{O}$ that represents an observable O is obtained by first writing the classical expression of such observable in Cartesian coordinates (e.g., $O = O(x,p)$) and then substituting the coordinate x in such expression by the *coordinate operator* $\hat{x}$ as well as the momentum p by the *momentum operator* $\hat{p} = -i\hbar\partial/\partial x$.

Definition 3: An operator $\hat{O}$ is *linear* if and only if (iff),

$$\hat{O}(af(x) + bg(x)) = a\hat{O}f(x) + b\hat{O}g(x),$$

where a and b are constants.

Definition 4: An operator $\hat{O}$ is *hermitian* iff, **R1(164)**

$$\int dx \phi_n^*(x) \hat{O} \psi_m(x) = \left[\int dx \psi_m^*(x) \hat{O} \phi_n(x) \right]^*,$$

where the asterisk represents the complex conjugate of the expression embraced by brackets.

Definition 5: A function $\phi_n(x)$ is an *eigenfunction* of $\hat{O}$ iff,

$$\hat{O}\phi_n(x) = O_n \phi_n(x),$$

where O_n is a number called *eigenvalue.*

Property 1: The eigenvalues of a hermitian operator are real.**R1(166)(167)**

Proof: Using Definition 4, we obtain

$$\int dx \phi_n^*(x) \hat{O} \phi_n(x) - \left[\int dx \phi_n^*(x) \hat{O} \phi_n(x) \right]^* = 0,$$

therefore,

$$[O_n - O_n^*] \int dx \phi_n(x)^* \phi_n(x) = 0.$$

Since $\phi_n(x)$ are square integrable functions, then,

$$O_n = O_n^*.$$

Property 2: Different eigenfunctions of a hermitian operator (i.e., eigenfunctions with different eigenvalues) are orthogonal (i.e., the *scalar product* of two different eigenfunctions is equal to zero). Mathematically, if $\hat{O}\phi_n = O_n \phi_n$, and $\hat{O}\phi_m = O_m \phi_m$, with $O_n \neq O_m$, then $\int dx \phi_n^* \phi_m = 0$.

Proof:

$$\int dx\phi_m^*\hat{O}\phi_n - \left[\int dx\phi_n^*\hat{O}\phi_m\right]^* = 0,$$

and

$$[O_n - O_m]\int dx\phi_m^*\phi_n = 0.$$

Since $O_n \neq O_m$, then $\int dx\phi_m^*\phi_n = 0$.

$\boxed{Postulate\ 3}$:*The only possible experimental results of a measurement of an observable are the eigenvalues of the operator that corresponds to such observable.*

$\boxed{Postulate\ 4}$: *The average value of many measurements of an observable* O*, when the system is described by function* $\psi(x)$*, is equal to the expectation value* $\bar{O}$*, which is defined as follows,*

$$\bar{O} = \frac{\int dx\psi(x)^*\hat{O}\psi(x)}{\int dx\psi(x)^*\psi(x)}.$$

$\boxed{Expansion\ Postulate}$***: R1(191), R5(15)), R4(97)*** here

The eigenfunctions of a linear and hermitian operator form a complete basis set. Therefore, any function $\psi(x)$ that is continuous, single valued, and square integrable can be expanded as a linear combination of eigenfunctions $\phi_n(x)$ of a linear and hermitian operator $\hat{A}$ as follows,

$$\psi(x) = \sum_j C_j\phi_j(x),$$

where C_j are numbers (e.g., complex numbers) called *expansion coefficients*.

Exercise 1: Show that $\bar{A} = \sum_j C_jC_j^*a_j$, when $\psi(x) = \sum_j C_j\phi_j(x)$,

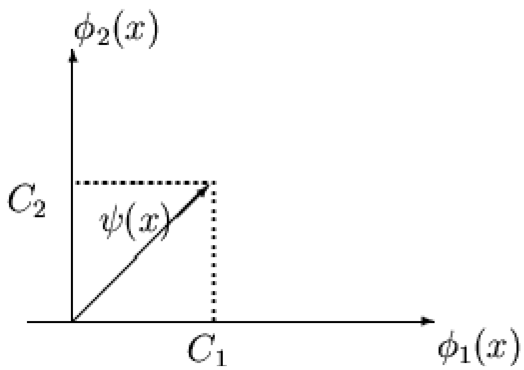
$$\hat{A}\phi_j(x) = a_j\phi_j(x), \quad \text{and} \quad \int dx\phi_j(x)^*\phi_k(x) = \delta_{jk}.$$

Note that (according to Postulate 3) eigenvalues a_j are the only possible experimental results of measurements of $\hat{A}$, and that (according to Postulate 4) the expectation value $\bar{A}$ is the average value of many measurements of $\hat{A}$ when the system is described by the expansion $\psi(x) = \sum_j C_j\phi_j(x)$. Therefore, the product $C_jC_j^*$ can be interpreted as the probability weight associated with eigenvalue a_j (i.e., the probability that the outcome of an observation of $\hat{A}$ will be a_j).

Hilbert-Space

According to the Expansion Postulate (together with Postulate 1), the state of a system described by the *function* $\Psi(x)$ can be expanded as a linear combination of eigenfunctions $\phi_j(x)$ of a linear and hermitian operator (e.g., $\Psi(x) = C_1\phi_1(x) + C_2\phi_2(x) + \ldots$). Usually, the space defined by these eigenfunctions (i.e., functions that are continuous, single valued and square integrable) has an infinite number of dimensions. Such space is called *Hilbert-Space* in honor to the mathematician Hilbert who did pioneer work in spaces of infinite dimensionality.**R4(94)**

A representation of $\Psi(x)$ in such space of functions corresponds to a vector-function,



where C_1 and C_2 are the projections of $\Psi(x)$ along $\phi_1(x)$ and $\phi_2(x)$, respectively. All other components are omitted from the representation because they are orthogonal to the ``plane" defined by $\phi_1(x)$ and $\phi_2(x)$.

Continuous Representation

Certain operators have a continuous spectrum of eigenvalues. For example, the coordinate operator is one such operator since it satisfies the equation $\hat{x}\ \delta(x_0 - x) = x_0\ \delta(x_0 - x)$, where the eigenvalues x_0 define a *continuum*. Delta functions $\delta(x_0 - x)$ define a continuum representation and, therefore, an expansion of $\psi(x)$ in such representation becomes,

$$\psi(x) = \int dx_0 C_{x_0} \delta(x_0 - x),$$

where $C_{x_0} = \psi(x_0)$, since

$$\int dx \delta(x - \beta)\psi(x) = \int dx \int d\alpha C_\alpha \delta(x - \beta)\delta(\alpha - x) = \psi(\beta).$$

According to postulates 3 and 4 (see Exercise 1), the probability of observing the system with coordinate eigenvalues between x_0 and $x_0 + dx_0$ is
$P(x_0) = C_{x_0} C_{x_0}^* dx_0 = \psi(x_0)\psi(x_0)^* dx_0$ (see Note 1).

In general, when the basis functions $\phi(\alpha, x)$ are *not* necessarily delta functions but nonetheless define a continuum representation,

$$\psi(x) = \int d\alpha C_\alpha \phi(\alpha, x),$$

with $C_\alpha = \int dx \phi(\alpha, x)^* \psi(x)$.

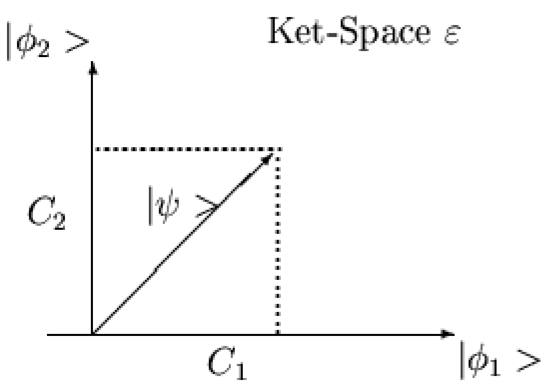
Note 2: According to the Expansion Postulate, a function $\psi(x)$ is uniquely and completely defined by the coefficients C_j, associated with its expansion in a complete set of eigenfunctions $\phi_j(x)$. However, the coefficients of such expansion would be different if the same basis functions ϕ_j depended on different coordinates (e.g., $\phi_j(x')$ with $x' \neq x$). In order to eliminate such ambiguity in the description it is necessary to introduce the concept of *vector-ket* space.**R4(108)**

Vector-Ket Space ε

The vector-ket space is introduced to represent states in a convenient space of *vectors* $|\phi_j>$, instead of working in the space of *functions* $\phi_j(x)$. The main difference is that the coordinate dependence does not need to be specified when working in the vector-ket space. According to such representation, function $\psi(x)$ is the *component* of vector $|\psi>$ associated with index x (*vide infra)* . Therefore, for any function $\psi(x) = \sum_j C_j\phi_j(x)$, we can define a ket-vector $|\psi>$ such that,

$$|\psi> = \sum_j C_j|\phi_j>.$$

The representation of $|\psi>$ in space ε is,



Note that the expansion coefficients C_j depend only on the kets $|\psi_j>$ and not on any specific vector component. Therefore, the ambiguity mentioned above is removed.

In order to learn how to operate with kets we need to introduce the *bra space* and the concept of *linear functional*. After doing so, this section will be concluded with the description of *Postulate 5*, and the *Continuity Equation*.

Linear functionals

A functional χ is a mathematical operation that transforms a function $\psi(x)$ into a number. This concept is extended to the vector-ket space ε, as an operation that transforms a vector-ket into a number as follows,

$$\chi(\psi(x)) = n, \text{ or } \chi(|\psi>) = n,$$

where n is a number. A *linear* functional satisfies the following equation,

$$\chi(a\psi(x) + bf(x)) = a\chi(\psi(x)) + b\chi(f(x)),$$

where a and b are constants.

Example: The scalar product,**R4(110)**

$$n = \int dx \psi^*(x)\phi(x),$$

is an example of a linear functional, since such an operation transforms a function $\phi(x)$ into a number n. In order to introduce the scalar product of kets, we need to introduce the *bra-space*.

Bra Space ε^*

For every ket $|\psi>$ we define a linear functional $<\psi|$, called *bra-vector*, as follows:

$$<\psi|(|\phi>) = \int dx \psi^*(x)\phi(x).$$

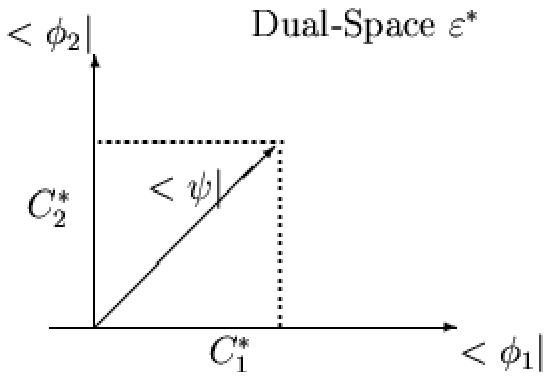
Note that functional $<\psi|$ is linear because the scalar product is a linear functional. Therefore,

$$<\psi|(a|\phi> + b|f>) = a<\psi|(|\phi>) + b<\psi|(|f>).$$

Note: For convenience, we will omit parenthesis so that the notation $<\psi|(|\phi>)$ will be equivalent to $<\psi||\phi>$. Furthermore, whenever we find two bars next to each other we can merge them into a single one without changing the meaning of the expression. Therefore,

$$<\psi||\phi> = <\psi|\phi> .$$

The space of bra-vectors is called dual space ε^* simply because given a ket $|\psi> = \sum_j C_j|\phi_j>$, the corresponding bra-vector is $<\psi| = \sum_j C_j^* <\phi_j|$. In analogy to the ket-space, a bra-vector $<\psi|$ is represented in space ε^* according to the following diagram:



where C_j^* is the projection of $< \psi \mid$ along $< \phi_j \mid$.

Projection Operator and Closure Relation

Given a ket $\mid \psi >$ in a certain basis set $|\phi_j >$,

$$|\psi >= \sum_j C_j |\phi_j >, \tag{1}$$

where $< \phi_k|\phi_j >= \delta_{kj}$,

$$C_j =< \phi_j|\psi > . \tag{2}$$

Substituting Eq. (2) into Eq.(1), we obtain

$$|\psi >= \sum_j |\phi_j >< \phi_j|\psi > . \tag{3}$$

From Eq.(3), it is obvious that

$$\sum_j |\phi_j >< \phi_j| = \hat{1}, \qquad Closure\ Relation$$

where $\hat{1}$ is the identity operator that transforms any ket, or function, into itself.

Note that $\hat{P}_j = |\phi_j >< \phi_j|$ is an operator that transforms any vector $|\psi >$ into a vector pointing in the direction of $|\phi_j >$ with magnitude $< \phi_j|\psi >$. The operator $\hat{P}_j$ is called the *Projection Operator*. It projects $|\phi_j >$ according to,

$$\hat{P}_j|\psi >=< \phi_j|\psi > |\phi_j > .$$

Note that $\hat{P}_j^2 = \hat{P}_j$, where $\hat{P}_j^2 = \hat{P}_j\hat{P}_j$. This is true simply because $< \phi_j|\phi_j >= 1$.

$\boxed{Postulate\ 5}$**:** *The evolution of* $\psi(x,t)$ *in time is described by the* *time dependent Schrodinger equation:*

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = \hat{H}\psi(x,t),$$

where $\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \hat{V}(x)$, *is the operator associated with the total energy of the system,* $E = \frac{p^2}{2m} + V(x)$.

Continuity Equation

Exercise 2*: Prove that*

$$\frac{\partial(\psi^*(x,t)\psi(x,t))}{\partial t} + \frac{\partial}{\partial x}j(x,t) = 0,$$

where

$$j(x,t) = \frac{\hbar}{2mi}\left(\psi^*(x,t)\frac{\partial\psi(x,t)}{\partial x} - \psi(x,t)\frac{\partial\psi^*(x,t)}{\partial x}\right).$$

In general, for higher dimensional problems, the change in time of probability density, $\rho(\mathbf{x},t) = \psi^*(\mathbf{x},t)\psi(\mathbf{x},t)$, *is equal to minus the divergence of the probability flux* $\mathbf{j}$,

$$\frac{\partial \rho(\mathbf{x},t)}{\partial t} = -\nabla \cdot \mathbf{j}.$$

This is the so-called [Continuity Equation.]

Note: *Remember that given a vector field* $\mathbf{j}$*, e.g.,* $\mathbf{j}(x,y,z) = j_1(x,y,z)\hat{\imath} + j_2(x,y,z)\hat{\jmath} + j_3(x,y,z)\hat{k}$*, the divergence of* $\mathbf{j}$ *is defined as the dot product of the ``del'' operator* $\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$ *and vector* $\mathbf{j}$ *as follows:*

$$\nabla \cdot \mathbf{j} = \frac{\partial j_1}{\partial x} + \frac{\partial j_2}{\partial y} + \frac{\partial j_3}{\partial z}.$$
